

(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2004-505612

(P2004-505612A)

(43) 公表日 平成16年2月26日(2004.2.26)

(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	4 B O 2 4
C 1 2 M 1/00	C 1 2 M 1/00 A	4 B O 2 9
C 1 2 Q 1/68	C 1 2 Q 1/68 A	4 B O 6 3

審査請求 未請求 予備審査請求 有 (全 131 頁)

(21) 出願番号	特願2001-573044 (P2001-573044)	(71) 出願人	500185748 ユニバーシティ・オブ・サザン・カリフォルニア
(86) (22) 出願日	平成13年4月2日 (2001.4.2)		
(85) 翻訳文提出日	平成14年9月30日 (2002.9.30)		
(86) 国際出願番号	PCT/US2001/010658		アメリカ合衆国90007-4344カリフォルニア州ロサンゼルス、サウス・ホープ・ストリート3716番、スイート313
(87) 国際公開番号	W02001/075172		
(87) 国際公開日	平成13年10月11日 (2001.10.11)	(74) 代理人	100077517 弁理士 石田 敬
(31) 優先権主張番号	60/193,839	(74) 代理人	100092624 弁理士 鶴田 準一
(32) 優先日	平成12年3月31日 (2000.3.31)	(74) 代理人	100087871 弁理士 福本 積
(33) 優先権主張国	米国 (US)	(74) 代理人	100082898 弁理士 西山 雅也

最終頁に続く

(54) 【発明の名称】 食道腺ガンに関する後成的配列

(57) 【要約】

ガン、その中でも胃腸腺ガンと食道腺ガンの診断法または予後予測法を開示する。特に、本発明により、胃腸腺ガン、食道腺ガン、バレット食道、正常な扁平粘膜上皮において、どの遺伝子が高メチル化され、どの遺伝子がメチル化されていないかなどを調べる CpG 島の標準的なメチル化アッセイによって検査することのできるメチル化パターンが提供される。

【特許請求の範囲】

【請求項1】

組織サンプルからガンまたはガン関連疾患を診断する方法であって、

(a) 診断する検査組織または検査領域から組織サンプルを取得し、

(b) その組織サンプルのメチル化アッセイを行なって、APC、ARF、CALCA、CDH1、CDKN2A、CDKN2B、ESR1、GSTP1、HIC1、MGMT、MLH1、MYOD1、RB1、TGFB2、THBS1、TIMP3、CTNNB1、PTGS2、TYMSおよびMTHFR、ならびにこれらの組み合わせからなるグループの中から選択した少なくとも1つの遺伝子配列内に位置するゲノムCpG配列のメチル化状態を決定し、

(c) 少なくとも部分的に、上記ゲノムCpG配列のメチル化状態に基づいて、ガンの診断または予後予測を行なう操作を含む方法。

10

【請求項2】

APC、ARF、CALCA、CDH1、CDKN2A、CDKN2B、ESR1、GSTP1、HIC1、MGMT、MLH1、MYOD1、RB1、TGFB2、THBS1、TIMP3、CTNNB1、PTGS2およびTYMSからなるグループの中から選択した少なくとも1つの遺伝子配列内に位置するゲノムCpG配列が、CpG島のゲノムCpG配列に対応している、請求項1に記載の方法。

【請求項3】

APC、ARF、CALCA、CDH1、CDKN2A、CDKN2B、ESR1、GSTP1、HIC1、MGMT、MLH1、MYOD1、RB1、TGFB2、THBS1、TIMP3、CTNNB1、PTGS2、TYMSおよびMTHFRの遺伝子配列が、表IIに掲載した配列番号1～60、64および65に対応する特異的オリゴヌクレオチド・プライマーまたはプローブ、あるいはその一部によって規定される配列である、請求項1に記載の方法。

20

【請求項4】

上記CpG島が、APC、ARF、CALCA、CDH1、CDKN2A、CDKN2B、ESR1、GSTP1、HIC1、MGMT、MLH1、MYOD1、RB1、TGFB2、THBS1、TIMP3、CTNNB1、PTGS2およびTYMSの遺伝子のうちの1つまたはそれ以上のプロモーター領域内に存在している、請求項2に記載の方法。

30

【請求項5】

APC、ARF、CALCA、CDH1、CDKN2A、CDKN2B、ESR1、GSTP1、HIC1、MGMT、MLH1、MYOD1、RB1、TGFB2、THBS1、TIMP3、CTNNB1、PTGS2およびTYMSの遺伝子の配列が、表IIに掲載した配列番号1～54、58～60、64および65に対応する特異的オリゴヌクレオチド・プライマーまたはプローブ、あるいはその一部によって規定される配列と関係した任意のCpG島配列に対応しており、関係した上記CpG島配列が、配列番号1～54、58～60、64および65に対応する特異的オリゴヌクレオチド・プライマーおよびプローブによって規定される配列の少なくとも1つのヌクレオチドを含み、CpGジヌクレオチドの頻度が観測値/予測値の比 >0.6 に対応することと、GC含量 >0.5 であることの両方の条件を満たす、ゲノムDNAの隣接配列である、請求項2に記載の方法。

40

【請求項6】

上記ゲノムCpG配列が、APC、CDKN2A、MYOD1、CALCA、ESR1、MGMTおよびTIMP3、ならびにこれらの組み合わせからなるグループの中から選択した少なくとも1つの遺伝子配列内に位置する、請求項1に記載の方法。

【請求項7】

APC、CDKN2A、MYOD1、CALCA、ESR1、MGMTおよびTIMP3からなるグループの中から選択した少なくとも1つの遺伝子配列内に位置する上記ゲノムCpG配列が、CpG島のゲノムCpG配列に対応する、請求項6に記載の方法。

50

【請求項 8】

A P C、C D K N 2 A、M Y O D 1、C A L C A、E S R 1、M G M T および T I M P 3 の遺伝子配列が、それぞれ、表 I I に掲載した配列番号 1 9 ~ 2 1、配列番号 1 ~ 3、配列番号 7 ~ 9、配列番号 1 0 ~ 1 2、配列番号 4 ~ 6、配列番号 1 6 ~ 1 8 および配列番号 1 3 ~ 1 5 に対応する特異的オリゴヌクレオチド・プライマーおよびプローブによって規定される配列である、請求項 6 に記載の方法。

【請求項 9】

上記 C p G 島が、A P C、C D K N 2 A、M Y O D 1、C A L C A、E S R 1、M G M T および T I M P 3 の遺伝子のうちの 1 つまたはそれ以上のプロモーター領域内に存在している、請求項 7 に記載の方法。

10

【請求項 10】

A P C、C D K N 2 A、M Y O D 1、C A L C A、E S R 1、M G M T および T I M P 3 の遺伝子配列が、それぞれ、表 I I に掲載した配列番号 1 9 ~ 2 1、配列番号 1 ~ 3、配列番号 7 ~ 9、配列番号 1 0 ~ 1 2、配列番号 4 ~ 6、配列番号 1 6 ~ 1 8 および配列番号 1 3 ~ 1 5 に対応する特異的オリゴヌクレオチド・プライマーおよびプローブ、またはその一部によって規定される配列と関係した任意の C p G 島配列に対応しており、関係した上記 C p G 島配列が、配列番号 1 9 ~ 2 1、配列番号 1 ~ 3、配列番号 7 ~ 9、配列番号 1 0 ~ 1 2、配列番号 4 ~ 6、配列番号 1 6 ~ 1 8 および配列番号 1 3 ~ 1 5 に対応する特異的オリゴヌクレオチド・プライマーおよびプローブによって規定される配列の少なくとも 1 つのヌクレオチドを含み、C p G ジヌクレオチドの頻度が観測値 / 予測値の比 > 0 . 6 に対応することと、G C 含量 > 0 . 5 であることの両方の条件を満たす、ゲノム D N A 中の隣接配列である、請求項 7 に記載の方法。

20

【請求項 11】

上記のガンまたはガン関連疾患を、胃腸腺ガンまたは食道腺ガン、胃腸異形成または食道異形成、胃腸化成または食道化成、バレット腸組織、正常な食道扁平粘膜上皮の前ガン疾患、およびこれらの組み合わせからなるグループの中から選択する、請求項 1 に記載の方法。

【請求項 12】

上記ガンが食道腺ガンであり、そしてゲノム C p G 配列のメチル化状態に基づいてガンの診断または予後予測をすることにより、その腺ガンの悪性度またはステージを分類する、請求項 11 に記載の方法。

30

【請求項 13】

上記のガンまたはガン関連疾患を、胃腸腺ガンまたは食道腺ガン、胃腸異形成または食道異形成、胃腸化成または食道化成、バレット腸組織、正常な食道扁平粘膜上皮の前ガン疾患、およびこれらの組み合わせからなるグループの中から選択する、請求項 6 に記載の方法。

【請求項 14】

上記ガンが食道腺ガンであり、その場合、ゲノム C p G 配列のメチル化状態に基づいてガンの診断または予後予測をすることにより、その腺ガンの悪性度またはステージを分類する、請求項 13 に記載の方法。

40

【請求項 15】

ゲノム C p G 配列のメチル化状態を明らかにするのに用いるメチル化アッセイを、“メチライト(商標)”、M s - S N u P E、M S P、C O B R A、M C A および D M H、ならびにこれらの組み合わせからなるグループの中から選択する、請求項 2 に記載の方法。

【請求項 16】

ゲノム C p G 配列のメチル化状態を明らかにするのに用いるメチル化アッセイを、“メチライト(商標)”、M s - S N u P E、M S P、C O B R A、M C A および D M H、ならびにこれらの組み合わせからなるグループの中から選択する、請求項 6 に記載の方法。

【請求項 17】

ゲノム C p G 配列のメチル化状態を明らかにするのに用いるメチル化アッセイの少なくとも

50

も一部が、APC、ARF、CALCA、CDH1、CDKN2A、CDKN2B、ESR1、GSTP1、HIC1、MGMT、MLH1、MYOD1、RB1、TGFB2、THBS1、TIMP3、CTNNB1、PTGS2、TYMSおよびMTHFRからなるグループの中から選択した少なくとも1つの遺伝子配列内に位置するCpG含有配列を含んで成るアレイまたはマイクロアレイに基づいている、請求項1に記載の方法。

【請求項18】

APC、ARF、CALCA、CDH1、CDKN2A、CDKN2B、ESR1、GSTP1、HIC1、MGMT、MLH1、MYOD1、RB1、TGFB2、THBS1、TIMP3、CTNNB1、PTGS2およびTYMSの遺伝子配列が、表IIに掲載した配列番号1～54、58～60、64および65に対応する特異的オリゴヌクレオチド・プライマーまたはプローブ、またはその一部によって規定される配列と関係した任意のCpG島配列に対応しており、関係した上記CpG島配列が、配列番号1～54、58～60、64および65に対応する特異的オリゴヌクレオチド・プライマーおよびプローブによって規定される配列の少なくとも1つのヌクレオチドを含み、CpGジヌクレオチドの頻度が観測値/予測値の比 > 0.6 に対応することと、GC含量 > 0.5 であることの両方の条件を満たす、ゲノムDNAの隣接配列である、請求項17に記載の方法。

【請求項19】

APC、ARF、CALCA、CDH1、CDKN2A、CDKN2B、ESR1、GSTP1、HIC1、MGMT、MLH1、MYOD1、RB1、TGFB2、THBS1、TIMP3、CTNNB1、PTGS2、TYMSおよびMTHFRの遺伝子配列が、表IIに掲載した配列番号1～60、64および65に対応する特異的オリゴヌクレオチド・プライマーおよびプローブ、またはその一部によって規定される配列、あるいはこれらに対応する配列である、請求項17に記載の方法。

【請求項20】

決定されるゲノムCpG配列のメチル化状態が、高メチル化、低メチル化または正常なメチル化である、請求項1に記載の方法。

【請求項21】

ガンまたはガン関連疾患を診断するのに役立つキットであって、

(a) APC、ARF、CALCA、CDH1、CDKN2A、CDKN2B、ESR1、GSTP1、HIC1、MGMT、MLH1、MYOD1、RB1、TGFB2、THBS1、TIMP3、CTNNB1、PTGS2、TYMSおよびMTHFRからなるグループの中から選択した少なくとも1つの遺伝子配列内に位置する配列の任意の領域とハイブリダイズするプローブまたはプライマーを含む容器と、

(b) メチル化されたCpGを含む核酸の検出の少なくとも一部を、このプローブまたはプライマーに基づいて行なうのに必要な標準的メチル化アッセイ用添加試薬とを含んで成る1または複数の容器を含む担持手段を有するキット。

【請求項22】

上記標準的メチル化アッセイ用添加試薬が、“メチライト(商標)”、MS-SNuPE、MSP、COBRA、MCAおよびDMH、ならびにこれらの組み合わせからなるグループの中から選択した、メチル化アッセイを行なうための標準的な試薬である、請求項21に記載のキット。

【請求項23】

上記プローブまたはプライマーが、表IIに掲載した配列番号1～60、64および65からなるグループの中から選択した配列の少なくとも約12～15個のヌクレオチドを含んで成る、請求項21に記載のキット。

【請求項24】

ガンまたはガン関連疾患を診断するのに役立つキットであって、

(a) 配列番号1～60、64、65からなるグループの中から選択した配列の少なくとも約12～15個のヌクレオチドからなる配列と、配列番号1～54、58～60、64および65と関係したCpG島配列内に位置する任意の配列とを含んで成るアレイまたは

マイクロアレイ、を含んで成る1または複数の容器を含む担持手段を有するキット。

【発明の詳細な説明】

【0001】

〔発明の属する技術分野〕

関連出願の相互参照

本出願は、「EPIGENETIC SEQUENCES FOR ESOPHAGEAL ADENOCARCINOMA」という名称で2000年3月31日に出願されたアメリカ合衆国仮特許出願番号第06/193,839号に対する優先権を主張する。

【0002】

連邦政府から資金援助を受けた研究であることに関する宣言

この研究は、P.W.L.に対するNIH/NCI助成金R01CA75090の援助を受けた。アメリカ合衆国は、35U.S.C. § 202(c)(6)に従い、本発明において所定の権利を有する。

【0003】

本発明の技術分野

本発明は、胃腸腺ガン、中でも食道腺ガン(“EAC”)の診断的または予後的なアッセイを提供する。さらに詳細には、本発明は、CpG島のメチル化状態を調べる標準的なメチル化アッセイによって検査で、且つ胃腸ガン、正常な扁平上皮細胞、およびEACの中の2つ以上の遺伝子の相対的メチル化状態を含んで成る多遺伝子後成的フィンガープリントまたはメチル化パターンを提供する。

【0004】

本発明の背景

DNAのメチル化とガン。DNAのメチル化パターンは、ヒトのガンにおいてしばしば変化している。これらのメチル化の変化には、ゲノム全体での低メチル化や、部分的な高メチル化が含まれる(JonesとLaird、Nat. Genet.、第21巻、163~167ページ、1999年)。正常な組織ではCpG島は一般にメチル化されないが、ガン細胞における異常な高メチル化は、このCpG島においてしばしば起こる。プロモーターCpG島(すなわち遺伝子のプロモーター領域に位置するCpG島)の高メチル化は、ヒトの多くのタイプのガンにおける転写サイレンシングと関係している。

【0005】

遺伝子のメチル化パターンからは、ガン細胞に関するさまざまな種類の有用な情報を得ることができる。第1に、それぞれのタイプの腫瘍(すなわち乳ガン、結腸ガン、食道ガンなど)には、メチル化されやすい特徴的な遺伝子群がある(Costello他、Nat. Genet.、第24巻、132~138ページ、2000年)。例えばRB1は、網膜芽細胞腫で高メチル化されている(Stirzaker他、Cancer Res.、第57巻、2229~2237ページ、1997年; Sakai他、Am. J. Hum. Genet.、第48巻、880~888ページ、1991年)が、急性骨髄性白血病ではそうになっていない(KornblauとQiu、Leuk. Lymphoma、第35巻、283~288ページ、1999年; Melki他、Cancer Res.、第59巻、3730~3740ページ、1999年)ことが知られている。

【0006】

第2に、一人の患者の個々の腫瘍は、別の患者の同じタイプの腫瘍と比べると、その腫瘍の進行度を反映した独特な後成的フィンガープリントを持っている(Costello他、Nat. Genet.、第24巻、132~138ページ、2000年)。

【0007】

しかし一般には、ガンにおける後成的変化に関するほとんどの研究は、主として、ごく少数の既知の遺伝子群(JonesとLaird、Nat. Genet.、第21巻、163~167ページ、1999年; BaylinとHerman、Trends Genet.、第16巻、168~174ページ、2000年)、あるいは未知のCpG島の全体的分析(Costello他、Nat. Genet.、第24巻、132~138ページ、2000年)が知られている。

10

20

30

40

50

ージ、2000年)に焦点を絞ったものであり、そのため、適切な診断法および/または予後予測法の枠組みは提供されていない。

【0008】

食道腺ガン(“EAC”)。食道腺ガン(“EAC”)は、正常な扁平粘膜上皮が化生によって特殊な柱状上皮(腸上皮化生(IM)またはバレット食道)になり、それが最終的には異形成へと進行し、その後に悪性腫瘍になるという多段階ステップを経て発症する(Barrett他、Nat. Genet.、第22巻、106~109ページ、1999年;Zhuang他、Cancer Res.、第56巻、1961~1964ページ、1996年)。西洋諸国では過去30年間にEACの発生が急速に増えてきている(Devesa他、Cancer、第83巻、2049~2053ページ、1998年;Janowski他、Am. J. Pathol.、第154巻、965~973ページ、1999年)。

10

【0009】

残念なことに、このモデルの後成的研究は、これまでのところ、数個の遺伝子についてのDNAメチル化分析に限られている(Wong他、Cancer Res.、第57巻、2619~2622ページ、1997年;Klump他、Gastroenterology、第115巻、1381~1386ページ、1998年;Eads他、Cancer Res.、第60巻、5021~5026ページ、2000年)。

【0010】

CpG島メチル化の表現型(“CIMP”)。結腸直腸腫瘍と胃腫瘍の一部には、CpG島メチル化のある表現型(“CIMP”)が見られることが以前に報告されており(Toyota他、Proc. Natl. Acad. Sci. USA、第96巻、8681~8686ページ、1999年;Toyota他、Cancer Res.、第59巻、5438~5442ページ、1999年)。この表現型は、1つの腫瘍の多数の遺伝子座に影響を与える広範囲の異常な高メチル化という変化を特徴とする。これは、一群の腫瘍においてメチル化されている遺伝子の数の頻度分布が2つのピークを持つことに反映されている(Toyota他、Proc. Natl. Acad. Sci. USA、第96巻、8681~8686ページ、1999年)。CIMP腫瘍は、はっきりと識別できる一群の腫瘍であり、ガンにおいてだけメチル化される遺伝子、またはタイプCの遺伝子において対応するCpG島が高い割合で高メチル化されていることを特徴としている。CIMPは現在のところ、新規ではっきりと識別できる、しかし主要な腫瘍発生の経路であると考えられている(Toyota他、Proc. Natl. Acad. Sci. USA、第96巻、8681~8686ページ、1999年;Toyota他、Cancer Res.、第59巻、5438~5442ページ、1999年)。

20

30

【0011】

しかし、EACの進行におけるCIMP経路の役割は、もし役割があるとしてのことだが、まだ明らかになっていない。というのも、これまでの後成的研究で分析されたのは、1つの遺伝子(Wong他、Cancer Res.、第57巻、2619~2622ページ、1997年;Klump他、Gastroenterology、第115巻、1381~1386ページ、1998年)、または数個の遺伝子(Eads他、Cancer Res.、第60巻、5021~5026ページ、2000年)だけだからである。

40

【0012】

したがって、ガンの検出、化学的予測および予後予測に関する新規な方法が、当業界で要求されている。ガンなどの病気のさまざまな段階を経て進行する間に多数の遺伝子座で見られるCpG島メチル化の変化の新規なコーディネート・パターンを明らかにすることが、当業界で要求されている。特定の腫瘍に特異的で、しかも個々の患者に特異的な後成的パターンまたはフィンガープリントを明らかにすることが、当業界で要求されている。ガン処置のための診断法および/または予後予測法で利用できるバイオマーカーまたはプローブ、例えばEAC特異的なバイオマーカーまたはプローブが、当業界で要求されている。食道腺ガンでCIMPが見られるかどうかを明らかにすることが、当業界で要求されて

50

いる。腫瘍のステージを決定する新規な方法が、当業界で要求されている。本発明は、これらの要求に対処するものである。

【0013】

本発明の要約

本発明は、組織サンプルからガンまたはガン関連疾患を診断する方法であって、(a) 診断する検査組織または検査領域から組織サンプルを取得し、(b) その組織サンプルのメチル化アッセイを行なって、APC、ARF、CALCA、CDH1、CDKN2A、CDKN2B、ESR1、GSTP1、HIC1、MGMT、MLH1、MYOD1、RB1、TGFB2、THBS1、TIMP3、CTNNB1、PTGS2、TYMSおよびMTHFR、ならびにこれらの組み合わせからなるグループの中から選択した少なくとも1つの遺伝子配列内に位置するゲノムCpG配列のメチル化状態を決定し、(c) 少なくとも部分的に、上記ゲノムCpG配列のメチル化状態に基づいたガンの診断的または予後的な予測を行なうことを含んで成る方法を提供する。APC、ARF、CALCA、CDH1、CDKN2A、CDKN2B、ESR1、GSTP1、HIC1、MGMT、MLH1、MYOD1、RB1、TGFB2、THBS1、TIMP3、CTNNB1、PTGS2およびTYMSからなるグループの中から選択した少なくとも1つの遺伝子配列内に位置するゲノムCpG配列は、CpG島のゲノムCpG配列に対応していることが好ましい。APC、ARF、CALCA、CDH1、CDKN2A、CDKN2B、ESR1、GSTP1、HIC1、MGMT、MLH1、MYOD1、RB1、TGFB2、THBS1、TIMP3、CTNNB1、PTGS2、TYMSおよびMTHFRの遺伝子配列は、表IIに掲載した配列番号1~60、64および65に対応する特異的オリゴヌクレオチド・プライマーまたはプローブ、またはその一部によって規定される配列であることが好ましい。CpG島は、前記遺伝子のプロモーター領域に位置することが好ましい。APC、ARF、CALCA、CDH1、CDKN2A、CDKN2B、ESR1、GSTP1、HIC1、MGMT、MLH1、MYOD1、RB1、TGFB2、THBS1、TIMP3、CTNNB1、PTGS2、およびTYMSの遺伝子配列は、表IIに掲載した配列番号1~54、58~60、64および65に対応する特異的オリゴヌクレオチド・プライマーまたはプローブ、あるいはその一部によって規定される配列と関係した任意のCpG島配列に対応することが好ましい。ここに、関係したCpG島配列とは、配列番号1~54、58~60、64および65に対応する特異的オリゴヌクレオチド・プライマーおよびプローブによって規定される配列の少なくとも1つのヌクレオチドを含み、CpGジヌクレオチドの頻度が観測値/予測値の比 > 0.6 に対応することと、GC含量 > 0.5 であることの両方の条件を満たす、ゲノムDNAの隣接配列である。

【0014】

ゲノムCpG配列は、APC、CDKN2A、MYOD1、CALCA、ESR1、MGMTおよびTIMP3、ならびにこれらの組み合わせからなるグループの中から選択した少なくとも1つの遺伝子配列内に位置することが好ましい。APC、CDKN2A、MYOD1、CALCA、ESR1、MGMT、TIMP3からなるグループの中から選択した少なくとも1つの遺伝子配列内に位置するゲノムCpG配列は、CpG島のゲノムCpG配列に対応することが好ましい。APC、CDKN2A、MYOD1、CALCA、ESR1、MGMTおよびTIMP3の遺伝子配列は、それぞれ、表IIに掲載した配列番号19~21、配列番号1~3、配列番号7~9、配列番号10~12、配列番号4~6、配列番号16~18および配列番号13~15に対応する特異的オリゴヌクレオチド・プライマーおよびプローブによって規定される配列であることが好ましい。CpG島は、遺伝子のプロモーター領域に位置することが好ましい。APC、CDKN2A、MYOD1、CALCA、ESR1、MGMTおよびTIMP3の遺伝子配列は、それぞれ、表IIに掲載した配列番号19~21、配列番号1~3、配列番号7~9、配列番号10~12、配列番号4~6、配列番号16~18および配列番号13~15に対応する特異的オリゴヌクレオチド・プライマーおよびプローブ、またはその一部によって規定される配列と関係した任意のCpG島配列に対応することが好ましい。ここに、関係したCpG島配

10

20

30

40

50

列とは、配列番号19～21、配列番号1～3、配列番号7～9、配列番号10～12、配列番号4～6、配列番号16～18および配列番号13～15に対応する特異的オリゴヌクレオチド・プライマーおよびプローブによって規定される配列の少なくとも1つのヌクレオチドを含み、CpGジヌクレオチドの頻度が観測値/予測値の比 >0.6 に対応することと、GC含量 >0.5 であることの両方の条件を満たす、ゲノムDNAの隣接配列である。

【0015】

好ましくは、ガンまたはガン関連疾患は、胃腸腺ガンまたは食道腺ガン、胃腸異形成または食道異形成、胃腸化成または食道化成、パレット腸組織、正常な食道扁平粘膜上皮の前ガン疾患、およびこれらの組み合わせからなるグループの中から選択される。ガンは食道腺ガンであることが好ましく、その場合、ゲノムCpG配列のメチル化状態に基づいてガンの診断または予後予測をすることにより、その腺ガンの悪性度またはステージを分類することが好ましい。

10

【0016】

ゲノムCpG配列のメチル化状態を明らかにするのに用いるメチル化アッセイは、“メチライト(Methylight:商標)”、Ms-SNuPE、MSP、COBRA、MCA、およびDMH、ならびにこれらの組み合わせからなるグループの中から選択することが好ましい。

【0017】

ゲノムCpG配列のメチル化状態を明らかにするのに用いるメチル化アッセイの少なくとも一部は、APC、ARF、CALCA、CDH1、CDKN2A、CDKN2B、ESR1、GSTP1、HIC1、MGMT、MLH1、MYOD1、RB1、TGFB2、THBS1、TIMP3、CTNNB1、PTGS2、TYMSおよびMTHFRからなるグループの中から選択した少なくとも1つの遺伝子配列内に位置するCpG配列を含んで成るアレイまたはマイクロアレイに基づいていることが好ましい。APC、ARF、CALCA、CDH1、CDKN2A、CDKN2B、ESR1、GSTP1、HIC1、MGMT、MLH1、MYOD1、RB1、TGFB2、THBS1、TIMP3、CTNNB1、PTGS2、およびTYMSの遺伝子配列は、表IIに掲載した配列番号1～54、58～60、64および65に対応する特異的オリゴヌクレオチド・プライマーまたはプローブ、あるいはその一部によって規定される配列と関係した任意のCpG島配列に対応することが好ましい。ここに、関係したCpG島配列とは、配列番号1～54、58～60、64および65に対応する特異的オリゴヌクレオチド・プライマーおよびプローブによって規定される配列の少なくとも1つのヌクレオチドを含み、CpGジヌクレオチドの頻度が観測値/予測値の比 >0.6 に対応することと、GC含量 >0.5 であることの両方の条件を満たす、ゲノムDNAの隣接配列である。APC、ARF、CALCA、CDH1、CDKN2A、CDKN2B、ESR1、GSTP1、HIC1、MGMT、MLH1、MYOD1、RB1、TGFB2、THBS1、TIMP3、CTNNB1、PTGS2、TYMSおよびMTHFRの遺伝子配列は、表IIに掲載した配列番号1～60、64、65に対応する特異的オリゴヌクレオチド・プライマーおよびプローブ、またはその一部によって規定される配列、あるいはこれらに対応する配列であることが好ましい。

20

30

40

【0018】

明らかにされるゲノムCpG配列のメチル化状態は、高メチル化、低メチル化、正常なメチル化のいずれかであることが好ましい。

【0019】

本発明は、ガンまたはガン関連疾患の診断または予後に有用なキットであって、(a) APC、ARF、CALCA、CDH1、CDKN2A、CDKN2B、ESR1、GSTP1、HIC1、MGMT、MLH1、MYOD1、RB1、TGFB2、THBS1、TIMP3、CTNNB1、PTGS2、TYMSおよびMTHFRからなるグループの中から選択した少なくとも1つの遺伝子配列内に位置する配列の任意の領域とハイブリ

50

ダイズするプローブまたはプライマーを含む容器、および (b) メチル化された CpG を含む核酸の検出の少なくとも一部を、このプローブまたはプライマーに基づいて行なうのに必要な追加の標準的メチル化アッセイ試薬、を含んで成る1または複数の容器を備える担持手段を含んで成るキットも提供する。追加の標準的メチル化アッセイ試薬は、“メチライト(商標)”、MS-SNUPE、MSP、COBRA、MCAおよびDMH、ならびにこれらの組み合わせからなるグループの中から選択した、メチル化アッセイを行なうための標準的な試薬であることが好ましい。上記プローブまたはプライマーは、表IIに掲載した配列番号1~60、64および65からなるグループの中から選択した配列の少なくとも約12~15個のヌクレオチドを含んで成ることが好ましい。

【0020】

10

本発明はさらに、ガンまたはガン関連疾患の診断または予後に有用なキットであって、(a) 配列番号1~60、64、65からなるグループの中から選択した配列の少なくとも約12~15個のヌクレオチドからなる配列と、配列番号1~54、58~60、64および65と関係したCpG島配列内に位置する任意の配列とを含んで成る1または複数の容器を備える担持手段を有するキットも提供する。

【0021】

本発明の詳細な説明

定義：

“EAC”という語は食道腺ガンを意味するが、この語には、正常な扁平粘膜上皮が化生によって特殊な柱状上皮(腸上皮化生(IM)またはバレット食道)になり、それがついには異形成へと進行し、その後悪性腫瘍になるという多段階プロセスを経る食道腺ガンのさまざまな組織学的段階も含まれる(Barrett他、Nat. Genet.、第2巻、106~109ページ、1999年;Zhuang他、Cancer Res.、第56巻、1961~1964ページ、1996年)。

20

【0022】

“CIMP”という語はCpG島メチル化の表現型を意味し、1つの腫瘍の多数の遺伝子座に影響を与える広範囲の異常な高メチル化という変化を特徴とする。これは、一群の腫瘍(16)においてメチル化されている遺伝子の数の頻度分布が2つのピークを持つことに反映されている。CIMP腫瘍は、はっきりと識別できる一群の腫瘍であり、ガンにおいてだけメチル化される遺伝子、またはタイプCの遺伝子で対応するCpG島が高い割合で高メチル化されていることを特徴としている。CIMPは現在のところ、新規ではっきりと識別できるしかし主要な腫瘍発生経路であると考えられている(Toyota他、Proc. Natl. Acad. Sci. USA、第96巻、8681~8686ページ、1999年;Toyota他、Cancer Res.、第59巻、5438~5442ページ、1999年)(上記の“背景”を参照のこと)。

30

【0023】

“PMR”という語は、メチル化された基準の割合を意味し、以下の実施例1に記載したようにして計算する。

【0024】

“GC含量”は、特定のDNA配列内の $[(C \text{ 塩基の数} + G \text{ 塩基の数}) / \text{各断片のバンドの長さ}]$ を意味する。

40

【0025】

“観測値/予測値の比”(“O/E比”)は、特定のDNA配列内のCpGジヌクレオチドの頻度を表わし、 $[CpG \text{ 部位の数} / (C \text{ 塩基の数} \times G \text{ 塩基の数})] \times \text{各断片のバンドの長さ}$ に対応する。

【0026】

“CpG島”は、(1) CpGジヌクレオチドの頻度が“観測値/予測値の比” > 0.6 に対応することと、(2) “GC含量” > 0.5 であることの両方の条件を満たす、ゲノムDNAの隣接配列を意味する。CpG島は、一般に、長さが約0.2kb~約1kbであるが、常にそうであるとは限らない。本発明の特定の配列番号の配列と関係するCpG

50

島の配列は、その特定の配列番号の配列の少なくとも1つのヌクレオチドを含むとともに、CpGジヌクレオチドの頻度が観測値/予測値の比 >0.6 に対応することと、GC含量 >0.5 であることの両方の条件を満たす、ゲノムDNAの隣接配列である。

【0027】

“メチル化状態”は、DNA配列内の1つまたは複数のCpGジヌクレオチドに5-メチルシトシン(“5-mCyt”)が存在しているか不在であるかを意味する。

【0028】

“高メチル化”は、テストするDNAサンプルのDNA配列内の1つまたは複数のCpGジヌクレオチドにおける5-mCytの量が、正常な対照DNAサンプル中の対応するCpGジヌクレオチドにおいて見られる5-mCytの量と比べて多いことに対応するメチル化状態を意味する。

10

【0029】

“低メチル化”は、テストするDNAサンプルのDNA配列内の1つまたは複数のCpGジヌクレオチドにおける5-mCytの量が、正常な対照DNAサンプル中の対応するCpGジヌクレオチドにおいて見られる5-mCytの量と比べて少ないことに対応するメチル化状態を意味する。

【0030】

“メチル化アッセイ”は、DNA配列内のCpGジヌクレオチドのメチル化状態を明らかにするための任意のアッセイを意味する。

【0031】

“MS-AP-PCR”(メチル化感受性のある任意のプライマーを用いたポリメラーゼ連鎖反応)は、CpGジヌクレオチドを含むことが非常に確からしい領域に焦点を絞ってCGリッチなプライマーを用いてゲノム全体をスキャンすることのできる公知の方法である。これについては、Gonzalgo他、Cancer Research、第57巻、594~599ページ、1997年に記載されている。

20

【0032】

“メチライト”は、蛍光に基づいたリアルタイムPCR法として公知であり、Eads他、Cancer Res.、第59巻、2302~2306ページ、1999年に記載されている。

【0033】

“Ms-SNuPE”(メチル化感受性単一ヌクレオチドプライマー伸長)は、GonzalgoとJones、Nucleic Acids Res.、第25巻、2529~2531ページ、1997年に記載されている公知のアッセイである。

30

【0034】

“MSP”(メチル化特異的PCR)は、Herman他、Proc. Natl. Acad. Sci. USA、第93巻、9821~9826ページ、1996年とアメリカ合衆国特許第5,786,146号に記載されている公知のメチル化アッセイである。

【0035】

“COBRA”(亜硫酸水素塩と制限酵素を組み合わせた分析)は、XiongとLaird、Nucleic Acids Res.、第25巻、2532~2534ページ、1997年に記載されている公知のメチル化アッセイである。

40

【0036】

“MCA”(メチル化されたCpG島の増幅)は、Toyota他、Cancer Res.、第59巻、2307~2312ページ、1999年とWO 00/26401A1に記載されているメチル化アッセイである。

【0037】

“DMH”(ディファレンシャル・メチル化ハイブリダイゼーション)は、Huang他、Hum. Mol. Genet.、第8巻、459~470ページ、1999年とYan他、Clin. Cancer Res.、第6巻、1432~1438ページ、2000年に記載されている公知のメチル化アッセイである。

50

【0038】

遺伝子と関連参考文献：

“APC”は、大腸腺維症遺伝子である(Eads他、Cancer Res.、第59巻、2302~2306ページ、1999年；Hiltunen他、Int. J. Cancer、第70巻、644~648ページ、1997年)。

【0039】

“ARF”は、P14の細胞周期を調節する腫瘍抑制遺伝子である(Esteller他、Cancer Res.、第60巻、129~133ページ、2000年；RobertsonとJones、Mol. Cell. Biol.、第18巻、6457~6473ページ、1998年)。

10

【0040】

“CALCA”は、カルシトニン遺伝子である(Melki他、Cancer Res.、第59巻、3730~3740ページ、1999年；Hakkarainen他、Int. J. Cancer、第69巻、471~474ページ、1996年)。

【0041】

“CDH1”は、E-カドヘリン遺伝子である(Melki他、Cancer Res.、第59巻、3730~3740ページ、1999年；Ueki他、Cancer Res.、第60巻、1835~1839ページ、2000年)。

【0042】

“CDKN2A”は、P16遺伝子である(JonesとLaird、Nat. Genet.、第21巻、163~167ページ、1999年；Melki他、Cancer Res.、第59巻、3730~3740ページ、1999年；BaylinとHerman、Trends Genet.、第16巻、168~174ページ、2000年；Cameron他、Nat. Genet.、第21巻、103~107ページ、1999年；Ueki他、Cancer Res.、第60巻、1835~1839ページ、2000年)。

20

【0043】

“CDKN2B”は、P15遺伝子である(Melki他、Cancer Res.、第59巻、3730~3740ページ、1999年；Cameron他、Nat. Genet.、第21巻、103~107ページ、1999年)。

30

【0044】

“CTNNB1”は、 β -カテニン遺伝子である。

【0045】

“ESR1”は、エストロゲン受容体遺伝子である(JonesとLaird、Nat. Genet.、第21巻、163~167ページ、1999年；BaylinとHerman、Trends Genet.、第16巻、168~174ページ、2000年)。

【0046】

“GSTP1”は、グルタチオン-S-トランスフェラーゼP1遺伝子である(Melki他、Cancer Res.、第59巻、3730~3740ページ、1999年；Tchou他、Int. J. Oncol.、第16巻、663~676ページ、2000年)。

40

【0047】

“HIC1”は、ガン1において高メチル化されている遺伝子である(Melki他、Cancer Res.、第59巻、3730~3740ページ、1999年；Wales他、Nat. Med.、第1巻、570~577ページ、1995年)。

【0048】

“MGMT”は、O6-メチルグアニン-DNAメチルトランスフェラーゼ遺伝子である(Esteller他、Cancer Res.、第59巻、793~797ページ、1999年)。

50

【0049】

“MLH1”は、Mut Lホモログ1遺伝子である（JonesとLaird、Nat. Genet.、第21巻、163～167ページ、1999年；BaylinとHerman、Trends Genet.、第16巻、168～174ページ、2000年；Cameron他、Nat. Genet.、第21巻、103～107ページ、1999年；Esteller他、Am. J. Pathol.、第155巻、1767～1772ページ、1999年；Ueki他、Cancer Res.、第60巻、1835～1839ページ、2000年）。

【0050】

“MTHFR”は、メチル-テトラヒドロ葉酸レダクターゼ遺伝子である（Pereira他、Oncol. Rep.、第6巻、597～599ページ、1999年）。 10

【0051】

“MYOD1”は、筋原決定基1遺伝子である（Eads他、Cancer Res.、第59巻、2302～2306ページ、1999年；Cheng他、Br. J. Cancer、第75巻、396～402ページ、1997年）。

【0052】

“PTGS2”は、シクロオキシゲナーゼ2遺伝子である（Zimmermann他、Cancer Res.、第59巻、198～204ページ、1999年）。

【0053】

“RB1”は、網膜芽細胞腫遺伝子である（Stirzaker他、Cancer Res.、第57巻、2229～2237ページ、1997年；Sakai他、Am. J. Hum. Genet.、第48巻、880～888ページ、1991年）。 20

【0054】

“TGFBR2”は、トランスフォーミング増殖因子受容体II遺伝子である（Kang他、Oncogene、第18巻、7280～7286ページ、1999年；Hougaard他、Br. J. Cancer、第79巻、1005～1011ページ、1999年）。

【0055】

“THBS1”は、トロンボスポンジン1遺伝子である（Ueki他、Cancer Res.、第60巻、1835～1839ページ、2000年；Li他、Oncogene 30、第18巻、284～289ページ、1999年）。

【0056】

“TIMP3”は、メタロプロテアーゼ3組織インヒビター遺伝子である（Cameron他、Nat. Genet.、第21巻、103～107ページ、1999年；Ueki他、Cancer Res.、第60巻、1835～1839ページ、2000年；Bachman他、Cancer Res.、第59巻、798～802ページ、1999年）。

【0057】

“TYMS1”は、チミジル酸シンターゼ遺伝子である（L. Herrera編、『家族性大腸腺種症』の中のSakamoto他、315～324ページ、ニューヨーク：アラン・R. リス社、1990年）。 40

【0058】

概要

本発明は、多彩な遺伝子へのアプローチを含んでおり、遺伝子間で共通して見られるメチル化の状態に関して新規で治療に役に立つ見通しを与える。本発明の特別な実施態様では、食道腺ガン（EAC）の異なる組織学的ステージにおける新規な後成的フィンガープリントを提供する。

【0059】

より詳細には、本発明は、ターゲットを絞った方法と総合的な方法の両方の利点を合わせ持っており、定量的ハイスループット・メチル化アッセイである“メチライト（商標）” 50

(Eads 他、Cancer Res.、第59巻、2302~2306ページ、1999年；Eads 他、Cancer Res.、第60巻、5021~5026ページ、2000年；Eads 他、Nucleic Acids Res.、第28巻、E32ページ、2000年)を利用して異なる20種類の遺伝子を分析することにより、(i)食道腺ガン(EAC)におけるメチル化の変化の特徴をより詳細に明らかにすること；(ii)EACの異なる組織学的ステージについて後成的フィンガープリントを作ること；(iii)病気の診断および予防に有効な後成的バイオマーカーを同定すること；(iv)CI-MPが食道腺ガンの腫瘍発生に関与しているかどうかを明らかにすることを実現する。

【0060】

バレット食道および/またはそれに付随する腺ガンのステージが異なる51人の患者から採取した合計で104個の組織サンプルを分析した。中でも、これら組織サンプルのうちの84個につき20種類の遺伝子すべてのスクリーニングを行なったところ、異なる組織においてメチル化パターンがはっきりと異なることが明らかになった。

10

【0061】

本発明の目的にとって最も多くの情報をもたらした遺伝子は、有意な高メチル化の頻度が中間的な値の遺伝子である(すなわちサンプルの約15%(CDKN2A)から約60%(MGMT)の範囲の遺伝子)。このような遺伝子群は、さらに細かい3つのクラスに分類することができた。すなわち、(1)正常な食道粘膜および胃におけるメチル化がない(CDKN2A、ESR1、MYOD1)、(2)正常な食道粘膜および胃においてメチル化がある(CALCA、MGMT、TIMP3)、(3)正常な食道粘膜でメチル化があまり頻繁には起こっていないが、正常な胃のサンプルではすべてメチル化が起きている(APC)という分類である。

20

【0062】

他の遺伝子は、これよりも情報が少なかった。というのも、組織のタイプに関係なく、高メチル化の頻度が約5%未満だった(ARF、CDH1、CDKN2B、GSTP1、MLH1、PTGS2、THBS1)か、高メチル化がまったくない(CTNNB1、RB1、TGFB2、TYMS1)か、高メチル化が常に見られた(HIC1、MTHFR)からである。

【0063】

それぞれの遺伝子クラスは、EACの異なる進行段階において独特の後成的変化をしている。これは、CpG島が高メチル化されることに対する多重保護障壁が段階的に失われることと合致している。異常な高メチル化は、同じ組織の異なる多数の遺伝子座で起こる。これは、EACの腫瘍発生においてメチル化の制御が全体的に異常になることと合致している。しかし、CpG島メチル化の表現型("CI-MP")が見られる腫瘍のグループを識別できるというはっきりとした証拠はなかった。

30

【0064】

さらに、異形成またはガンを伴っているという証拠のある患者から採取した正常組織および化生組織では、病気がもはや進行していない患者から採取した同様の組織におけるよりも高メチル化の程度が有意に大きかった。これら2つのグループの患者から採取したサンプルを組織学的に識別することはできないが、分子レベルでは識別できるという事実は、本発明により、そのような高メチル化の発生が、前ガン状態のバレット食道患者の中から病気がさらに進行するリスクを有する患者を同定するための臨床上的新規かつ有用なツールとなることを示唆している。

40

【0065】

表Iは、EACに関してメチライト(商標)アッセイにより分析した遺伝子の名称と機能のリストである。遺伝子は、HUGO(ヒトゲノム国際機構)によって与えられた名称に基づき、アルファベット順に掲載してある。遺伝子は、CpG島を有するかどうか、他の腫瘍においてメチル化することが知られているかどうかに従い、3つのグループに分類した。各遺伝子の機能に関する簡単な説明も記載してある。

【表1】

50

表 1. EACにおいてメチライト（商標）アッセイによって分析した遺伝子名のリスト

遺伝子の記号	名称	高メチル化された腫瘍*	提案されている機能
CpG 島：腫瘍においてメチル化されている			
APC	大腸腺種症	結腸ガン	WNT シグナル伝達経路； β -カテニンの分解；腫瘍の抑制
ARF	P14	結腸ガン、リンパ腫	細胞周期の調節；腫瘍の抑制

【表 2】

表1. のつづき

CALCA	カルシトニン	乳ガン、結腸ガン、白血病	アデニル酸シクラーゼを通じたカルシウム・レベルの調節
CDH1	E-カドヘリン	急性骨髄性白血病、膀胱ガン、乳ガン、結腸ガン、胃ガン、甲状腺ガン	細胞接着
CDKN2A	P16	急性骨髄性白血病、膀胱ガン、結腸ガン、胃ガン、リンパ腫、黒色腫	細胞周期の調節；腫瘍の抑制
CDKN2B	P15	結腸ガン、血液悪性腫瘍	細胞周期の調節；腫瘍の抑制
ESR1	エストロゲン受容体 α	急性骨髄性白血病、結腸ガン、乳ガン、肺ガン、白血病	哺乳類の細胞におけるホルモン受容体；直腸ガン細胞における成長抑制
GSTP1	グルタチオン-S-トランスフェラーゼPI	乳ガン、前立腺ガン、肝細胞ガン	酸化剤と求電子性発ガン物質からの保護
H1C1	ガン1における高メチル化	脳腫瘍、乳ガン、結腸ガン、腎臓ガン、白血病、リンパ腫	ジンクフィンガー転写因子；腫瘍抑制の可能性
MGMT	O6-メチルグアニン-DNAメチルトランスフェラーゼ	脳腫瘍、結腸ガン、リンパ腫、肺非小細胞ガン	DNA修復
MLH1	Mut Lホモログ1	結腸ガン、子宮内膜ガン、胃ガン	DNAミスマッチ修復
MYOD1	筋原決定基	乳ガン、結腸ガン、卵巣ガン	筋肉の分化
RB1	網膜芽細胞腫1	網膜芽細胞腫	細胞周期の調節；腫瘍の抑制
TGFBR2	トランスフォーミング増殖因子 β 受容体II	結腸ガン、胃ガン、肺非小細胞ガン	セリン-トレオニンキナーゼ受容体；細胞のシグナル伝達
THBS1	トロンボスポンジン1	結腸ガン、グリオブラストーマ	血管新生の抑制

10

20

30

40

【表3】

表I. のつづき

TIMP3	メタロプロテアーゼ3組織インヒビター	脳腫瘍、乳ガン、結腸ガン、腎臓ガン、肺ガン、膵臓ガン	転移、血管新生
CpG 島：メチル化の状態が未知			
CTNNB1	β -カテニン	未知	WNT シグナル伝達経路 (APC によって調節されている) ; E-カドヘリンを通じた細胞接着
PTGS2	シクロオキシゲナーゼ2	未知	アラキドン酸からのプロスタグランジン合成
TYMS	チミジル酸シンターゼ	未知	ヌクレオチド/DNA の合成
非 CpG 島：メチル化の状態が未知			
MTHFR	メチル-テトラヒドロ葉酸レダクターゼ	未知	葉酸塩の代謝、メチオニンの合成 ; 結腸ガンにおいて多型を発生させやすくする

*上記の“定義”に記載した個々の遺伝子に関する参考文献を参照のこと。

10

20

【0066】

ガンの診断法および予後予測法

本発明は、ここに示した20種類の遺伝子配列 (APC、ARF、CALCA、CDH1、CDKN2A、CDKN2B、ESR1、GSTP1、HIC1、MGMT、MLH1、MYOD1、RB1、TGFB2、THBS1、TIMP3、CTNNB1、PTGS2、TYMS、MTHFR ; 表Iと表IIを参照のこと ; 上記の“定義”を参照のこと) のうちの1または複数のメチル化状態、またはメチル化によってこれら遺伝子において変化したDNA配列を明らかにすることに基づいたガンの診断アッセイおよび予後予測アッセイを提供する。20種類あるこれら遺伝子配列領域は、配列番号1~60、64、65 (以下の表IIを参照のこと) に対応するオリゴマー・プライマーおよびプローブによって規定される。配列番号61~63は、この分析で用いるACTB“対照”遺伝子領域に対応する (以下の実施例1を参照のこと)。

30

【0067】

さらに、20種類あるこれら遺伝子配列領域のうちの19種類は、その遺伝子の (GC含量とO/E比に基づいた) CpG島またはCpG領域に対応する ; 19種類の遺伝子とは、APC、ARF、CALCA、CDH1、CDKN2A、CDKN2B、ESR1、GSTP1、HIC1、MGMT、MLH1、MYOD1、RB1、TGFB2、THBS1、TIMP3、CTNNB1、PTGS2、TYMSである (上記の表Iを参照のこと)。したがって、所定のCpG島の一部におけるメチル化状態が一般にその島全体を代表しているという事実に基づくと、本発明にはさらに、(ガンの診断および予後予測への応用においては、配列番号1~60、64、65に対応するプライマーおよびプローブ (以下の表IIを参照のこと) によって規定される) これら19種類の遺伝子配列領域と関係する19の完全なCpG島の中にある任意の配列の新規な使用方法も含まれている。ここに、これら19種類の遺伝子配列と関係するCpG島配列は、これら19種類の遺伝子配列のうちの1つの少なくとも1つのヌクレオチドを含み、CpGジヌクレオチドの頻度が

40

50

観測値 / 予測値の比 > 0.6 に対応することと、GC 含量 > 0.5 であることの両方の条件を満たす、ゲノム DNA 中の連続配列である。

【0068】

一般に、このようなアッセイには、検査組織から組織サンプルを採取し、その組織サンプルに由来する DNA についてメチル化アッセイを行なってその DNA に付随するメチル化状態を明らかにし、その結果に基づいて診断または予後予測をする操作が含まれる。

【0069】

メチル化アッセイは、DNA サンプルの DNA 配列内の 1 つまたは複数の CpG ジヌクレオチドのメチル化状態を明らかにするのに用いる。本発明によれば、可能なメチル化状態としては、正常な状態（すなわちガンでない対照の状態）と比べて高メチル化の状態と低メチル化の状態がある。高メチル化と低メチル化は、検査サンプルの DNA 配列内の 1 つまたは複数の CpG ジヌクレオチドにおける 5 - メチルシトシン（“5 - m C y t”）の量が、正常な対照 DNA サンプル中の対応する CpG ジヌクレオチドにおいて見られる 5 - m C y t の量と比べてそれぞれ多いこと、少ないことに対応するメチル化状態を意味する。

10

【0070】

診断または予後予測の少なくとも一部は、ガンでない正常な組織から得られた対照データと比較することによって明らかになる、サンプルの DNA 配列のメチル化状態に基づく。

【0071】

メチル化アッセイの方法

メチル化アッセイの方法はさまざまなものが従来技術で知られており、本発明でも利用することができる。これらアッセイにより、DNA 配列内の 1 つまたは複数の CpG ジヌクレオチド（例えば CpG 島）のメチル化状態を明らかにすることができる。このようなアッセイとしては、特に、亜硫酸水素塩で処理した DNA のシーケンシング、（配列特異的な増幅を行なうための）PCR、サザンブロット分析、メチル化感受性制限酵素の利用などの方法が挙げられる。

20

【0072】

例えば DNA のメチル化パターンと 5 - メチルシトシンの分布を分析するためのゲノム・シーケンシングは、亜硫酸水素塩処理を行なうことにより簡単になった（Fromme 他、Proc. Natl. Acad. Sci. USA、第 89 巻、1827 ~ 1831 ページ、1992 年）。さらに、亜硫酸水素塩を使用して変換した DNA から増幅した PCR 産物を制限酵素で消化させる。この方法は、例えばサルディとホーンズビー（Nucl. Acids Res.、第 24 巻、5058 ~ 5059 ページ、1996 年）によって、あるいは COBRA（亜硫酸水素塩と制限酵素を組み合わせた分析）（Xiong と Laird、Nucleic Acids Res.、第 25 巻、2532 ~ 2534 ページ、1997 年）として記載されている。

30

【0073】

“メチライト（商標）”（蛍光に基づいたリアルタイム PCR 法）（Eads 他、Cancer Res.、第 59 巻、2302 ~ 2306 ページ、1999 年）、メチル化感受性単一ヌクレオチドプライマー伸長（“Ms - SnuPE”；Gonzalogo と Jones、Nucleic Acids Res.、第 25 巻、2529 ~ 2531 ページ、1997 年）、メチル化特異的 PCR（“MSP”；Herman 他、Proc. Natl. Acad. Sci. USA、第 93 巻、9821 ~ 9826 ページ、1996 年とアメリカ合衆国特許第 5,786,146 号）、メチル化された CpG 島の増幅（“MCA”；Toyota 他、Cancer Res.、第 59 巻、2307 ~ 2312 ページ、1999 年）などのアッセイは、単独で、あるいは他の方法と組み合わせて利用することが好ましい。本発明のさまざまな実施態様で用いることのできるメチル化アッセイとしては、以下のアッセイが挙げられるが、これだけに限定されるわけではない。

40

【0074】

COBRA（亜硫酸水素塩と制限酵素を組み合わせた分析）。COBRA 分析は、少量の

50

ゲノムDNA中の特定の遺伝子座におけるDNAのメチル化レベルを明らかにするのに有用な定量的メチル化アッセイである(XiongとLaird、Nucleic Acids Res.、第25巻、2532~2534ページ、1997年)。一言で述べるならば、制限酵素による消化を利用して、亜硫酸水素ナトリウムで処理したDNAのPCR産物中にある、メチル化に依存した配列の違いを明らかにする。メチル化に依存した配列の違いは、フロマー他(Proc. Natl. Acad. Sci. USA、第89巻、1827~1831ページ、1992年)によって記載された方法に従って標準的な亜硫酸水素塩で処理することにより、ゲノムDNAにまず最初に導入される。次に、亜硫酸水素塩で変換したこのDNAを、興味の対象であるCpG島に特異的なプライマーを用いてPCR増幅した後、制限エンドヌクレアーゼで消化させ、電気泳動をかけ、標識された特異的ハイブリダイゼーション・プローブを用いて検出する。元のDNAサンプル中のメチル化レベルは、消化されたPCR産物と消化されなかったPCR産物の相対量により、幅広いDNAメチル化レベルにわたって定量的に一次式で表現される。さらに、この方法は、顕微解剖したパラフィン包埋組織サンプルから得られるDNAに安心して適用することができる。COBRA分析のための(例えばCOBRAに基づいた典型的なメチル化キットに見られるような)一般的な試薬としては、特定の遺伝子のためのPCRプライマー(またはメチル化によって変化したDNA配列またはCpG島);制限酵素と適切な緩衝液;遺伝子ハイブリダイゼーション用オリゴ;対照ハイブリダイゼーション用オリゴ;オリゴ・プローブのためのキナーゼ標識キット;放射性ヌクレオチド(蛍光法やリン光法など、公知の他の標識法も用いることができる)などが挙げられるが、これだけに限定されるわけではない。さらに、亜硫酸水素塩による変換のための試薬としては、DNA変性用緩衝液;スルホン化用緩衝液;DNA再生用の試薬またはキット(例えば、沈殿、限外濾過、アフィニティ・カラム);脱スルホン化用緩衝液;DNA再生成分などが挙げられる。

【0075】

Ms-SnuPE(メチル化感受性単一ヌクレオチドプライマー伸長)。Ms-SnuPE法は、DNAを亜硫酸水素塩で処理した後、単一ヌクレオチド・プライマーにより伸長させることに基づいて特定のCpG部位におけるメチル化の違いを評価する定量的方法である(GonzalgoとJones、Nucleic Acids Res.、第25巻、2529~2531ページ、1997年)。一言で述べるならば、ゲノムDNAを亜硫酸水素ナトリウムと反応させてメチル化されていないシトシンをウラシルに変換し、その一方で5-メチルシトシンは変化させずにそのままにしておく。次に、亜硫酸水素塩で変換したDNAに対して特異的なPCRプライマーを用いて望む標的配列を増幅し、得られた産物を単離し、それを興味の対象であるCpG部位におけるメチル化分析のための鋳型として用いる。少量のDNA(例えば顕微解剖した病理切片)を分析することができるため、CpG部位におけるメチル化状態を明らかにするのに制限酵素を利用せずにすむ。(例えばMs-SnuPEに基づいた典型的なメチル化キットに見られるような)Ms-SnuPE分析のための典型的な試薬としては、特定の遺伝子用のPCRプライマー(またはメチル化によって変化したDNA配列またはCpG島);最適化されたPCR緩衝液とデオキシヌクレオチド;ゲル抽出キット;正の制御用プライマー;特定の遺伝子用のMs-SnuPEプライマー;(Ms-SnuPE反応用の)反応緩衝液;放射性ヌクレオチドなどが挙げられるが、これだけに限定されるわけではない。さらに、亜硫酸水素塩による変換のための試薬としては、DNA変性用緩衝液;スルホン化用緩衝液;DNA再生のための試薬またはキット(例えば、沈殿、超濾過、アフィニティ・カラム);脱スルホン化用緩衝液;DNA再生成分などが挙げられる。

【0076】

MSP(メチル化特異的PCR)。MSPにより、メチル化感受性制限酵素を使用するかどうかには関係なく、CpG島内のほとんどすべてのCpG部位群のメチル化状態を評価することができる(Herman他、Proc. Natl. Acad. Sci. USA、第93巻、9821~9826ページ、1996年とアメリカ合衆国特許第5,78

6, 146号)。一言で述べるならば、メチル化されているシトシンを除き、亜硫酸水素ナトリウムを用いてメチル化されていないすべてのシトシンをウラシルに変換することによりDNAを変化させた後、メチル化されていないDNAではなくメチル化されたDNAに対して特異的なプライマーを用いて増幅する。MSPは、ほんの少量のDNAしか必要とせず、所定のCpG島遺伝子座の0.1%メチル化された対立遺伝子に対して感受性があり、パラフィン包埋サンプルから抽出したDNAに対して実行することができる。(例えばMSPに基づいた典型的なキットに見られるような)MSP分析のための典型的な試薬としては、特定の遺伝子用のメチル化された、またはメチル化されていないPCRプライマー(またはメチル化によって変化したDNA配列またはCpG島);最適化されたPCR緩衝液とデオキシヌクレオチド;特異的なプローブなどが挙げられるが、これだけに限定されるわけではない。

10

【0077】

MCA(メチル化されたCpG島の増幅)。MCA法は、ゲノムDNA中の変化したメチル化パターンをスクリーニングし、その変化に関係した特別な配列を単離するのに用いることのできる方法である(Toyota他、Cancer Res., 第59巻、2307~2312ページ、1999年)。一言で述べるならば、シトシンのメチル化に対する認識部位での感受性が異なる複数の制限酵素を用い、原発腫瘍、細胞系、正常な組織からのゲノムDNAを消化させた後、AP-PCR増幅を行なう。PCR産物を高分解能ポリアクリルアミドゲル上で分離した後、メチル化の程度が異なる断片をクローニングし、配列を決定する。次に、クローニングした断片をプローブとして用いてサザンブロット分析を行ない、これら断片のメチル化の違いを確認する。(例えばMCAに基づいた典型的なキットに見られるような)MCA分析のための典型的な試薬としては、ゲノムDNAに自由に結合するPCRプライマー;PCR緩衝液とヌクレオチド、制限酵素と適切な緩衝液;遺伝子ハイブリダイゼーション用のオリゴまたはプローブ;対照ハイブリダイゼーション用のオリゴまたはプローブが挙げられるが、これだけに限定されるわけではない。

20

【0078】

DMH(ディファレンシャル・メチル化ハイブリダイゼーション)。DMHは、アレイに基づいた公知のメチル化アッセイであり、ファン他、Hum. Mol. Genet., 第8巻、459~470ページ、1999年とヤン他、Clin. Cancer Res., 第6巻、1432~1438ページ、2000年に記載されている。DMHにより、ガン細胞系のCpG島の高メチル化をゲノム全体でスクリーニングすることができる。一言で述べるならば、CpG島のタグを固体支持体(例えばナイロン膜、シリコンなど)上に配置し、検査サンプル(例えば腫瘍)または基準サンプルからのメチル化されたCpG DNAである“アンプリコン”を用いて調べる。スクリーニングされたCpG島アレイに関して検査サンプルと基準サンプルの間に見られる信号強度の差は、検査用DNA中の対応する配列がメチル化により変化したことを反映している。

30

【0079】

メチライト(商標)。好ましい実施態様では、メチライト(商標)アッセイを用いて1または複数のCpG配列のメチル化状態を明らかにする。メチライト(商標)アッセイは、PCRステップの後にさらに操作を行なう必要のない、蛍光に基づいたリアルタイムPCR(タックマン(商標))法を利用したハイスループットの定量的メチル化アッセイである(Eads他、Cancer Res., 第60巻、5021~5026ページ、2000年;Eads他、Cancer Res., 第59巻、2302~2306ページ、1999年;Eads他、Nucleic Acids Res., 第28巻、E32ページ、2000年)。一言で述べるならば、メチライト(商標)法は、ゲノムDNAの混合サンプルを用意し、それを標準的な方法に従って亜硫酸水素ナトリウムと反応させて変換し(亜硫酸水素塩プロセスによって、メチル化されていないシトシン残基がウラシルに変換される)、メチル化に依存した違いを有する配列が混合したプールにする。次に、蛍光に基づいたPCRを、“偏りのない”(既知のCpGメチル化部位と重なっていないプライマーを用いた)PCR反応として、または“偏りのある”(既知のCpGジヌクレオチ

40

50

ドと重なったPCRプライマーを用いた)反応として行なう。配列の識別は、増幅プロセスまたは蛍光検出プロセスにおいて、あるいはその両方においてなされる可能性がある。

【0080】

メチライト(商標)アッセイは、ゲノムDNAサンプル中のメチル化パターンの定量的検査に用いることができる。この場合には、配列の識別は、プローブがハイブリダイゼーションするときになされる。この定量的方法では、PCR反応により、特定の推定メチル化部位と重なる蛍光プローブの存在下で、偏りのない増幅が行なわれる。入れたDNAに対する偏りのない対照は、プライマーとプローブのいずれもCpGジヌクレオチドとまったく重なっていない反応によって与えられる。また、ゲノムのメチル化に関する定性的検査は、偏りのあるPCRプールを、既知のメチル化部位を“カバー”していない対照オリゴヌクレオチド(蛍光に基づいた“MSP”法)と結合させることによって、またはメチル化部位の可能性のある部位をカバーしているオリゴヌクレオチドと結合させることによって実現される。

10

【0081】

メチライト(商標)プロセスは、増幅プロセスにおいて“タックマン(TaqMan:商標)”プローブとともに用いることができる。例えば、二本鎖ゲノムDNAを亜硫酸水素ナトリウムで処理し、次いでタックマン(商標)プローブを用いて2通りあるPCR反応の一方を行なう;例えば、偏りのあるプライマーとタックマン(商標)プローブを用いるか、あるいは偏りのないプライマーとタックマン(商標)プローブを用いる。タックマン(商標)プローブは、蛍光性“レポーター”分子と“クエンチャー”分子で二重に標識されており、比較的GC含量の多い領域に対して特異的になるよう設計されているため、PCRサイクルにおいて順プライマーまたは逆プライマーよりも約10 高い温度で溶解する。そのため、タックマン(商標)プローブは、PCRのアニーリング/伸長ステップを通じ、完全にハイブリダイズしたままになる。Taqポリメラーゼという酵素はPCRの間に新しい鎖を合成するため、場合によってはアニールされたタックマン(商標)プローブに達することがある。次に、Taqポリメラーゼの5' 3'エンドヌクレアーゼ活性により、タックマン(商標)プローブが消化されて蛍光性レポーター分子が放出される。そのため、リアルタイム蛍光検出システムを用いて、今やクエンチされていない信号を定量的に検出することができる。

20

【0082】

(例えばメチライト(商標)に基づいた典型的なメチル化キットに見られるような)メチライト(商標)分析のための典型的な試薬としては、特定の遺伝子(またはメチル化によって変化したDNA配列またはCpG島)のためのPCRプライマー;タックマン(商標)プローブ;最適化されたPCR緩衝液とデオキシヌクレオチド;Taqポリメラーゼなどが挙げられるが、これだけに限定されるわけではない。メチライト(商標)アッセイの異なる4つの方法(“A”から“D”)に関しては、後に詳しく説明する。定量的メチライト(商標)プロセス“B”を利用することが好ましい。

30

【0083】

メチル化された核酸をメチライト(商標)に基づいて検出することは比較的迅速にでき、その検出は、特異的オリゴヌクレオチド・プローブが増幅により置換されることに基づいている。好ましい実施態様では、増幅と検出は、実際には、特異的な二重標識タックマン(商標)オリゴヌクレオチド・プローブを用いて蛍光に基づいたリアルタイム定量的PCR(“RT-PCR”)により測定を行なうとき、同時に行なわれる。その後何らかの操作をすることも分析することも必要ない。この置換可能なプローブは、修飾されていない元の核酸サンプル中に存在するメチル化されたCpG部位とメチル化されていないCpG部位を識別できるように特別な設計にすることができる。

40

【0084】

メチライト(商標)は、メチル化特異的PCR法(“MSP”;アメリカ合衆国特許第5,786,146号)と同様、メチル化パターンを明らかにする上で、従来のPCRに基づいた方法やその他の方法(例えばサザンブロット分析)よりも有意に優れている。メチ

50

ライト（商標）は、サザンブロット分析よりも実質的に高感度であり、メチライト（商標）を用いると、非常に少数の核酸サンプル中やパフィン包埋サンプル中のメチル化された少数（少ない割合）の対立遺伝子を容易に検出できる。さらに、ゲノムDNAの場合には、分析できるのがメチル化感受性制限エンドヌクレアーゼによって認識されるDNA配列だけではないため、CpGリッチな領域のより広い範囲にわたってメチル化パターンを細かくマッピングすることができる。メチライト（商標）を使用しなかったならば、従来のPCRに基づいたメチル化法では不可避なメチル化感受性制限酵素による不完全な消化の結果として偽の結果が生じる可能性があるが、メチライト（商標）を使用すると、偽の結果をすべて排除することができる。

【0085】

10

メチライト（商標）は、メチル化の量を測定する定量的方法として利用することができ、しかも他の方法よりも実質的に早い。メチライト（商標）は、PCR後の操作または処理をまったく必要としない。そのため、亜硫酸水素塩で処理したDNAの分析に関係する仕事を大きく減らせるだけでなく、PCR産物を取り扱うことで後に実施する反応を汚染させる可能性を避けることもできる。

【0086】

一実施態様では、メチライト（商標）を利用することにより、修飾されていない元のDNA配列内のどのCpG配列もカバーしないプライマーを用いて、可能なあらゆるメチル化状態を偏りなく増幅する。すべてのメチル化パターンが同等に増幅されるからには、DNAのメチル化パターンに関する定量的情報は、配列の違いを検出することのできる任意の方法（例えば蛍光に基づいたPCR）によって、得られたPCRプールから取り出されることになる。

20

【0087】

メチライト（商標）では、1つまたは複数のCpG特異的タックマン（商標）プローブを利用しており、それぞれのタックマン（商標）プローブが、所定の増幅されたDNA領域中の特定のメチル化部位に対応している。次に、変化した単一のDNAサンプルについての複数のアリコートを用いた並列増幅反応においてこのプローブ群を利用し、修飾されていない元のゲノムDNAサンプル中に存在しているすべてのメチル化パターンを同時に明らかにする。これは、ゲノムDNAサンプルを直接シーケンシングするよりも少ない時間と費用で実現され、しかも実質的に感受性がより大きい。さらに、メチライト（商標）の一実施態様では、このようなメチル化パターンを定量的に評価することができる。

30

【0088】

本発明は、この明細書に記載してあるように、さまざまなメチル化アッセイを利用して実施することができる。メチライト（商標）の実施態様に関しては、メチル化に依存した核酸修飾試薬（例えば亜硫酸水素塩）を用いて核酸サンプル（例えばゲノムDNAサンプル）中のCpGのメチル化状態を定性的かつ定量的に明らかにする4つの方法と、その方法に関係した診断キットがある。これら4つの方法をこの明細書ではプロセス“A”、プロセス“B”、プロセス“C”、プロセス“D”として説明する。結局、メチル化されたCpG配列の識別が、増幅段階またはプローブのハイブリダイゼーション段階で、あるいはその両方でなされるように設計する。例えばプロセスCとDでは、修飾されたメチル化されていない核酸と修飾されたメチル化された核酸を識別できる“偏りのある”プライマーを用いることで、PCR増幅の段階でメチル化されたCpG配列を識別する。プロセスBでは、修飾された核酸を偏りなしに増幅するために（CpGメチル化部位をカバーしていない）“偏りのない”プライマーを用いるが、プローブとしては、修飾されたメチル化されていない核酸と修飾されたメチル化された核酸を識別できるプローブを用い、検出段階で（例えば蛍光（または発光）プローブのハイブリダイゼーション段階でだけ）メチル化されたCpG配列を定量的に識別する。プロセスAそれ自体は、増幅段階でも検出段階でもメチル化されたCpG配列を識別しないが、入れたDNAに対する対照反応を提供することによって他の3つの方法の裏付けを与えてそれらの方法を意味あるものにする。

40

【0089】

50

メチライト（商標）プロセスD。本発明によるメチライト（商標）の第1の実施態様では、メチル化されたCpGを含む核酸を定性的に検出する方法であって、核酸を含むサンプルを、メチル化されていないシトシンを修飾する修飾剤と接触させて、変換された核酸を生成させ；特異的オリゴヌクレオチド・ハイブリダイゼーション・プローブの存在下で、2つのオリゴヌクレオチド・プライマーを用いてこの変換された核酸を増幅し（プライマーとプローブの両方とも、修飾されたメチル化されていない核酸と修飾されたメチル化された核酸を識別できる）；増幅によるプローブの置換に基づいて、“メチル化された”核酸を検出する操作を含む方法が提供される。

【0090】

この明細書で用いる“修飾する”という用語は、修飾剤を用いてメチル化されていないシトシンを別のヌクレオチドに変換することを意味する。この変換によって、元の核酸サンプル中でメチル化されていないシトシンがメチル化されたシトシンから識別される。修飾剤は、メチル化されていないシトシンをウラシルに変化させることが好ましい。メチル化されていないシトシンを変化させるのに用いる修飾剤は亜硫酸水素ナトリウムであることが好ましいが、メチル化されていないシトシンを選択的に変化させるがメチル化されたシトシンは変化させない他の等価な修飾剤も、本発明の方法で代わりに用いることができる。亜硫酸水素ナトリウムは、メチル化されたシトシンとは反応せずにシトシンの5, 6 - 二重結合と簡単に反応してスルホン化されたシトシン中間体を生成させる。この中間体は、アルカリ条件のもとで脱アミノ化反応によりウラシルになる。Taqポリメラーゼはウラシルをチミンとして認識し、5 - メチルシチジン (m^5C) をシチジンとして認識するため、亜硫酸水素ナトリウムによる処理とPCR増幅を順番に組み合わせることにより、最終的に、メチル化されていないシトシン残基はチミンに変換され (C U T)、メチル化されたシトシン残基 (“ m^5C ”) はシトシンに変換される (m^5C m^5C C)。したがってゲノムDNAを亜硫酸水素ナトリウムで処理すると、メチル化されていないシトシンがウラシルに変換されることにより、メチル化に依存した配列の違いが生まれ、PCRを行なうと、得られた産物は、修飾されていない核酸内でメチル化されたシトシンが生成される位置だけにシトシンを含むようになる。

【0091】

この明細書で用いるオリゴヌクレオチド“プライマー”とは、修飾された核酸または修飾されていない核酸に相補的な鎖と配列特異的なハイブリダイゼーション（アニーリング）をすることのできる、直線状で一本鎖になったオリゴマーのデオキシリボ核酸分子またはリボ核酸分子を意味する。特異的プライマーは、この明細書で用いられているように、DNAであることが好ましい。本発明のプライマーには、適切な配列で十分な長さがあり、増幅プロセスにおいて重合（プライマーの伸長）を特異的かつ効果的に開始させるオリゴヌクレオチドが含まれる。本発明の方法で用いられているように、オリゴヌクレオチド・プライマーは、一般に、12 ~ 30個またはそれ以上のヌクレオチドを含んでいるが、ヌクレオチドの数はそれよりも少なくてもよい。プライマーは18 ~ 30個のヌクレオチドを含んでいることが好ましい。正確な長さは、（増幅中の）温度、緩衝液、ヌクレオチドの組成など、多くの因子に依存するであろう。プライマーは一本鎖であることが好ましいが、二本鎖のプライマーも、鎖が最初に分離するのであれば用いることができる。プライマーは、適切な任意の方法を用いて調製することができる。例えば、従来のホスホトリエステル法、ホスホジエステル法や、従来技術でよく知られている自動化された方法がある。

【0092】

この明細書に記載した本発明の実施態様におけるように、特異的プライマーは、興味の対象であるゲノム遺伝子座の各鎖と実質的に相補的になるように設計することが好ましい。一般に、一方のプライマーは、遺伝子座のマイナス（-）鎖（水平に置いた二本鎖DNA分子の“下方の”鎖）と相補的であり、他方のプライマーは、プラス（+）鎖（“上方の”鎖）と相補的である。プロセスDの実施態様で用いられているように、プライマーは、DNAがメチル化される可能性のある部位（CpGヌクレオチド）と重なるように設計し

、修飾されたメチル化されていないDNAをメチル化されたDNAから特異的に識別できるようにすることが好ましい。この配列識別は、完全に一致したオリゴヌクレオチドと一致しないオリゴヌクレオチドでアニーリング温度に差があることに基づいていることが好ましい。プロセスDの実施態様では、プライマーは、一般に、1～数個のCpG配列と重なるように設計する。プライマーは、1～5個のCpG配列と重なるように設計することが好ましく、非常に好ましいのは1～4個のCpG配列と重なることである。対照的に、本発明の実施例における定量的実施態様では、プライマーはどのCpG配列とも重なっていない。

【0093】

十分に“メチル化されていない”（修飾されたメチル化されていない核酸鎖と相補的な）プライマー・セットの場合には、アンチセンス・プライマーは、対応する（-）鎖配列内にグアノシン残基（“G”）の代わりにアデノシン残基（“A”）を含んでいる。アンチセンス・プライマー中の置換されたこれらA残基は、メチル化されていないC残基（“C”）を亜硫酸水素塩で変化させた後に増幅を行なうことによって生成する対応する（+）鎖領域であり、ウラシル残基およびチミジン残基（“U”と“T”）と相補的になるであろう。この場合、センス・プライマーは、アンチセンス・プライマー伸長産物と相補的になるよう設計し、対応する（+）鎖配列内のメチル化されていないC残基の代わりにT残基を含むようにすることが好ましい。このセンス・プライマー中の置換されたこれらT残基は、元の（+）鎖の中の修飾されたC（U）残基と相補的な位置においてアンチセンス・プライマー伸長産物中に組み込まれたA残基と相補的になるであろう。

【0094】

十分にメチル化された（メチル化されたCpGを含む核酸鎖と相補的な）プライマーの場合には、アンチセンス・プライマーは、元の（+）鎖中のメチル化されたC残基（すなわち^mCpG配列）と相補的な対応する（-）鎖配列内にG残基の代わりにA残基を含むことはなかろう。同様に、この場合には、センス・プライマーは、対応する（+）鎖の^mCpG配列中にあるメチル化されたC残基の代わりにT残基を含むことはなかろう。しかし十分にメチル化されたプライマーによってカバーされる領域内のCpG配列中に存在しておらず、しかもメチル化されていないC残基は、メチル化されていないプライマーについて上で説明したように、十分にメチル化されたプライマー・セットの中に含まれるであろう。

【0095】

プロセスDの実施態様におけるように、増幅プロセスにより、特異的オリゴヌクレオチド・ハイブリダイゼーション・プローブの存在下で、亜硫酸水素塩で変換された核酸を2つのオリゴヌクレオチド・プライマーを用いて増幅することが好ましい。プライマーとプローブの両方とも、修飾されたメチル化されていない核酸と修飾されたメチル化された核酸を識別できる。さらに、“メチル化された”核酸の検出は、増幅により得られるプローブの蛍光に基づいている。一実施態様では、蛍光は、ポリメラーゼ酵素の5' 3'エキソヌクレオチド活性によってプローブが分解されることによって発生する。別の実施態様では、蛍光は、2つの互いに隣接したハイブリダイゼーション・プローブ（ライトサイクラー（Lightcycler：商標））の間で、またはハイブリダイゼーション・プローブとプライマーの間で蛍光エネルギーが移動することによって発生する。別の実施態様では、蛍光は、プライマー自身から発生する（サンライズ（Sunrise：商標）法）。増幅プロセスは、オリゴヌクレオチド・プライマーを用いた酵素による連鎖反応であることが好ましく、この連鎖反応により、標的となる遺伝子座から、関係する反応ステップの数からして指数関数的な量の増幅産物を生成させる。

【0096】

上で説明したように、プライマー・セットのうちの一方は（-）鎖と相補的であり、他方は（+）鎖と相補的である。プライマーは、興味の対象となる増幅領域、すなわち“アンプリコン”を挟むように選択する。標的となる変性した核酸にプライマーをハイブリダイズさせた後、DNAポリメラーゼとヌクレオチドを用いてプライマーを伸長させると、ア

10

20

30

40

50

ンプリコンに対応する新しい核酸鎖が合成される。DNAポリメラーゼは、従来技術で一般に用いられているTaqポリメラーゼであることが好ましい。しかし、5' 3'ヌクレアーゼ活性を有する同等なポリメラーゼを代わりに使用することもできる。新しいアンプリコンの配列もプライマーおよびポリメラーゼの鑄型であるため、変性、プライマーのアニールング、伸長というサイクルを繰り返すことにより、アンプリコンが指数関数的に生成される。連鎖反応の産物は、アンプリコン配列に対応していて、端部が、使用した特異的プライマーの端部によって決まる別々の核酸のコピーである。利用する増幅法は、PCR法(Mullis他、Cold Spring Harb. Symp. Quant. Biol.、第51巻、263~273ページ; Gibbs、Anal. Chem.、第62巻、1202~1214ページ、1990年)であることが好ましいが、それよりも好ましいのは、従来技術でよく知られている、この方法を自動化した方法である。 10

【0097】

メチル化に依存した配列の違いは、蛍光に基づいた定量的PCR法(リアルタイム定量的PCR、Heid他、Genome Res.、第6巻、986~994ページ、1996年; Gibson他、Genome Res.、第6巻、995~1001ページ、1996年)(例えば“タックマン(商標)”法、“ライトサイクラー(商標)”法、“サンライズ(商標)”法)で検出することが好ましい。タックマン(商標)法とライトサイクラー(商標)法の場合には、配列の識別は、(1)増幅ステップと(2)蛍光検出ステップという2つのステップの一方または両方においてなされる。サンライズ(商標)法の場合には、増幅ステップと蛍光検出ステップが一致している。FRETハイブリダイゼーションの場合には、ライトサイクラー(商標)に関するプローブのフォーマットとして、配列の違いを識別するためにFRETオリゴヌクレオチドの一方または両方を用いることができる。非常に好ましい増幅法は、この明細書に記載した本発明のすべての実施態様で用いられているように、二重標識した蛍光オリゴヌクレオチド・プローブを利用した、蛍光に基づいたリアルタイム定量的PCR法(Heid他、Genome Res.、第6巻、986~994ページ、1996年)である(ABIプリズム7700配列検出システム(パーキン・エルマー・アプライド・バイオシステムズ社、フォスター・シティ、カリフォルニア州)を使用したタックマン(商標)PCR)。 20

【0098】

“タックマン(商標)”PCR反応では、ペアの増幅用プライマーと、タックマン(商標)プローブと呼ばれる、応答を引き出すための伸長不能なオリゴヌクレオチドとを用いる。このプローブは、順プライマーと逆プライマー(すなわちセンス・プライマーとアンチセンス・プライマー)の間に位置するGCリッチな配列とハイブリダイズするように設計されている。タックマン(商標)プローブは、さらに、蛍光性“レポーター部分”と、タックマン(商標)オリゴヌクレオチドのヌクレオチドに結合しているリンカー部分(例えばホスホラミダイト)と共有結合する“クエンチャー部分”とを備えている。適切なレポーター分子とクエンチャー分子の具体例としては、5'蛍光レポーター染料6FAM(“FAM”; 2', 7'-ジメトキシ-4', 5'-ジクロロ-6'-カルボキシ-フルオレセイン)、TET(6'-カルボキシ-4', 7', 2', 7'-テトラクロロフルオレセイン)、3'クエンチャー染料TAMRA(6'-カルボキシテトラメチルローダミン)(Livak他、PCR Methods Appl.、第4巻、357~362ページ、1995年; Gibson他、Genome Res.、第6巻、995~1001ページ、1996年; Heid他、Genome Res.、第6巻、986~994ページ、1996年)が挙げられる。 30 40

【0099】

適切なタックマン(商標)プローブを設計する1つの方法では、“プライマー・エクスプレス(Primer Express)”などの支援用ソフトウェア・ツールを用いる。この“プライマー・エクスプレス”は、特異的配列(AT塩基対と比べてCG塩基対は結合が強い)であるために、またはプライマーの長さのために(プライマーの融点と比べて)融点に少なくとも10の差を与えるGCリッチな配列内のCpG島の位置に関する 50

変数を決定することができる。

【0100】

タックマン(商標)プローブは、本発明のどの方法を利用するかに応じ、既知のCpGメチル化部位をカバーしている場合としていない場合が可能である。プロセスDの実施態様では、タックマン(商標)プローブは、1~5個のCpG配列を重ねることにより、修飾されたメチル化されていない核酸と修飾されたメチル化された核酸を識別できるように設計することが好ましい。十分にメチル化されていないプライマー・セットと十分にメチル化されたプライマー・セットについて上に説明したように、タックマン(商標)プローブは、修飾されていない核酸と相補的になるように、あるいは適切な塩基置換により、亜硫酸水素塩処理で変換した配列と相補的になるように設計することができる。なお亜硫酸水素塩処理で変換した配列は、修飾されていない元の核酸サンプル中で十分にメチル化されていない配列か、十分にメチル化されている配列である。

10

【0101】

タックマン(商標)PCR反応におけるそれぞれのオリゴヌクレオチド・プライマーまたはプローブは、0個から異なる多数個のCpGジヌクレオチドにまたがることが可能であるため、それぞれ、亜硫酸水素塩処理の後に、2つの異なる配列となることができる(^mCpGまたはUpG)。例えば、1つのオリゴヌクレオチドが3つのCpGジヌクレオチドにまたがっている場合、ゲノムDNAに生じる可能な異なる配列の数は、 $2^3 = 8$ 種類である。順プライマーと逆プライマーのそれぞれが3つのCpGジヌクレオチドにまたがっており、プローブ・オリゴヌクレオチド(またはFRETフォーマットの場合には両方のオリゴヌクレオチド)が別の3つのCpGジヌクレオチドにまたがっている場合、配列の組み合わせの合計数は $8 \times 8 \times 8 = 512$ 通りになる。理論上は、これら512通りの配列それぞれの相対量を定量的に分析するのに別々のPCR反応を設計することができる。実際には、メチル化に関する定量的情報のほとんどは、はるかに少数の異なる配列を分析することで得られる。したがって、最も単純な形態として、仮想的な実施例における最も極端な配列である、十分にメチル化された配列と十分にメチル化されていない配列に対する反応を設計することによって、本発明の方法を実施することができる。これら2つの反応の比、またはメチル化された配列の反応と対照反応(プロセスA)の比から、この遺伝子座におけるDNAメチル化レベルの指標が得られよう。

20

【0102】

プロセスDのメチライト(商標)の実施態様におけるメチル化の検出は、この明細書に記載した他のメチライト(商標)の実施態様におけるのと同様、増幅によるプローブの置換に基づいている。理論上は、プローブの置換プロセスは、プローブをそのままにするか、あるいはプローブを消化させるように設計することができる。プローブの置換は、この明細書におけるのと同様、増幅中にプローブの消化によって起こることが好ましい。PCRサイクルの伸長段階において、蛍光ハイブリダイゼーション・プローブを、DNAポリメラーゼの5'→3'ヌクレアーゼ活性によって開裂させる。プローブが開裂すると、レポーター部分の放出はもはやクエンチャー部分に効果的には伝わらないため、518nmにおけるレポーター部分の蛍光放出スペクトルが大きくなる。クエンチャー部分(例えばTAMRA)の蛍光強度は、PCR増幅の間を通じてほんのわずかしき変化しない。タックマン(商標)PCR反応の効率にはいくつかの因子が影響を与える可能性がある。そのような因子としては、例えば、磁性や塩の濃度; 反応条件(時間と温度); プライマーの配列; PCRの標的のサイズ(すなわちアンプリコンのサイズ)と組成などがある。所定のゲノム遺伝子座について最適の蛍光強度を得るためにこれらの因子を最適化する方法は、PCRの当業者には明らかであるが、好ましい条件をこの明細書の“実施例”においてさらに詳しく説明する。アンプリコンのサイズは、50~8,000塩基対またはそれ以上が可能であるが、これよりも短くてもよい。一般に、アンプリコンは100~1,000塩基対であり、好ましいのは100~500塩基対である。反応は、96ウエルの光学トレイおよびキャップと、配列検出装置(ABIプリズム)とを用いてPCR増幅を行なうことによりリアルタイムで監視し、PCR増幅中の熱サイクルにおける96ウエルすべて

30

40

50

の蛍光スペクトルを連続的に測定できるようにすることが好ましい。入れる核酸の量を制御し、トレイごとのデータを規格化するため、プロセスDはプロセスAと合わせて実行することが好ましい。

【0103】

メチライト（商標）プロセスC。メチライト（商標）プロセスを変更し、PCR産物の検出段階で配列を識別することを回避できる。そこで、追加の定性的方法では、プライマーだけがCpGジヌクレオチドをカバーするように設計し、増幅段階においてだけ配列が識別されるようにする。この実施態様で用いるプローブはやはりタックマン（商標）プローブであるが、このプローブは、修飾されていない元の核酸中に存在するどのCpG配列とも重ならないように設計することが好ましい。プロセスCのこの実施態様は、MSP法をハイスループットで蛍光に基づいたリアルタイムの方法にしたものであり、メチル化されたCpG配列を検出するのに必要な時間を短くすることによって顕著な改善が達成されている。反応は、96ウエルの光学トレイおよびキャップと、配列検出装置（ABIプリズム）とを用いてPCR増幅を行なうことによりリアルタイムで監視し、PCR増幅中の熱サイクルにおける96ウエルすべての蛍光スペクトルを連続的に測定できるようにすることが好ましい。入れる核酸の量を制御し、トレイごとのデータを規格化するため、プロセスCは（下に示す）プロセスAと合わせて実行することが好ましい。

10

【0104】

メチライト（商標）プロセスB。本発明の好ましい実施態様では、やはりメチライト（商標）プロセスを変更し、PCR産物の検出段階で配列を識別することを回避できる。定量的プロセスBの一実施態様では、プローブだけがCpGジヌクレオチドをカバーするように設計し、プローブのハイブリダイゼーション段階においてだけ配列が識別されるようにする。タックマン（商標）プローブを使用することが好ましい。その場合、もともと増幅に偏りがあるのでない限り、亜硫酸水素塩による変換ステップで生成されるさまざまな配列が同じ効率で増幅される（Warnecke他、Nucleic Acids Res.、第25巻、4422～4426ページ、1997年）。特定の1つのメチル化パターンに関係する異なる配列のそれぞれについて別々のプローブを設計すると（例えば3個のCpGの場合には $2^3 = 8$ 個のプローブ）、いろいろなPCR産物が混合したプール内のそれぞれの配列の相対的割合を定量的に決定することができよう。反応は、96ウエルの光学トレイおよびキャップと、配列検出装置（ABIプリズム）とを用いてPCR増幅を行なうことによりリアルタイムで監視し、PCR増幅中の熱サイクルにおける96ウエルすべての蛍光スペクトルを連続的に測定できるようにすることが好ましい。入れる核酸の量を制御し、トレイごとのデータを規格化するため、プロセスBは下に示すプロセスAと合わせて実行することが好ましい。

20

30

【0105】

メチライト（商標）プロセスA。メチライト（商標）プロセスAそれ自体は、増幅段階でも検出段階でもメチル化されたCpG配列を識別しないが、入れたDNAの対照反応を提供することによって他の3つの方法の裏付けを与えてそれらの方法を意味あるものにするとともに、トレイごとのデータを規格化する。したがって、プライマーとプローブのいずれもどのCpGジヌクレオチドとも重なっていない場合には、反応は偏りのない増幅であり、蛍光に基づいた定量的リアルタイムPCRを利用した増幅の測定結果は、入れたDNAの量に対する対照となる。プロセスAは、増幅プロセスによる亜硫酸水素塩による処理が何らかの差をもたらすことを回避できるよう、プライマーおよびプローブ中にCpGジヌクレオチドがないだけでなく、アンプリコンにもCpGをまったく含んでいないことが好ましい。プロセスAのためのアンプリコンは、遺伝子増幅や遺伝子欠失などでコピー数がしばしば変化することのないDNA領域であることが好ましい。

40

【0106】

定性的メチライト（商標）法（プロセス“B”の実施態様）で得られた結果を以下の実施例で説明する。ヒトの腫瘍に関する多数のサンプルをこの方法で分析し、良好な結果を得た。

50

【0107】

ガンの診断法および予後予測法と、そのためのキット

一般に、本発明の診断法および/または予後予測法には、検査組織から組織サンプルを取得し、その組織サンプルに由来するDNAについてメチル化アッセイを行なって付随するメチル化状態を明らかにし、その結果に基づいて診断または予後予測を行なう操作が含まれる。

【0108】

好ましい実施態様では、ガンの診断法および予後予測法は、配列番号1~60、64、65(以下の表I Iを参照のこと)に対応するオリゴマー・プライマーおよびプローブによって規定されるここに示した20種類の遺伝子配列(APC、ARF、CALCA、CDH1、CDKN2A、CDKN2B、ESR1、GSTP1、HIC1、MGMT、MLH1、MYOD1、RB1、TGFB2、THBS1、TIMP3、CTNNB1、PTGS2、TYMS、MTHFR、またはメチル化によってこれら遺伝子において修飾されたDNA配列)のうちの1または複数のメチル化状態を明らかにすることに基づいている。配列番号61~63は、この分析で用いるACTB“対照”遺伝子領域に対応する(以下の実施例1を参照のこと)。

10

【0109】

さらに、所定のCpG島の一部におけるメチル化状態が一般にその島全体を代表しているという事実に基づくと、この明細書で用いるAPC、ARF、CALCA、CDH1、CDKN2A、CDKN2B、ESR1、GSTP1、HIC1、MGMT、MLH1、MYOD1、RB1、TGFB2、THBS1、TIMP3、CTNNB1、PTGS2、TYMSに関するCpG島の他の配列領域に対応する他のプライマーまたはプローブも用いることができる。

20

【0110】

そこで、公知の1または複数のメチル化アッセイ(上に説明したものも含む)を実行するのに必要な試薬をこのようなプライマーおよび/またはプローブ、またはその一部と組み合わせ、CpGを含む核酸のメチル化状態を明らかにする。

【0111】

例えば、メチライト(商標)、Ms-SNUPE、MCA、COBRA、MSPというメチル化アッセイを単独で、あるいは組み合わせ、配列番号1~65の配列またはその一部を含むプライマーまたはプローブを使用することで、APC、ARF、CALCA、CDH1、CDKN2A、CDKN2B、ESR1、GSTP1、HIC1、MGMT、MLH1、MYOD1、RB1、TGFB2、THBS1、TIMP3、CTNNB1、PTGS2、TYMS、MTHFRに対応する20種類の遺伝子配列領域のうちの1または複数の中にあるCpGジヌクレオチドのメチル化状態、あるいはこれら20種類の配列領域のうちの19種類(すなわちMTHFRを除くすべて)の場合には、これら配列に付随する他のCpG島配列のメチル化状態を明らかにする。なお、これら19種類の遺伝子配列に関する他のCpG島配列は、これら19種類の遺伝子配列領域のうちの1つの少なくとも1つのヌクレオチドを含み、CpGジヌクレオチドの頻度が観測値/予測値の比 > 0.6 に対応することと、GC含量 > 0.5 であることの両方の条件を満たす、ゲノムDNA中の連続配列である。

30

40

【0112】

実施例1

EACの進行とともに増加するCpG島の高メチル化

この実施例では、合計で20の遺伝子座の発ガンに関与することが知られているか、あるいは他の腫瘍でメチル化されていることが明らかになっているということで選んだ19の異なる遺伝子(表Iと上記の“定義”を参照のこと)と関係したCpG島のメチル化状態と、1つの非CpG島配列(MTHFR対照配列)のメチル化状態の分析結果を示す。

【0113】

バレット食道および/またはそれに付随する腺ガンがさまざまなステージにある患者31

50

人から採取した84個の組織サンプル中の20種類の遺伝子に関するメチル化の定量的データから、病気のステージが進むにつれてCpG島が高メチル化される頻度と程度が一般に大きくなることがわかった。そこで、組織ごとに、高メチル化の頻度と程度の両方に基づき、遺伝子をメチル化状態に従って異なるクラスに分けた(図1)。

【0114】

材料と方法

サンプルの採取と組織病理学的検査。腺ガンになっているか、IMが最も進んだ段階にある、合計で51人の患者(年齢は39歳~86歳)から、多数の組織サンプル(正常な食道(NE)、正常な胃(S)、腸上皮化生(IM)、異形成(DYS)および/または腺ガン(T))を採取した。

10

【0115】

分析した最初のサンプル群は31人の患者から採取したバイオプシー用組織である。採取してすぐに分割し、各サンプルの一部を液体窒素にただちに浸して凍らせ、パラフィン中に埋めた。それを病理学者(K.W.)が組織病理学的に調べた。すべての患者について、病気になっている領域から正常な食道の組織を10cmまたはそれ以上採取した。診断がはっきりしない場合には、凍らせた組織の凍結切片の検査を行なった。接合部を胃粘膜の皺の近位縁であると定義すると、ガンの中心が解剖学的な胃食道接合部よりも上である場合には、ガンの原発部位の分類を食道にした。TNM病期分類を用いて各腺ガンのステージを分類した。

【0116】

20の症例について追試研究を行なうため、第2のサンプル群を採取した。2種類のIMサンプル群を採取した。一方のサンプル群は、最も進行したステージのIMだけを有する患者(8人)からのものであり、他方のサンプル群は、IMとそれに付随した異形成/腺ガンが食道の別の領域にある患者(12人)からのものである。各サンプルについてH&Eスライド(5ミクロンの切片)を用意し、病理学者(K.W.)が検査してIM組織を確認し、異常のある位置を特定した。分析に用いたパラフィン・ブロック中に異形成または腺ガンの何らかの徴候が見られたケースをこの追試研究では除外した。5ミクロンのH&E切片に隣接する30ミクロンの切片から、IM組織を注意深く顕微解剖して他のタイプの細胞と分離した。サンプル中に存在する最も進行した組織病理学的病変がどのステージにあるかに従って、すべてのサンプルを分類した。この研究に関しては、南カリフォルニア大学ケック医学部の審査委員会から承認を得た。

20

30

【0117】

核酸の単離。単純化したプロテイナーゼK消化法により、凍らせたバイオプシー用組織からゲノムDNAを単離した(Laird他、Nucleic Acids Res.、第19巻、4293ページ、1991年)。パラフィン包埋組織からのDNAを一晩かけて50℃にて溶解用緩衝液(100mMのトリスHCl、pH8; 10mMのEDTA; 1mg/mlのプロテイナーゼK)中に抽出した(Shibata他、Am. J. Pathol.、第141巻、539~543ページ、1992年)。

【0118】

亜硫酸水素ナトリウムによる変換。亜硫酸水素ナトリウムによるゲノムDNAの変換を、以前に報告されているようにして行なった(Olek他、Nucleic Acids Res.、第24巻、5064~5066ページ、1996年)。ビーズを50℃にて14時間にわたって培養し、完全に変換されるようにした。亜硫酸水素ナトリウムによる処理は、メチル化されていないシトシンがウラシルに変換されるのに対し、メチル化されたシトシン残基はそのままになる(Frommer他、Proc. Natl. Acad. Sci. USA、第89巻、1827~1831ページ、1992年)。

40

【0119】

メチライト(商標)分析。亜硫酸水素ナトリウムによる変換の後、この明細書に記載したようにして、また以前に報告されているようにして(Eads他、Cancer Res.、第60巻、5021~5026ページ、2000年; Eads他、Cancer R

50

e s .、第 59 卷、2302 ~ 2306 ページ、1999 年；E a d s 他、N u c l e i c A c i d s R e s .、第 28 卷、E 32 ページ、2000 年）、蛍光に基づいたリアルタイム PCR アッセイであるメチライト（商標）によりメチル化分析を行なった。亜硫酸水素塩で変換した DNA に対して特異的になるように設計したプライマーとプローブを 2 セット用いた。用いたのは、興味の対象である遺伝子用のメチル化されたセットと、入れた DNA に対する規格化用の - アクチン（ACTB）という基準セットである。ヒト精子の DNA と、S s s I（ニュー・イングランド・バイオラズ社）で処理した精子の DNA（高度にメチル化されている）とを用い、以前に報告されているようにして（E a d s 他、C a n c e r R e s .、第 60 巻、5021 ~ 5026 ページ、2000 年）、メチル化された DNA に対する反応の特異性を別々に確認した。

10

【0120】

サンプルの遺伝子 / ACTB の比を、S s s I で処理した精子の DNA の遺伝子 / ACTB の比で割り、100 を掛けることにより、特定の遺伝子座における十分にメチル化された分子の割合を計算した。PMR（メチル化された基準の割合）という略号を用い、この測定結果を示した。亜硫酸水素塩による処理の後、付随している異形成の状態がサンプル中でどうなっているかを知らない人が、上に説明したようにして、顕微解剖したパラフィン包埋サンプルに関するメチル化分析を行なった。

【0121】

表 I I には、本発明のメチル化分析で用いたメチライト（商標）プライマーとプローブの配列（配列番号 1 ~ 65）を、Genbank の配列データに基づき（配列番号 64 と 65 は除く、下の説明を参照のこと）、リストにしてある。以下の 3 つのオリゴをすべての反応で用いた。すなわち、用いたのは、5' 蛍光レポーター染料（6FAM）と 3' クエンチャー染料（TAMTRA）を有する 1 つのオリゴヌクレオチド・プライマーが隣接した、遺伝子座特異的な 2 つの PCR プライマー（Livak 他、PCR Methods Appl .、第 4 巻、357 ~ 362 ページ、1995 年）である。各配列についての Genbank 登録番号を、この配列内の対応する PCR アンプリコンの位置とともにリストにしてある。メチライト・アンプリコンを含む 200 塩基対についての GC 含量 %、CpG の観測値 / 期待値の比、CpG : GpC の比を、それぞれの遺伝子について示してある。反応のタイプは、メチル化反応については“M”で、対照反応については“C”で示してある。亜硫酸水素塩で処理した DNA 鎖（上（“T”）と下（“B”））とアンプリコンの方向（平行（“P”）または反平行（“A”））も示してある。すべてのプライマーとプローブの配列を 5' から 3' の方向に示してある。それぞれのプライマーまたはプローブの配列の後ろにある括弧内の数字は、関係する配列番号に対応している。星印 1 つ（*）は、この Genbank の配列とは異なる塩基が 2 つ、われわれの CDKN2A プライマーの中に存在していることを示す。そうなったのも、予備的ハイスループット Genbank の入力データが、出願人がプライマーを設計したときに利用できる唯一の配列だったからである。正しいプライマーは、以下のものでなければならない。順プライマーは、TGGAGGT T T T C G G T T G A T T G G T T（配列番号 64）、逆プライマーは、A A C A A C G C C C G C A C C T C C T（配列番号 65）。Genbank の配列と異なる塩基に下線を施してある。星印 2 つ（**）は、開始部位がはっきりしていないことを意味する。

20

30

40

【表 4】

表 11. メチライト (商標) プライマーとプローブの配列

HUGO による遺伝子名	Genbank の登録番号	亜硫酸水素塩 (T 鎖または B 鎖)	アンプリコンの位置 (Genbank での数値)	転写開始位置と比較したアンプリコンの位置 (塩基対)	GC 含量 (%)	CpG の観測値/予測値の比	CpG:GpC の比	順プライマーの配列 (5'-3')	プローブの配列 (5'-3')	逆プライマーの配列 (5'-3')

【表 5】

表11. のつづき

“クラス A” 遺伝子**										
CDKN2A*	NM000077	T	66 塩基対 -133 塩基対	+9/ +69	72	1.68	0.81	TGGAAT TTTCGG TTGATT GGTT[1]	6FAM- ACCCGA CCCCGA ACCGCG -TAMRA [3]	AACAAC GTCCGC ACCTCC T[2]
ESR1	X62462	T	2784 塩基対 -2884 塩基対	+14/ +114	69	1.52	0.83	GGCGTT CGTTTT GGGATT G[4]	6FAM-C GATAAA ACCGAA CGACCC GACGA- TAMRA[6]	GCCGAC ACGCGA ACTCTA A[5]
MYOD1	AF027148	T	9889 塩基対 -9962 塩基対	-375/ -302	61	1.12	0.74	GAGCGC GGGTAG TTAGCG [7]	6FAM-C TCCAAC ACCCGA CTACTA TATCCG CGAAA- TAMRA[9]	TCCGAC ACGCC TTTCC[8]
“クラス B” 遺伝子										
CALCA	X15943	T	1706 塩基対 -1806 塩基対	-128/ -28	60	0.72	0.53	GTTTTG GAAGTA TGAGGG TGACG[10]	6FAM-A TTCCGC CAATAC ACAACA ACCAAT AAACG- TAMRA[12]	TTCCCG CCGCTA TAAATC G[11]

10

20

30

【表 6】

表II. のつづき

MGMT	U95038	B	206 塩 基 対 -297 塩 基対	+392/ +483	67	0.88	0.55	GTAACG TATAAC GAAAAT CGTAAC AACC[1 6]	6FAM-C CTTACC TCTAAA TACCAA CCCCAA ACCCG- TAMRA[18]	AGTATG AAGGGT AGGAAG AATTCG G[17]
TIMP3	U33110	T	1051 塩 基 対 -1143 塩基対	+1051/ +1143	78	1.04	0.63	GCGTCG GAGGTT AAGGTT GTT[13]	6FAM-A ACTCGC TCGCC GCCGAA -TAMRA [15]	CTCTCC AAAATT ACCGTA CGCG[1 4]
“クラス C” 遺伝子										
APC	U02509	B	759 塩 基 対 -832 塩 基対	+7/ +80	69	1.36	0.71	GAACCA AAACGC TCCCCA T[19]	6FAM-C CCGTCG AAAACC CGCCGA TTA-TA MRA[21]	TTATAT GTCGGT TACGTG CGTTTA TAT[20]
“クラス D” 遺伝子										
ARF	AF0823 38	T	5447 塩 基 対 -5515 塩基対	-203/ -135	78	2	0.89	ACGGGC GTTTTT GGTAGT T[22]	6FAM-C GACTCT AAACCC TACGCA CGCGAA A-TAMR A[24]	CCGAAC CTCCAA AATGTC GA[23]
CDH1	L34545	T	842 塩 基 対 -911 塩 基対	-180/ -111	69	1.2	0.68	AATTTT AGGTTA GAGGGT TATCGC GT[25]	6FAM-C GCCCAC CCGACC TCGCAT -TAMRA [27]	TCCCCA AAACGA AACTAA CGAC[2 6]

10

20

30

40

【表 7】

表11. のつづき

CDKN2B	S75756	T	350 塩 基 対 -430 塩 基対	+222/ +302	64	1.44	0.82	AGGAAG GAGAGA GTGCGT CG [28]	6FAM-T TAACGA CACTCT TCCCTT CTTTCC CACG-T AMRA [3 0]	CGAATA ATCCAC CGTTAA CCG [29]
GSTP1	M24485	T	1146 塩 基 対 -1245 塩基対	-79/ +21	75	2	0.93	GTCGGC GTCGTG ATTTAG TATTG [31]	6FAM-A AACCTC GCGACC TCCGAA CCTTAT AAAA-T AMRA [3 3]	AAACTA CGACGA CGAAAC TCCAA [32]
MLH1	U26559	T	254 塩 基 対 -341 塩 基対	-662/ -575	68	2.08	1.04	CGTTAT ATATCG TTCGTA GTATTC GTGTTT [35]	6FAM-C GCGACG TCAAAC GCCACT ACG-TA MRA [36]	CTATCG CCGCCT CATCGT [34]
PTGS2	AF0442 06	T	6779 塩 基 対 -6924 塩基対	-362/ -217	62	0.8	0.71	CGGAAG CGTTCG GGTAAA G [37]	6FAM-T TTCCGC CAAATA TCTTTT CTTCTT CGCA-T AMRA [3 9]	AATTCC ACCGCC CCAAAC [38]
THBS1	J04835	B	1642 塩 基 対 -1716 塩基対	-636/ -562	75	2.08	1	CGACGC ACCAAC CTACCG [40]	6FAM-A CGCCGC GCTCAC CTCCCT -TAMRA [42]	GTTTTG AGTTGG TTTTAC GTTTCGT T [41]

10

20

30

40

表II. のつづき

“クラスE” 遺伝子										
CTNNB1	X89448	T	583 塩基対 -413/ -332 -664 塩基対	75	1.92	1	GAAAAG GCGCGT CGAGT [43]	6FAM-C GCGCGT TTCCCG AACCG- TAMRA [45]	TCCCCT ATCCCA AACCCG [44]	
RB1	L11910	T	1781 塩基対 -279/ -160 -1900 塩基対	68	2.24	1.27	TTAGTT CGCGTA TCGATT AGCG [46]	6FAM-T CAGGTC CGCGAA ACTCCC GA-TAM RA [48]	ACTAAA CGCCGC GTCCAA [47]	
TGFBR2	U52240	T	256 塩基対 -58/ +19 -333 塩基対	75	2.08	0.81	GCGCGG AGCGTA GTTAGG [49]	6FAM-C ACGAAC GACGCC TTCCCG AA-TAM RA [51]	CAAACC CCGCTA CTCGTC AT [50]	
TYMS	D00517	T	1052 塩基対 +301/ +377 -1128 塩基対	75	2.16	0.9	CGGCGT TAGGAA GGACGA T [52]	6FAM-C CGAATA CCGACA AAATAC CGATAC CCGT-T AMRA [54]	TCTCAA ACTATA ACGCGC CTACAT [53]	

10

20

30

【表 9】

表II. のつづき

“クラスF” 遺伝子										
HIG-1	L41919	T	562 塩 基 対 -662 塩 基対	-55/ +45**	71	1.12	0.5	GTTAGG CGGTTA GGGCGT C[58]	6FAM-C AACATC GTCTAC CCAACA CACTCT CCTACG -TAMRA [60]	CCGAAC GCCTCC ATCGTA T[59]
“クラスG” 遺伝子										
MTHFR	AF1059 77	T	123 塩 基 対 -212 塩 基対	+72/ +45	57	0.56	0.54	TGGTAG TGAGAG TTTTAA AGATAG TTCGA[55]	6FAM-T CTCATA CCGCTC AAAATC CAAACC CG-TAM RA[57]	CGCCTC ATCTTC TCCCGA [56]
“対照” 遺伝子										
ACTB	Y00474	T	390 塩 基 対 -522 塩 基対	-1599/ -1467*	58	0.16	0.14	TGGTGA TGGAGG AGGTTT AGTAAG T[61]	6FAM-A CCACCA CCCAAC ACACAA TAACAA ACACA- TAMRA[63]	AACCAA TAAAC CTACTC CTCCCT TAA[62]

括弧内の数字は配列番号 1~63 に対応している。

**遺伝子“クラス”は本発明に従って定義した。

10

20

30

40

50

【0122】

統計。統計処理をするため、以前に報告されているようにして (Eads 他、Cancer Res., 第60巻、5021~5026ページ、2000年)、メチライト(商標)により得られたPMR値(上の説明を参照のこと)を4PRMを境にして“2つに分けた”。2つに分けることにより、グラフ表示が容易になる。また、2つに分けることにより、高メチル化のレベルが異なる遺伝子座が数値に及ぼす影響が和らげられる結果、遺伝子相互間で高メチル化の頻度を比較するときの信頼性が高まる。特に、2つに分けることにより、各クラスでメチル化された遺伝子が数値に及ぼす影響が等しくなり、メチル化の頻度を遺伝子相互間で比較するのが容易になる。

【0123】

4PRMという二分点は、すべてのCpG島について正常な組織と悪性組織を最もよく区別できることから選択した(Eads 他、Cancer Res., 第60巻、5021~5026ページ、2000年)。しかしこの明確な二分点は、統計に有意に影響することはない、結論を変えることもないため、他の二分点も本発明の範囲に含まれる(下の説明を参照のこと)。

【0124】

そこで、4PRMまたはそれ以上のサンプルはメチル化されたとして値1を与え、4PRM未満のサンプルはメチル化されていないとして値0を与えた。次に、各クラス(下の“

後成的遺伝子クラス”を参照のこと)、または全部で19の遺伝子についてメチル化された遺伝子の積算値を、サンプル数が均等でない場合に使用できるようにしたフィッシャーの保護された最小有意差(PLSD)検定(SAS統計ビュー・ソフトウェア)における連続変数として用い、p値を得た。組織のタイプ、付随する異形成の存在、腫瘍のステージなどのさまざまなパラメータを名義変数として用いた。IMにおける高メチル化と、付随する異形成および/またはガンの存在に関する上記の“追試”研究におけるIMサンプルを、メチル化されたクラスA遺伝子が1個以下と2個以上の地点でさらに2つに分けた。次に、フィッシャーの検定を厳密に行なって統計的有意性を明らかにした。

【0125】

結果

CpG島の高メチル化とEACの進行。合計で20の遺伝子座についての、19の異なる遺伝子に付随する一群のCpG島のメチル化状態と、1つの非CpG島配列のメチル化状態とを、定量的ハイスループット・メチライト(商標)アッセイで分析した(Eads他、Cancer Res.、第59巻、2302~2306ページ、1999年; Eads他、Nucleic Acids Res.、第28巻、E32ページ、2000年)。メチル化反応の効率は、それぞれの分析においてメチル化されていない対照DNAとメチル化された対照DNAを含めることによって制御した(Eads他、Cancer Res.、第60巻、5021~5026ページ、2000年)。20種類の遺伝子は、腫瘍発生への関与が知られているという理由で、または他の腫瘍でメチル化されていることがわかっているという理由で選択した(表Iと上記の“定義”を参照のこと)。MTHFR遺伝子内に位置する1つの領域を、CpG島としての条件(上記の“定義”を参照のこと)を満たさない単一コピー配列のための“非CpG島”対照として含めた。島の外にあるCpGジヌクレオチドは、CpG島内のCpGジヌクレオチドとは異なり、おそらく普通にメチル化される。

【0126】

図1は、さまざまなステージにあるバレット食道および/またはそれに付随する腺ガンの患者31人から採取した84個の組織サンプル中の20種類の遺伝子について、メチル化を定量分析したデータを示している。メチル化の分析は、メチライト・アッセイを利用して行なった(Eads他、Cancer Res.、第59巻、2302~2306ページ、1999年; Eads他、Nucleic Acids Res.、第28巻、E32ページ、2000年)。特定の遺伝子座において十分にメチル化された分子の割合(PMR=メチル化された基準の割合)は、サンプルの遺伝子/ACTB比を、SsisIで処理した精子DNAの遺伝子/ACTB比で割り、その数値を100倍することによって計算した。次に、得られた値を4%PRMの位置で2つに分け、グラフを描きやすくするとともに、組織特異的パターンが明らかになるようにした。多数の正方形は、4通りのグレーで示した強度レベル(図1の最下部を参照のこと)のいずれかであり、それぞれ、PMRが4未満、4~20、21~50、51以上のサンプルを表わしている。なお、グレーが濃くなるほど、PMR値が大きくなっている。組織のタイプは左側に示してある。TNM病期分類を、“1”、“2”、“3”、“4”で表わしてある。異形成および/または腺ガンが離れた位置に発生した患者がいる場合には、図の右側に“はい”、いない場合には“いいえ”と表示してある。“N”は、その特定のサンプルの分析におけるメチル化反応については、対照とした遺伝子ACTBが1PMRという最小値を検出できるほど十分なレベルに達していなかったことを意味している。

【0127】

病気が重くなるにつれて、CpG島の高メチル化の頻度と定量的レベルが一般に増加していた。しかし遺伝子が異常にメチル化する傾向は一定ではなかった。遺伝子は、高メチル化の頻度とレベルの両方が組織ごとに異なっていた。

【0128】

したがって、本発明によれば、遺伝子は、メチル化の状態に基づいてクラス(図1の上部に示したようにクラスA~G)に分けることができる。このようにすることで、腫瘍発生

10

20

30

40

50

のさまざまなステージにおいて遺伝子がメチル化している様子を目で見て評価することができた。それぞれの遺伝子クラスについての説明は、以下のセクションで行なう。

【0129】

後成的遺伝子クラス。DNAのメチル化レベルが異なる複数の遺伝子の状態を合わせて分析すると、適切なデータ処理をしなければ、グループの振る舞いは、DNAのメチル化状態が高いレベルにある遺伝子に引っ張られて偏ることが予測される。例えば、ほとんどの腫瘍サンプルについて、遺伝子“クラスB”に関する平均値は、主としてTIMP3の値によって左右されよう。というのも、この遺伝子は、このグループ内の他の2つの遺伝子よりもメチル化の程度が大きい傾向があったからである(図1を参照のこと)。

【0130】

そこで、図1を作成するのに用いたメチル化値は、それぞれの後成的クラス内でメチル化された遺伝子が数値に及ぼす影響を等しくするため、4PRMという二分点を有する二値変数にした。4PRM以上のサンプルはメチル化されたとして値1を与え、4PRM未満のサンプルはメチル化されていないとして値0を与えた(上記“材料と方法”の中の“統計”を参照のこと)。このように2つに分けることにより、高度にメチル化された遺伝子の効果が和らげられ、図2に示したようにメチル化の頻度を遺伝子相互間で比較するのが容易になり、図3(下方の図)に示したように、メチル化の頻度のクラス平均を計算することが可能になる。

【0131】

図2は、各遺伝子についてメチル化されたサンプルの割合を組織のタイプごとに示した図である。データは、4PRMの位置で2つに分けた。4PRM以上だとメチル化されていて、4PRM未満だとメチル化されていないことを意味する。遺伝子は、本発明により、図1に示したようにそれぞれの後成的遺伝子クラス(A~G)に従ってグループ分けした。“n”は、各組織について分析したサンプルの数である。

【0132】

4PRMという二分点が適切であることは、図1~図3に示したように、この点が組織のタイプを識別する能力を持っていることに基づいている(Klump他、Gastroenterology、第115巻、1381~1386ページ、1998年も参照のこと)。高度にメチル化された遺伝子が統計に及ぼす効果を和らげ、遺伝子相互間のメチル化の頻度の比較が簡単にできるようにし、メチル化の頻度のクラス平均の計算を簡単にできるような他の二分点も本発明の範囲に含まれる。例えばデータを10PRMで二分する場合には、(19遺伝子のうちで)メチル化された遺伝子の割合の平均値を正常な食道粘膜とIM組織($p = 0.0003$)、DYS組織($p < 0.0001$)、T組織($p < 0.0001$)の間で比べると、やはり統計的に有意な差がある。

【0133】

さらに、付随した異形成を伴う、あるいは伴っていないNEとIM(以下の実施例3を参照のこと)のメチル化の頻度に関する統計的に有意な発見はすべて、二分点を4PRMではなく10PRMにしても有意なままである。4PRMは、単一のCpGジヌクレオチドの4%というメチル化レベルとは比較できないことに注目することが大切である。これは、このサンプルでは、DNA分子の4%が、3つのメチライト(商標)プライマーによってカバーされるすべてのCpGジヌクレオチド(通常は約8個のCpG)において完全にメチル化されていたことを示唆する。メチライト(商標)アッセイは、存在する可能性のある他のすべてのメチル化パターンを無視するという性質がある(Eads他、Nucleic Acids Res.、第28巻、E32ページ、2000年)。

【0134】

したがって4PRMは、メチル化の平均レベルが4%よりも大きいことを示している可能性が大きい。メチライト(商標)アッセイで調べる高度にメチル化された分子は、CpG島の高メチル化によって完全に不活性状態になった対立遺伝子を示している可能性が大きい。これについてはこの明細書では触れなかった。

【0135】

10

20

30

40

50

20種類の遺伝子群のうち、もっとも多く情報をもたらした遺伝子は、高メチル化の頻度が中間的な値の遺伝子である(4PRMというメチル化カットオフ値よりも大きいサンプルの数が15%(CDKN2A)から60%(MGMT)まで)。このグループは正常な食道粘膜および胃にメチル化がない(クラス“A”)、正常な食道粘膜および胃にメチル化がある(クラス“B”)、正常な食道粘膜ではメチル化があまり頻繁には起こっていないが、正常な胃のサンプルではすべてメチル化が起こっている(クラス“C”)に従い、さらに3つの後成的遺伝子クラスに分類することができた。他の遺伝子はより少ない情報しかもたらさなかった。というのも、高メチル化の発生は、非常に稀である(クラス“D”)、まったくない(クラス“E”)、組織のタイプに関係なくすべてにある(クラス“F”と“G”)という状態だったからである(図1、図2、図3)。

10

【0136】

後成的遺伝子クラスAには、CDKN2A、ESR1、MYOD1の遺伝子が含まれる(図1、図2、図3)。正常な食道(NE)では、IM組織と比較すると、メチル化の頻度に関する統計的に有意な差が、ESR1($p = 0.0001$)とMYOD1($p = 0.0038$)において存在していたが、CDKN2A($p = 0.097$)には存在していなかった。CDKN2Aのメチル化の頻度は、腺ガン(T)のさらに進んだステージで有意に増加した($p < 0.0001$)。

【0137】

後成的遺伝子クラスBには、CALCA、MGMT、TIMP3の遺伝子が含まれる。クラスAと比べると、このクラスは正常な食道粘膜(NE)と胃(S)の組織でメチル化を示す(図1と図2)。TIMP3だけが、メチル化の頻度に関してNE値とIM値の間で有意な差を示した($p = 0.0074$)。

20

【0138】

後成的遺伝子クラスCには、APCの遺伝子が含まれる。この遺伝子は、クラスAおよびクラスBの遺伝子とは異なり、正常な胃のサンプルすべてでメチル化されていた(図1と図2)。これは、正常な胃の組織でAPCがメチル化されているという以前の文献の結果を確認したことになっている(Eads他、Cancer Res., 第60巻、5021~5026ページ、2000年)。正常な胃の組織(S)ではなく正常な食道組織(NE)でAPCがメチル化されるのを防いでいるメカニズムははっきりしていない。

【0139】

後成的遺伝子クラスDには、あまりメチル化されていないARF、CDH1、CDKN2B、GSTP1、MLH1、PTGS2、THBS1の遺伝子が含まれる(図1と図2)。このクラスの遺伝子のメチル化の頻度は腺ガン(T)においてわずかに増加していたが、統計的に有意というほどではなかった(図3)。興味深いことに、他の系ではまだ調べられていないPTGS2を除き、クラスDのそれ以外の遺伝子は、他のタイプの腫瘍で高メチル化されていることがしばしばある。(表II)。

30

【0140】

後成的遺伝子クラスEには、EACが進行する各ステージでメチル化されていなかったCTNNB1、RB1、TGFB2、TYMSの遺伝子が含まれる。クラスDのほとんどの遺伝子と同様、RB1とTGFB2は、他のタイプの腫瘍で高メチル化されていることが見いだされた(表Iと、上記の“定義”に掲載した参考文献を参照のこと)。すべてのサンプルが、対照遺伝子(ACTB)と比較して、入れたDNAに対してプラスの反応を示したことに注目されたい。したがって、DNAのメチル化が検出されないのは、入れたDNAが不足していたからではありえない。対照反応はそれぞれのサンプルで十分に起こったため、テストする所定の遺伝子について、1PRMという低いレベルまで検出することができた。すべてのメチル化反応が完全に特異的であることは、インビトロでメチル化されたヒトDNAを用いて確認した。

40

【0141】

後成的遺伝子クラスFには、組織のタイプに関係なく完全にメチル化されたHIC1遺伝子が含まれる(図1と図2)。HIC1は、他のタイプのガンでも一般にメチル化されて

50

おり (Jones と Laird、Nat. Genet.、第 21 巻、163 ~ 167 ページ、1999 年; Baylin と Herman、Trends Genet.、第 16 巻、168 ~ 174 ページ、2000 年)、乳ガン患者の正常な胸管組織や AML 患者の骨髄サンプルでもメチル化されていることがわかっている (Melki 他、Cancer Res.、第 59 巻、3730 ~ 3740 ページ、1999 年; Fujii 他、Oncogene、第 16 巻、2159 ~ 2164 ページ、1998 年)。しかし正常な組織の CpG 島で 100% メチル化されているという発見は予想外であった。したがって、HIC1 に対するメチライト (商標) アッセイの結果の有効性は、別の方法 (HpaII-PCR) (Singer-Sam 他、Nucleic Acids Res.、第 18 巻、687 ページ、1990 年) で確認した。

10

【0142】

後成的遺伝子クラス G には、この明細書で対照として用いる非 CpG 島遺伝子 MTHFR が含まれる。興味深いことに、100% の割合で見られる HIC1 のメチル化パターンは、対照である非 CpG 島遺伝子 MTHFR (クラス G) と似ているが、メチル化された分子の割合は HIC1 のほうが量的に多い (図 1)。

【0143】

EAC の進行の後成的パターン。それぞれのタイプの組織は、病気の進行とともに変化する独特の後成的パターンまたはフィンガープリントを示した (図 3 の上方の図)。

【0144】

図 3 は、本発明による後成的パターンを比較した図である。データは、4PRM の位置で 2 つに分けた。4PRM 以上だとメチル化されていて、4PRM 未満だとメチル化されていないことを意味する。誤差を示す棒は、平均値の標準偏差を表わしている。上方の図：各遺伝子クラス (A ~ F、または 19 ある CpG 島の合計) においてメチル化された遺伝子の割合の平均値を組織のタイプ別 (N、正常な食道; S、胃; IM、腸上皮化生; DYS、異形成; T、腺ガン) に示したもの。誤差棒は、平均値の標準偏差 (SEM) を示す。下方の図：遺伝子の各クラス (A ~ F) と、19 の CpG 島をすべて合わせた場合 (合計) につき、さまざまな組織でメチル化された遺伝子の割合の平均値の違いを統計的に分析した結果。p 値は、フィッシャーの保護された最小有意差 (PLSD) 検定をサンプル数が均等でない場合に使用できるようにして求めた (SAS スタットビュー (商標) ソフトウェア)。

20

30

【0145】

クラス A、B、C は、正常な食道粘膜 (NE) よりも IM 組織において有意に高い頻度でメチル化されていた (図 3 の上下の図)。さらに、IM から異形成 (DYS) または悪性 (T) への移行は、クラス A におけるメチル化の増加と関係していた (図 3 の上下の図)。どの遺伝子クラスについても、また 19 種類の遺伝子をすべて合計した場合でも (図 3 の上下の図)、異形成と腺ガンの間に有意な差がない。これは、こうした異常な後成的変化のほとんどが、EAC の進行の初期に起こることを示唆している。

【0146】

この実施例のまとめ。本発明によれば、バレット食道および / またはそれに付随する腺ガンがさまざまなステージにある患者 31 人から採取した 84 個の組織サンプル中の 20 種類の遺伝子 (上記の表 I と表 II) に関するメチル化の定量的データから、病気の段階が進むにつれて CpG 島が高メチル化される頻度と程度が一般に大きくなることがわかった (上記の図 1 ~ 図 3)。

40

【0147】

さらに、遺伝子を、腫瘍形成の間のメチル化の状態に基づき、新たな後成的クラスに分類した (図 1 ~ 図 3 に示したクラス A ~ G)。こうすることにより、腫瘍発生のさまざまなステージにおいて異なる遺伝子が一斉にメチル化されている様子をグラフに表わすことができた。すると単に目で見ただけで簡単に評価することができる。

【0148】

それぞれのタイプの組織は、病気が進行するにつれて変化する後成的パターンまたはフィ

50

ンガープリントを示した(図3の上方の図)。クラスA、B、Cは、正常な食道粘膜(NE)よりもIM組織において有意に高い頻度でメチル化されていた(図3の上下の図)。さらに、IMから異形成(DYS)または悪性(T)への移行は、クラスAにおけるメチル化の増加と関係していた(図3の上下の図)。

【0149】

実施例2

高メチル化は、EACの腫瘍の悪性度とステージを反映していた

この実施例では、食道腺ガンの悪性度またはステージが、CpG島の高メチル化の頻度が大きいほど進んでいることと相関しているかどうかを調べた。本発明によれば、EACの場合には、後成的遺伝子クラスAの遺伝子は、ステージII、III、IVの腫瘍のほうが、それよりも前のステージIの腫瘍よりも有意に多くメチル化されていることがわかる(図4)。

10

【0150】

材料と方法

TNM病期分類法。ガンに関するアメリカ両院合同委員会("AJCC")は、TNM(腫瘍：リンパ節転移、遠方転移)分類による病期分類を提示している。TNM病期分類法を用い、実施例1の組織からそれぞれの食道腺ガンのステージを分類した。

【0151】

メチル化と統計的分析。メチル化と統計的分析については、この明細書の実施例1のところで説明した。

20

【0152】

結果

クラスA遺伝子のメチル化の頻度は、腫瘍のステージが進むほど大きくなる。ある程度分化した腫瘍は、あまり分化していない腫瘍と比べてクラスA遺伝子のメチル化が有意に少ない($p = 0.045$)。しかも、図4(上下の図)から、ステージII、III、IVの腫瘍においては、それよりも前のステージIの腫瘍よりもメチル化されているクラスA遺伝子の平均数が有意に多いことがわかる。ステージIの腫瘍とステージII、III、IVの腫瘍の差は、他のどの後成的遺伝子クラスでも統計的に有意ではなかった。

【0153】

図4は、クラスA遺伝子のメチル化の頻度と腫瘍のステージの関係を、本発明に従って示した図である。データは、4PRMの位置で2つに分けた。4PRM以上だとメチル化されていて、4PRM未満だとメチル化されていないことを意味する。上方の図：メチル化されたクラスA遺伝子の平均数を腫瘍のステージ(I~IV)ごとに示したもの(図1を参照のこと)。誤差棒は、平均値の標準偏差(SEM)を示す。"n"は、腫瘍の各ステージについて分析したサンプルの数である。下方の図：腫瘍のステージそれぞれにつき、メチル化された遺伝子の数の平均値の違いを統計的に分析したもの。p値は、フィッシャーの保護された最小有意差(PLSD)検定をサンプル数が均等でない場合に使用できるようにして求めた(SASスタットビュー(商標)ソフトウェア)。

30

【0154】

この実施例のまとめ。本発明によれば、発ガンの進行を評価するのに後成的パターンまたはフィンガープリント(この明細書に記載した遺伝子クラスを含む)を利用できることに加え、メチル化されたクラスA遺伝子の平均数もEAC腫瘍の相対的ステージを評価するのに用いることができる。

40

【0155】

実施例3

付随する異形成を伴った、あるいは伴わない前ガン組織のメチル化

この実施例は、正常な食道(NE)におけるクラスB遺伝子のメチル化の頻度が、付随する異形成/腫瘍を伴った患者で有意に大きかったことを示している($p = 0.0037$)(図1)。さらに、クラスA遺伝子のメチル化は、異形成またはガンを伴っている患者からのIMサンプルにおいて、病気が以前よりも進行しているという何らの証拠もない患者

50

からのIMサンプルにおけるよりも頻度が大きいことがわかった ($p < 0.0001$) (図1と図5)。すなわち、IMサンプル中の後成的クラスAの遺伝子の高メチル化と、付随する異形成またはガンの存在の間には、有意に正の相関関係が存在していた (図5)。

【0156】

材料と方法

組織病理学。組織病理学的分類については、上記の実施例1の“材料と方法”で説明した。

【0157】

メチル化と統計的分析。メチル化と統計的分析については、実施例1で説明したのと同じである。

10

【0158】

結果

付随する異形成を伴った、あるいは伴わない前ガン組織のメチル化。本発明によれば、クラスAに関するいくつかのIMのケースとクラスBに関する正常な食道粘膜のいくつかのケースにおいてCpG島に高メチル化が起こっていることから、こうしたメチル化が、異形成のないこれらの組織における正常なメチル化パターンを表わしているのか、それとも病気をさらに進行させようとするメチル化の変化を反映しているのかという疑問が湧いた。後者の場合、すでに病気が進行した患者のこれら組織において、CpG島の高メチル化がより高い頻度で見いだされることが予測されよう。したがって、CpG島の高メチル化の頻度を、(この研究の)付随する異形成を伴った組織と付随する異形成を伴わない組織で比較した。

20

【0159】

最初の研究では、バレット食道(IM)が最も進行した段階であるかどうか、あるいは付随する異形成および/または腺ガンが食道の別の領域に存在しているかどうか (図1、“はい”、“いいえ”)に基づき、患者を分類した。実際には、正常な食道(NE)におけるクラスB遺伝子のメチル化の頻度は、付随する異形成を伴った患者において有意に大きかった ($p = 0.0037$) (図1)。さらに、クラスA遺伝子のメチル化は、異形成またはガンを伴っている患者からのIMサンプルにおいて、病気が以前よりも進行しているという何らの証拠もない患者からのIMサンプルにおけるよりも頻度が大きいことがわかった ($p < 0.0001$) (図1)。

30

【0160】

この分析に対して寄せられる可能性のある批判は、遺伝子のクラスを説明するのに臨床パラメータとの関係をテストするのに用いたのと同じサンプル群を用いているという点である。そこで、IMに関する別の20症例についての追試研究を、最初のデータ群とは完全に独立に実施した。

【0161】

20の症例に関するこの追試研究では、2つのグループからIMサンプルを採取した。1つは、最も進行した段階のIMだけを有する患者(8人)で、もう1つは、付随した異形成/腺ガンが食道の別の領域にあるIMを有する患者(12人)である。各サンプルについてH&Eスライド(5ミクロンの切片)を用意し、病理学者(K.W.)が検査してIM組織を確認し、異常のある位置を特定した。分析に用いたパラフィン・ブロック中に異形成または腺ガンの何らかの徴候が見られたケースをこの追試研究では除外した。5ミクロンのH&E切片に隣接する30ミクロンの切片から、IM組織を注意深く顕微解剖して他のタイプの細胞と分離した。サンプル中に存在する最高ステージの組織病理学的病巣に従ってすべてのサンプルを分類した。

40

【0162】

最初の研究により、病気がより進行したすべてのIMサンプル(“はい”)ではクラスAの少なくとも2つの遺伝子がメチル化されていたのに対し、付随する異形成または腺ガンがないすべてのIMサンプル(“いいえ”)ではクラスAのどの遺伝子もメチル化されていなかったことが明らかにされていた (図1、バレット食道(IM)の項)。したがって

50

、クラス A の 2 つ以上の遺伝子がメチル化されているという状態は、付随する異形成または腺ガンが存在しているリスクが大きいことの指標であるとされた。

【 0 1 6 3 】

われわれの最初の研究でのデータから、この関連性に関してフィッシャーの検定を厳密に行なったところ、p 値が 0 . 0 0 4 8 となった (図 5 、 左 図) 。 それとは独立な 2 0 症例での追試研究からは、p 値として 0 . 0 1 8 が得られた (図 5 、 右 図) 。

【 0 1 6 4 】

図 5 は、付随する異形成および / または腺ガンを伴っている場合 (“ Y ”) または伴っていない場合 (“ N ”) について、腸上皮化生 (“ I M ”) 組織においてメチル化されているクラス A の 2 つ以上の遺伝子の割合を示している。データは、4 P R M の位置で 2 つに分けた。4 P R M 以上だとメチル化されていて、4 P R M 未満だとメチル化されていないことを意味する。左側の図 : 図 1 に示した I M に関するデータの中のクラス A 遺伝子のメチル化。右側の図 : 顕微解剖した 2 0 の異なる I M サンプルに関して行なった完全に独立な追試研究における I M 組織中のクラス A 遺伝子のメチル化。誤差棒は、平均値の標準偏差 (S E M) を示す。 “ n ” は、それぞれの組織群について分析したサンプルの数である。

10

【 0 1 6 5 】

したがって、クラス A 遺伝子の高メチル化と、付随する異形成またはガンの存在の間には、有意な正の相関関係がある。この追試研究において、付随する異形成のない I M サンプル (図 5 、 右 図) の中で少なくとも 2 つの遺伝子がメチル化されているサンプルが、少ない頻度で存在していたことに注意されたい。これは、最初の研究でメチル化されたサンプルが存在していなかったことと対照的である (図 1 と、図 5 の左 図) 。このような結果になったのは、追試研究でのサンプルはパラフィン包埋切片から顕微解剖したからである可能性がある。したがって、サンプル中には、背景となるメチル化されていないストロマ細胞がより少ない。この場合、メチル化信号が他の正常な細胞によってそれほど弱められることはなく、したがって全 D N A に対するメチル化された分子の比が 4 P R M という閾値を超える可能性がある。これ以外にも、内視鏡によるサンプル採取の限界のため、病気が以前と比べて進行していないと判定されたいいくつかのケースにおいて、内視鏡検査の間に異形成またはガンの組織が見逃されていた可能性がある。これは、食道腺ガンの検出において詳しく報告されている問題である (P e t e r s 他、J . T h o r a c . C a r d i o v a s c . S u r g . 、第 1 0 8 巻、8 1 3 ~ 8 2 1 ページ、1 9 9 4 年) 。

20

30

【 0 1 6 6 】

実施例 4

E A C に関しては、C p G 島メチル化の表現型 (“ C I M P ”) が存在する明らかな証拠はない

この実施例は、E A C のこの研究に関し、直腸ガンと胃ガンで以前に明らかにされたように (T o y o t a 他、P r o c . N a t l . A c a d . S c i . U S A 、第 9 6 巻、8 6 8 1 ~ 8 6 8 6 ページ、1 9 9 9 年 ; T o y o t a 他、C a n c e r R e s . 、第 5 9 巻、5 4 3 8 ~ 5 4 4 2 ページ、1 9 9 9 年) 、独立した C I M P 腫瘍群が存在する明らかな証拠はないことを示している。しかし E A C における C p G 島の高メチル化が、特定のサンプル中の多数の遺伝子座で実際に起こった。さらに、1 つのサンプル中では、高メチル化された遺伝子座の数が、病気がさまざまな組織学的ステージを経て進行するにつれて増加した (図 6) 。 I M 組織において見られるピークが 2 つある分布 (図 6) は、上に説明したように、異形成またはガンも同時に発生していることで十分に説明できる。

40

【 0 1 6 7 】

材料と方法

組織病理学。組織病理学的分類については、上記の実施例 1 の “ 材料と方法 ” で説明した。

【 0 1 6 8 】

メチル化と統計的分析。メチル化と統計的分析については、実施例 1 で説明したのと同じ

50

である。

【0169】

結果

CIMP分析。直腸ガンと胃ガンの一部ではCpG島メチル化の表現型(“CIMP”)として、単一の腫瘍中の多数の遺伝子座に影響を与える広範囲にわたる異常な高メチル化変化を特徴とする表現型が見られることが以前から報告されている(Toyota他、Proc. Natl. Acad. Sci. USA、第96巻、8681~8686ページ、1999年;Toyota他、Cancer Res.、第59巻、5438~5442ページ、1999年)。これは、一群の腫瘍でメチル化された遺伝子数の頻度が2つのピークを有する分布になることに反映されている(Toyota他、Proc. Natl. Acad. Sci. USA、第96巻、8681~8686ページ、1999年)。CIMP腫瘍は、はっきりと識別できる一群の腫瘍であり、ガンにおいてだけメチル化される遺伝子またはタイプCの遺伝子で対応するCpG島が高い割合で高メチル化されていることを特徴としている。CIMPは現在のところ、腫瘍形成の道筋として、はっきりと識別できる、新規だがすでに主要なものになっていると考えられている(Toyota他、Proc. Natl. Acad. Sci. USA、第96巻、8681~8686ページ、1999年;Toyota他、Cancer Res.、第59巻、5438~5442ページ、1999年)。

【0170】

そこで、食道腺ガンにおいてCpG島メチル化の表現型(CIMP)が見られるかどうかという疑問を調べた。

【0171】

本発明のクラスAの遺伝子は、“タイプC”遺伝子の最もよい具体例となっている。というのも、正常な組織ではメチル化されていないからである。メチル化されたクラスA遺伝子の数の分布をEACについて調べた(図6)。

【0172】

図6は、本発明に従い、食道腺ガンの進行に伴うメチル化の頻度分布を調べたグラフである。データは、4PRMの位置で2つに分けた。4PRM以上だとメチル化されていて、4PRM未満だとメチル化されていないことを意味する。それぞれの組織においてメチル化されたCpG島の数が0~3(クラスA)、0~9(クラスA+D)、0~14(クラスA+B+C+D)であった患者の割合が示してある。クラスEとクラスFのCpG島は含まれていない。というのも、組織間でメチル化の頻度に違いが見られなかったからである。“n”は、それぞれの組織について分析したサンプルの数である。

【0173】

しかし腺ガン組織中では、メチル化された遺伝子の頻度は、CIMPで予想される2つのピークを有する分布を示さなかった(図6)(Toyota他、Cancer Res.、第59巻、5438~5442ページ、1999年)。同様の結果が、クラスAに加えて、やはりタイプCのメチル化を示すクラスDの遺伝子を含めた場合(図6の真中の図)と、クラスA、B、C、Dの遺伝子を合計した場合(図6の右図)に観察された。クラスEとGの遺伝子は、組織のタイプが違っててもメチル化に変化がなかった。

【0174】

クラスA~Dの14個の遺伝子のうちの10個がメチル化されたサンプルが1つあった(図1のケース3と図6)。しかしこのサンプルは、正常な食道粘膜においてメチル化されているためにCIMP表現型を構成する“タイプC”遺伝子の定義にあてはまらないクラスBの遺伝子を含めているから目立っているだけである。

【0175】

したがって、食道腺ガンに関するこの研究では、以前に直腸ガンと胃ガンで見られたような、独立したCIMP腫瘍群が存在する明らかな証拠は見つからなかった。

【0176】

しかしEACにおけるCpG島の高メチル化は、特定のサンプル中の多数の遺伝子座で実際に起こった。さらに、1つのサンプル中では、高メチル化された遺伝子座の数が、病気がさまざまな組織学的段階を経て進行するにつれて増加した(図6)。IM組織において見られるピークが2つある分布(図6)は、上で説明したように、異形成またはガンも同時に発生していることで十分に説明できる。

【0177】

実施例5

アレイまたはマイクロアレイに基づいた応用

マイクロアレイに基づいた実施態様も本発明の範囲に含まれる。例えば、アレイに基づいたそのような1つの実施態様では、ディファレンシャル・メチル化ハイブリダイゼーション(“DMH”)が用いられている(Huang他、Hum. Mol. Genet.、第8巻、459~470ページ、1999年; Yan他、Clin. Cancer Res.、第6巻、1432~1438ページ、2000年)。DMHにより、テストするサンプルと正常なサンプルをペアにしてスクリーニングしたり、特定の後成的変化のパターン(実施例1の“後成的パターン”を参照のこと)が、分析する組織サンプル中の病理学的パラメータと相関しているかどうかを明らかにしたりすることができる。これらサンプルに由来するメチル化されたCpG DNAのプールである“アンプリコン”(同上)は、本発明のCpG島タグを含むアレイ・パネルにおけるハイブリダイゼーション用のプローブとして用いられる。

10

【0178】

したがって、この明細書に開示した20種類の遺伝子配列のうちの19種類(すなわちAPC、ARF、CALCA、CDH1、CDKN2A、CDKN2B、ESR1、GSTP1、HIC1、MGMT、MLH1、MYOD1、RB1、TGFB2、THBS1、TIMP3、CTNNB1、PTGS2、TYMS(上記の表Iと表II;上記の“定義”を参照のこと)またはこれら遺伝子がメチル化によって変化したDNA配列と関係するCpG島配列のうちの1または複数を、アレイまたはマイクロアレイに基づいたアッセイの実施態様におけるCpG島タグとして用いることができる。これら19種類の遺伝子配列領域は、配列番号1~54、58~60、64、65(上記の表IIを参照のこと;配列番号61~63は、本発明の分析で用いるATCB“対照”遺伝子領域に対応する(上記の実施例1を参照のこと))に対応するオリゴヌクレオチド・プライマーまたはプローブによって規定される。関係したCpG島配列とは、(所定のCpG島の一部のメチル化状態が一般にその島全体を代表しているという事実に基づき)これらの特異的オリゴヌクレオチド・プライマーおよびプローブによって規定される配列の少なくとも1つのヌクレオチドを含み、CpGジヌクレオチドの頻度が観測値/予測値の比 > 0.6 に対応することと、GC含量 > 0.5 であることの両方の条件を満たす、ゲノムDNA中の連続配列である。

20

30

【0179】

次に、これらのCpG島タグを固体支持体(例えばナイロン膜、シリコンなど)上に配置し、検査用サンプル(例えば腫瘍)または基準サンプルからの、メチル化された一群のCpG DNAを表わすアンプリコンで調べる。スクリーニングされたCpG島に関する検査用サンプルと基準サンプルの信号強度の差は、検査用DNA中の対応する配列がメチル化により変化したことを反映している。

40

【0180】

得られたデータをこの明細書に開示した後成的パターンと比較すると、診断または予後予測ができる。

【0181】

したがって、この実施態様によれば、アレイまたはマイクロアレイを形成するために固体支持体に固定したCpG島タグの一部におけるパターン分析(上記の実施例1~4を参照のこと)を利用して、ガンが進行するさまざまなステージ(例えば、胃腸異形成および食道異形成、胃腸上皮化生および食道上皮化生、バレット食道、正常な食道扁平粘膜上皮に

50

おける前ガン状態)における進行状態を追試するとともに、食道腺ガンなどの腫瘍の組織学的悪性度またはステージを明らかにすることができる。

【0182】

本発明の他のアレイまたはマイクロアレイの実施態様は、当業者には明らかであろう。そのような実施態様としては、配列番号1~54、58~60、64、65(上記の表IIを参照のこと;配列番号61~63は、本発明の分析で用いるATCB“対照”遺伝子領域に対応する(上記の実施例1を参照のこと))に対応する、APC、ARF、CALCA、CDH1、CDKN2A、CDKN2B、ESR1、GSTP1、HIC1、MGMT、MLH1、MYOD1、RB1、TGFB2、THBS1、TIMP3、CTNNB1、PTGS2、TYMS(上記の表Iと表II;上記の“定義”を参照のこと)に対する特異的プライマーおよび/またはプローブが固体支持体上に配置されているものが挙げられるが、これだけに限定されるわけではない。

10

【0183】

考察

ガンの検出、化学的予測、予後予測に関する新規でより感度の高い方法が、従来技術において要求されている。ガンなどの病気が進行する間に多数の遺伝子座で見られるCpG島のメチル化変化の新規なコーディネート・パターン(すなわち新規な後成的パターン)を明らかにすることが、従来技術において要求されている。特定の腫瘍に特異的で、個々の患者に特異的な後成的パターンまたはフィンガープリントを明らかにすることが、従来技術において要求されている。ガンの処置のための診断法および/または予後予測法で用いることができるバイオマーカーまたはプローブ、例えばEAC特異的なバイオマーカーまたはプローブが、従来技術において要求されている。食道腺ガンがCIMPを示すかどうかを明らかにすることが、従来技術において要求されている。腫瘍のステージを決定する新規な方法が、従来技術において要求されている。本発明は、これらの要求に対処しようとするものである。

20

【0184】

蛍光に基づいたハイスループットのメチル化アッセイ(メチライト(商標))を利用して、食道腺ガン(“EAC”)が進行する間の19のCpG島と1つの非CpG島の新規な高メチル化パターンを調べ、その特徴を明らかにした。そうすることにより、遺伝子を、さまざまなタイプの組織において見られる6通りの後成的パターンに分けた。これは、病気が進行するにつれて明確に異なる非常に多数の組織学的ステージを有するシステムに関してこれまでになされた最も包括的なメチル化検査である。さらに、DNAの異常な高メチル化に関するこの分析は、一般的な制限因子である汚染源となる正常細胞が存在している中での感度が、遺伝子発現分析などの他の方法におけるよりも大きいという顕著な利点を有する。

30

【0185】

この明細書に記載してあるように、DNAの高メチル化は、EACが多段階で進行していくときの初期の後成的変化である。前ガン性の腸上皮化生(“IM”またはバレット食道)は、正常な組織(正常な扁平粘膜上皮)よりもすでに有意に多くメチル化されている。本発明は、いくつかの実施態様において、この腫瘍系においてさらに5つの遺伝子(MYOD1、MGMT、CALCA、TIMP3、HIC1)で頻繁に高メチル化が起こっているという新たな発見を提供している。

40

【0186】

正常な組織のMGMT、TIMP3、HIC1でメチル化が観察されたのは、われわれがメチル化を分析したのが特別な遺伝子領域であったからである可能性がある(Stoger他、Cell、第73巻、61~71ページ、1993年;Larsen他、Hum. Mol. Genet.、第2巻、775~780ページ、1993年;Jones, P. A.、Trends Genet.、第15巻、34~37ページ、1999年)。これら3つの遺伝子は、転写開始部位またはそれよりも下流に位置するCpG島において分析した(表II)。しかしこれではわれわれが観察したCALCAのメチル化を説明で

50

きない。というのも、われわれはこの遺伝子のプロモータ領域を分析したからである。C A L C A のメチル化が低いレベルであることは、A M L 患者の骨髄サンプルで以前に報告されている (M e l k i 他、C a n c e r R e s . , 第 5 9 巻、3 7 3 0 ~ 3 7 4 0 ページ、1 9 9 9 年)。これは、この遺伝子座が、ガン患者の正常な組織の中ではよりメチル化されやすくなっていることを示唆している。

【0187】

特に指摘しておくべきなのは、ステージ I の腫瘍よりも異形成組織で多くメチル化されている状態が、クラス A ($p < 0.0001$) とクラス B ($p = 0.0174$) の両方で見られることである (図 1)。これは、遺伝子の異常 (L O H、欠失、突然変異) が、ステージが進んだ異形成を有するバレット食道には存在するが、隣接する浸潤性 E A C には存在していないという発見 (B a r r e t t 他、N a t . G e n e t . , 第 2 2 巻、1 0 6 ~ 1 0 9 ページ、1 9 9 9 年) と似ている。ステージ I I ~ I V の腫瘍はクラス A の遺伝子に関して異形成と同じ頻度でメチル化されているように見えるため、ステージ I の腫瘍は、実際には、異形成組織よりステージが進んだ腫瘍とは異なる原発巣から進展してきたか、あるいは異形成の後にクローンが増大する間にステージ I I ~ I V の腫瘍とは独立に大きくなってきた可能性があることが示唆される。また、より可能性は小さいものの、ステージ I の腫瘍では、高メチル化の一時的逆転が起こった可能性がある。バレット食道における腫瘍は、化生 - 異形成 - 腫瘍という直線的な多段階経路を通じてクローン化により進行していくことが提案されている (Z h u a n g 他、C a n c e r R e s . , 第 5 6 巻、1 9 6 1 ~ 1 9 6 4 ページ、1 9 9 6 年)。しかし、遺伝子の変化が起こり、本発明により明らかにされたように直線的な順番でない後成的変化が起こるということは、クローン化による E A C の進行が元々予想されていたよりも複雑であることを示唆している (B a r r e t t 他、N a t . G e n e t . , 第 2 2 巻、1 0 6 ~ 1 0 9 ページ、1 9 9 9 年)。同様の観察結果が、膀胱ガンのさまざまなステージについて報告されている (S a l e m 他、C a n c e r R e s . , 第 6 0 巻、2 4 7 3 ~ 2 4 7 6 ページ、2 0 0 0 年)。

10

20

【0188】

ここでの分析では、メチル化された 1 0 種類の遺伝子を有する腫瘍が 1 つあることは別にして、C p G 島メチル化の表現型 (" C I M P ") であることを示唆する、広範囲にわたる一斉メチル化が見られる独立の腫瘍群が存在しているというはっきりとした証拠はなかった。同様の結果が、C I M P (ガンではメチル化され、正常な組織ではメチル化されていない; T o y o t a 他、P r o c . N a t l . A c a d . S c i . U S A , 第 9 6 巻、8 6 8 1 ~ 8 6 8 6 ページ、1 9 9 9 年; T o y o t a 他、C a n c e r R e s . , 第 5 9 巻、5 4 3 8 ~ 5 4 4 2 ページ、1 9 9 9 年) を特徴づける " タイプ C " 遺伝子を調べただけで得られた。興味深いことに、E A C の " タイプ C " 遺伝子は、直腸ガンについて報告されている遺伝子とは異なっている (同上)。例えば、E S R 1 は、老人の正常な結腸上皮でしばしばメチル化されているため、直腸ガンでは " タイプ C " 遺伝子ではなく (老化した正常な組織でメチル化されている) " タイプ A " 遺伝子に分類される (同上)。しかし食道腺ガンでは、E S R 1 は明らかに " タイプ C " 遺伝子のように振る舞う。これは、高メチル化を測定するのに用いる方法の違いか、あるいはそれ以上に組織のタイプの違いが原因となっている可能性がある。

30

40

【0189】

本発明によれば、個々の遺伝子の高メチル化には組織特異的かつ腫瘍特異的な傾向がある。例えば、A P C は正常な胃で高メチル化されているが、正常な食道粘膜では高メチル化されていない。高メチル化が腫瘍特異的であることは、胃ガンと肺ガンでしばしば高メチル化されている遺伝子 T G F B R 2、網膜芽繊維症でしばしば高メチル化されている R B 1 というクラス E の 2 つの遺伝子でメチル化が検出されないことに現われている (S t i r z a k e r 他、C a n c e r R e s . , 第 5 7 巻、2 2 2 9 ~ 2 2 3 7 ページ、1 9 9 7 年; H o u g a a r d 他、B r . J . C a n c e r , 第 7 9 巻、1 0 0 5 ~ 1 0 1 1 ページ、1 9 9 9 年)。

50

【0190】

CpG島の高メチル化が腫瘍特異性を有するという事実は、食道腺ガンと他のタイプの腫瘍では異なっていて、腫瘍発生の間にこれらCpG島のメチル化の変化を制御している腫瘍特異的なトランス作用因子が存在する可能性があることを示唆している。また、食道腺ガンにおいてDNAのメチル化によりこれら遺伝子が不活性になることには、特別な利点がない可能性がある。これがあてはまる可能性のある2つのシナリオがある。1つは、問題の遺伝子が別の遺伝的メカニズムによって不活化されていて、高メチル化にはもはや特別な利点がない場合である。もう1つは、遺伝子がこの特別な腫瘍系で腫瘍抑制において役割を果たしていない場合である。

【0191】

DNAのメチル化状態が変化するのは腫瘍発生において一般的であるが、裏に隠れているメカニズムははっきりしていない。異常なメチル化が、少なくとも直腸ガンではDNAメチルトランスフェラーゼ遺伝子の単なるアップレギュレーションにはよらないというのは、他の主要なプレイヤーが関与していることを示唆している(Eads他、Cancer Res.、第59巻、2302~2306ページ、1999年)。本発明は、こうした異常なメチル化の裏にあるプロセスを初めて覗き見るものである。

【0192】

本発明によれば、機能が互いに無関係のさまざまな遺伝子は、メチル化の変化に関しては、EACが進行しているさまざまな組織において、明確に異なるクラスの遺伝子として機能することができる。CpG島の高メチル化は、(確率的な要素があるにもかかわらず) 確率的なプロセスではなく、複数の異なる変化を含む段階的なプロセスであるらしい。これは、CpG島が高メチル化されないようにするいくつかの異なるメカニズムが存在していることと両立する。このシナリオでは、異なるCpG島で一斉に見られる変化は、病気進行のさまざまな段階において異なるタイプの保護要素が失われた結果であろう。この発見は、遺伝子に対するCpG島の位置に依存しているようには見えない。というのも、プロモーターと内部CpG島の両方がすべての遺伝子クラスで観察されたからである。この分析において、GC含量%、CpGの観測値/予測値の比、CpG:GpC比を分析することによりこれらCpG島の構造上の特徴も調べたが、遺伝子クラスとの関係は見いだせなかった(表II)。

【0193】

本発明によれば、IMまたはNEのサンプル自体は、付随する異形成またはガンがある場合もない場合も、組織学的には識別できなかったが、分子レベルでは識別できた。離れた位置に異形成またはガンも同時に有する人に由来するNEサンプルとIMサンプルでは、CpG島の高メチル化が統計的に高い頻度で起こっている。この発見は、この明細書において、完全に独立な別の研究の中でIM組織において確認した。これは、後成的マーカー、特にクラスAとBの遺伝子を病気のスクリーニング・ツールとして、また病気がさらに進行していくことの予測をするマーカーとして用いる際の強い支持材料となる。

【0194】

本発明のメチル化パターンは、ガンを初期の段階で検出するための方法と組成物を提供する。ガンがあるかガンになるリスクが高いかを正常な組織および/または前ガン状態の組織を用いて明らかにするこのような分子診断法は、初期治療の機会を提供する。さらに、診断または予後予測のためのマーカーとしてCpG島の高メチル化を用いることの利点は、汚染源となる正常細胞を含むサンプル中では遺伝子発現のないこと(例えばLOHや欠失の分析)を明らかにするのが難しいのと異なり、高メチル化をプラスの信号として汚染源となる正常細胞中で容易に検出できることにある。

【0195】

要約

本発明によれば、研究した19のCpG島(表Iと表II)は、さまざまなタイプの組織で6通りの後成的パターンに分類される。それぞれの分類クラスは、EACが進行していくさまざまな段階で独特の後成的変化をする。これらメチル化パターンは、ガンを初期段

10

20

30

40

50

階で検出するための方法と組成物を提供する。

【図面の簡単な説明】

【図 1】

さまざまな段階にあるバレット食道（“IM”）、異形成（“DYS”）および/またはそれに付随する食道腺ガン（“T”）の患者31人から採取した84個の組織サンプル中の20種類の遺伝子について、メチル化を本発明に従って定量分析した図である。メチル化の分析は、メチライト（商標）アッセイを利用して行なった（Eads他、Cancer Res.、第59巻、2302～2306ページ、1999年；Eads他、Nucleic Acids Res.、第28巻、E32ページ、2000年）。特定の遺伝子座において十分にメチル化された分子の割合（PMR = メチル化された基準の割合）は、サンプルの遺伝子/ACTB比を、SssIで処理した精子DNAの遺伝子/ACTB比で割り、その数値を100倍することによって計算した。次に、（この明細書に説明してあるように）得られた数値を4%PRMの位置で2つに分け、グラフを描きやすくするとともに、組織特異的パターンが明らかになるようにした。“N”は、その特定のサンプルの分析におけるメチル化反応については、対照とした遺伝子ACTBが、1PMRという最小値を検出できるほど十分なレベルに達していなかったことを意味している。

10

【図 2】

各遺伝子についてメチル化されたサンプルの割合を組織のタイプごとに示した図である。データは、4PRMの位置で2つに分けた。4PRM以上だとメチル化されていて、4PRM未満だとメチル化されていないことを意味する。遺伝子は、本発明により、図1に示したようにそれぞれの後成的遺伝子クラス（A～G）に従ってグループ分けした。“n”は、各組織について分析したサンプルの数である。

20

【図 3】

本発明による後成的パターンを比較した図である。データは、4PRMの位置で2つに分けた。4PRM以上だとメチル化されていて、4PRM未満だとメチル化されていないことを意味する。誤差を示す棒は、平均値の標準偏差を表わしている。上方の図：各遺伝子クラス（A～F、または19あるCpG島の合計）においてメチル化された遺伝子の割合の平均値を組織のタイプ別（N、正常な食道；S、胃；IM、腸上皮化生；DYS、異形成；T、腺ガン）に示したもの。誤差棒は、平均値の標準偏差（SEM）を示す。下方の図：遺伝子の各クラス（A～F）と、19のCpG島をすべて合わせた場合（合計）につき、さまざまな組織でメチル化された遺伝子の割合の平均値の違いを統計的に分析した結果。p値は、フィッシャーの保護された最小有意差（PLSD）検定をサンプル数が均等でない場合に使用できるようにして求めた（SASスタットビュー（Statview：商標）ソフトウェア）。

30

【図 4】

クラスA遺伝子のメチル化の頻度と腫瘍のステージの関係を、本発明に従って示した図である。データは、4PRMの位置で2つに分けた。4PRM以上だとメチル化されていて、4PRM未満だとメチル化されていないことを意味する。上方の図：メチル化されたクラスA遺伝子の平均数を腫瘍のステージ（I～IV）ごとに示したもの（図1を参照のこと）。誤差棒は、平均値の標準偏差（SEM）を示す。“n”は、腫瘍の各ステージについて分析したサンプルの数である。下方の図：腫瘍のステージそれぞれにつき、メチル化された遺伝子の数の平均値の違いを統計的に分析したもの。p値は、フィッシャーの保護された最小有意差（PLSD）検定をサンプル数が均等でない場合に使用できるようにして求めた（SASスタットビュー（商標）ソフトウェア）。

40

【図 5】

本発明に従い、腸上皮化生（“IM”）において、それに付随する異形成および/または腺ガンを伴っている場合（“Y”）または伴っていない場合（“N”）について、クラスAの2つ以上の遺伝子がメチル化されている割合を示している。データは、4PRMの位置で2つに分けた。4PRM以上だとメチル化されていて、4PRM未満だとメチル化されていないことを意味する。左側の図：図1の中のIMに関するデータの中でメチル化さ

50

れているクラスA遺伝子の割合。右側の図：顕微解剖した20の異なるIMサンプルに関する完全に独立な追試研究において、IM中でメチル化されているクラスA遺伝子の割合。誤差棒は、平均値の標準偏差(SEM)を示す。“n”は、それぞれの組織群について分析したサンプルの数である。

【図6】

本発明に従い、食道腺ガンの進行に伴うメチル化の頻度分布を調べたグラフである。データは、4PRMの位置で2つに分けた。4PRM以上だとメチル化されていて、4PRM未満だとメチル化されていないことを意味する。それぞれの組織においてメチル化されたCpG島の数が0~3(クラスA)、0~9(クラスA+D)、0~14(クラスA+B+C+D)であった患者の割合が示してある。クラスEとクラスFのCpG島は含まれていない。というのも、組織間でメチル化の頻度に違いが見られなかったからである。“n”は、それぞれの組織について分析したサンプルの数である。

10

【図1】

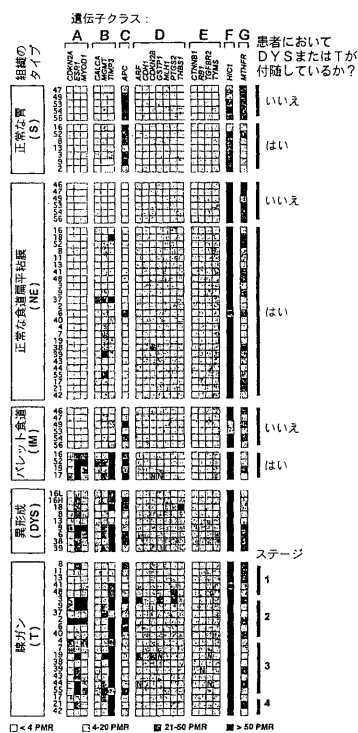


図 1

【図2】

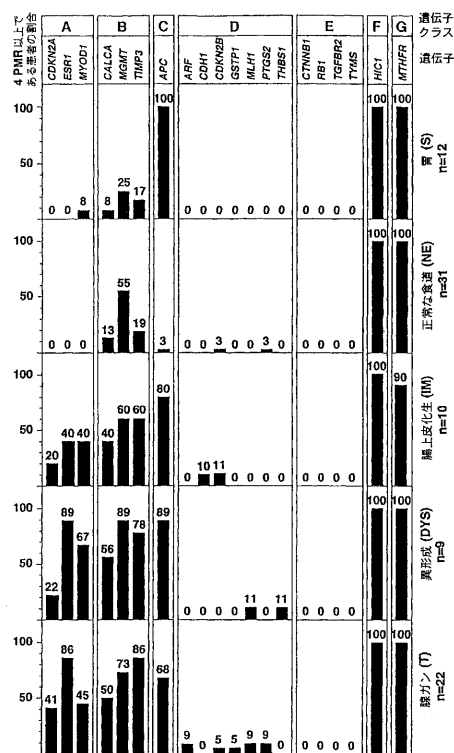


図 2

【 図 3 】

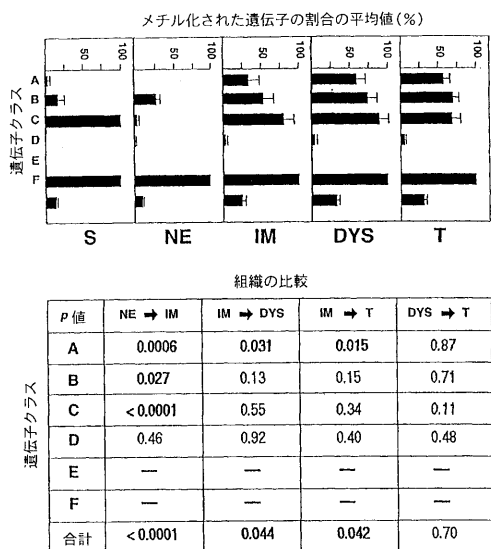


図 3

【 図 4 】

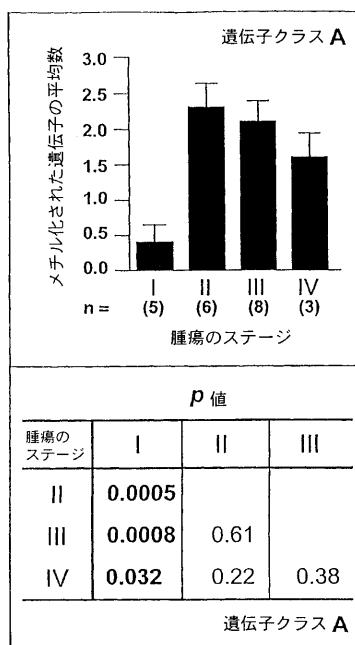


図 4

【 図 5 】

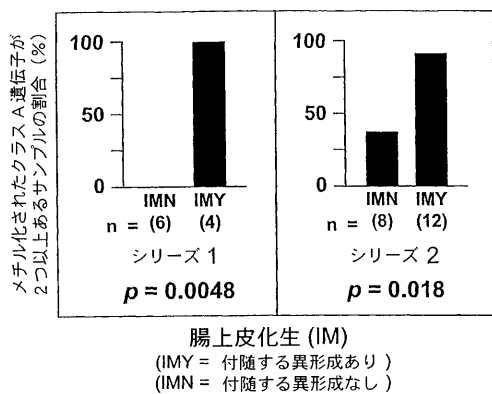


図 5

【 図 6 】

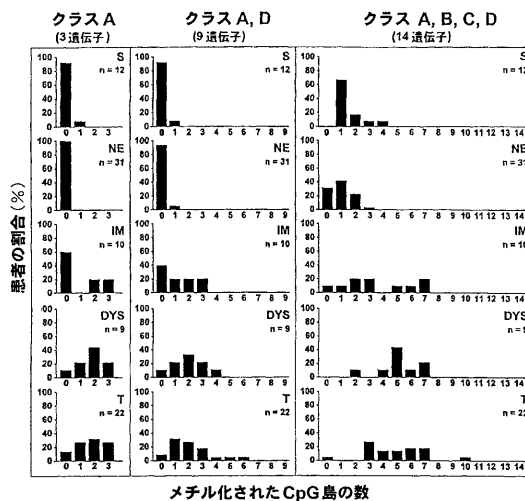


図 6

【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
11 October 2001 (11.10.2001)

PCT

(10) International Publication Number
WO 01/75172 A1

- (51) International Patent Classification: C12Q 1/68
- (21) International Application Number: PCT/US01/10658
- (22) International Filing Date: 2 April 2001 (02.04.2001)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
60/193,839 31 March 2000 (31.03.2000) US
- (71) Applicant: UNIVERSITY OF SOUTHERN CALIFORNIA [US/US]; 3716 Hope Street, Los Angeles, CA 98007-4344 (US).
- (72) Inventors: LAIRD, Peter; 649 Forrest Avenue, South Pasadena, CA 91010 (US); EADS, Cindy; 635 Prospect Street #B15, South Pasadena, CA 91010 (US).
- (74) Agent: DAVISON, Barry, L.; Davis Wright Tremaine, LLP, 2600 Century Square, 1501 Fourth Avenue, Seattle, WA 98101-1688 (US).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
- Published:
— with 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
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



WO 01/75172 A1

(54) Title: EPIGENETIC SEQUENCES FOR ESOPHAGEAL ADENOCARCINOMA

(57) Abstract: There is disclosed a diagnostic or prognostic assay for cancer, particularly gastrointestinal and esophageal adenocarcinoma. Specifically, the present invention provides a methylation pattern that can be assayed by standard methylation assays of CpG islands, including which genes are hypermethylated and which genes are unmethylated in gastrointestinal and esophageal adenocarcinomas, Barrett's esophagus, and normal squamous mucosa.

WO 01/75172

PCT/US01/10658

EPIGENETIC SEQUENCES FOR ESOPHAGEAL ADENOCARCINOMA

CROSS-REFERENCE TO RELATED APPLICATIONS

- 5 This application claims priority to United States *Provisional* Patent Application Serial No. 06/193,839, entitled EPIGENETIC SEQUENCES FOR ESOPHAGEAL ADENOCARCINOMA, filed 31 March 2000.

STATEMENT REGARDING FEDERALLY-SPONSORED RESEARCH

- 10 This work was supported by NIH/NCI grant R01 CA 75090 to P.W.L. The United States has certain rights in this invention, pursuant to 35 U.S.C. § 202(c)(6).

TECHNICAL FIELD OF THE INVENTION

- 15 The present invention provides a diagnostic or prognostic assay for gastrointestinal adenocarcinoma, and particularly esophageal adenocarcinoma ("EAC"). Specifically, the present invention provides a multi-genic *epigenetic fingerprint* or methylation pattern, that can be assayed by standard methylation assays of CpG island methylation status, and that comprises the relative methylation status of two or more genes in gastrointestinal carcinomas, normal squamous cells, and EAC.

20

BACKGROUND OF THE INVENTION

- DNA methylation and cancer.* DNA methylation patterns are frequently altered in human cancers. These methylation changes include genome-wide hypomethylation as well as regional hypermethylation (Jones & Laird, *Nat Genet.* 21:163-167, 1999).
- 25 Aberrant hypermethylation in cancer cells often occurs at CpG islands, which are generally protected from methylation in normal tissues. Hypermethylation of promoter CpG islands (that is, CpG islands located in promoter regions of genes) has been associated with transcriptional silencing in many types of human cancers.

- Methylation patterns of genes can provide different types of useful information about a cancer cell. First, each tumor type (*i.e.*, breast, colon, esophagus, etc.) has a characteristic set of genes with an increased propensity to become methylated (Costello et al., *Nat. Genet.* 24:132-138, 2000). For example, *RBI* is known to be hypermethylated in retinoblastoma (Stirzaker et al., *Cancer Res.* 57:2229-2237, 1997; Sakai et al., *Am. J. Hum. Genet.* 48:880-888, 1991), but not in acute myelogenous leukemia (Komblau & Qiu, *Leuk. Lymphoma.* 35:283-288, 1999; Melki et al., *Cancer Res.* 59:3730-3740, 1999).

- 30 Second, an individual tumor within a single patient has a unique epigenetic fingerprint reflective of the evolution of that tumor as compared to a tumor of the same type in a different patient (Costello et al., *Nat. Genet.* 24:132-138, 2000).

- 40 Generally, however, most studies of epigenetic alterations in cancer have focused primarily on either a very small set of known genes (Jones & Laird, *Nat Genet.* 21:163-

WO 01/75172

PCT/US01/10658

167, 1999; Baylin & Herman, *Trends Genet.* 16:168-174, 2000) or on the global analysis of unknown CpG islands (Costello et al., *Nat. Genet.* 24:132-138, 2000), and thus do not provide a suitable diagnostic and/or prognostic framework.

Esophageal adenocarcinoma ("EAC"). Esophageal adenocarcinoma ("EAC") arises from a multistep process whereby normal squamous mucosa undergoes metaplasia to specialized columnar epithelium (Intestinal Metaplasia (IM) or Barrett's esophagus), which then ultimately progresses to dysplasia and subsequent malignancy (Barrett et al., *Nat. Genet.* 22:106-109, 1999; Zhuang et al., *Cancer Res.* 56:1961-4, 1996). The incidence of EAC has increased rapidly in the Western World over the past three decades (Devesa et al., *Cancer.* 83:2049-2053, 1998; Jankowski et al., *Am. J. Pathol.* 154:965-973, 1999).

Unfortunately, epigenetic studies of this model have so far been limited to the DNA methylation analysis of a few genes (Wong et al., *Cancer Res.* 57:2619-2622, 1997; Klump et al., *Gastroenterology.* 115:1381-1386, 1998; Eads et al., *Cancer Res.* 60:5021-5026, 2000).

CpG island methylator phenotype ("CIMP"). It has previously been reported that a subset of colorectal and gastric tumors display a CpG island methylator phenotype ("CIMP"), characterized by widespread, aberrant hypermethylation changes affecting multiple loci in a single tumor (Toyota et al., *Proc. Natl. Acad. Sci. USA* 96:8681-8686, 1999; Toyota et al., *Cancer Res.* 59:5438-5442, 1999). This is reflected in a bimodal distribution of the frequency of the number of genes methylated in a group of tumors (Toyota et al., *Proc. Natl. Acad. Sci. USA* 96:8681-8686, 1999). CIMP tumors are a distinct group of tumors that are defined by a high degree of concordant CpG island hypermethylation of genes exclusively methylated in cancer, or type C genes. CIMP is now thought to be a new, distinct, yet major pathway of tumorigenesis (Toyota et al., *Proc. Natl. Acad. Sci. USA* 96:8681-8686, 1999; Toyota et al., *Cancer Res.* 59:5438-5442, 1999).

However, the role, if any, of the CIMP pathway in the tumor evolution of EAC is still uncharacterized, because the previous epigenetic studies only analyzed one (Wong et al., *Cancer Res.* 57:2619-2622, 1997; Klump et al., *Gastroenterology.* 115:1381-1386, 1998) or a few genes (Eads et al., *Cancer Res.* 60:5021-5026, 2000).

Therefore, there is a need in the art for novel methods of cancer detection, chemoprediction and prognostics. There is a need in the art to define novel coordinate patterns of CpG island methylation changes at multiple loci during different steps of a disease, such as cancer. There is a need in the art to determine tumor-type-specific, and patient-specific epigenetic patterns or fingerprints. There is a need in the art to provide biomarkers or probes, such as EAC-specific biomarkers or probes, that can be used in diagnostic and/or prognostic methods for the treatment of cancer. There is a need in the art to determine whether esophageal adenocarcinoma displays a CIMP. There is a need in

WO 01/75172

PCT/US01/10658

the art for novel methods for determining the stage of a tumor. The present invention addresses these needs.

SUMMARY OF THE INVENTION

5 The present invention provides a method for diagnosing cancer or cancer-related conditions from tissue samples, comprising: (a) obtaining a tissue sample from a test tissue or region to be diagnosed; (b) performing a methylation assay of the tissue sample, wherein the methylation assay determines the methylation state of genomic CpG sequences, wherein the genomic CpG sequences are located within at least one gene sequence selected from the group consisting of *APC*, *ARF*, *CALCA*, *CDHI*, *CDKN2A*, *CDKN2B*, *ESR1*, *GSTP1*, *HIC1*, *MGMT*, *MLH1*, *MYOD1*, *RBI*, *TGFBR2*, *THBS1*, *TIMP3*, *CTNNB1*, *PTGS2*, *TYMS* and *MTHFR*, and combinations thereof; and (c) making a diagnostic or prognostic prediction of the cancer based, at least in part, upon the methylation state of the genomic CpG sequences. Preferably, the genomic CpG sequences located within at least one gene sequence selected from the group consisting of *APC*, *ARF*, *CALCA*, *CDHI*, *CDKN2A*, *CDKN2B*, *ESR1*, *GSTP1*, *HIC1*, *MGMT*, *MLH1*, *MYOD1*, *RBI*, *TGFBR2*, *THBS1*, *TIMP3*, *CTNNB1*, *PTGS2* and *TYMS*, correspond to genomic CpG sequences of CpG islands. Preferably, the *APC*, *ARF*, *CALCA*, *CDHI*, *CDKN2A*, *CDKN2B*, *ESR1*, *GSTP1*, *HIC1*, *MGMT*, *MLH1*, *MYOD1*, *RBI*, *TGFBR2*, *THBS1*, *TIMP3*, *CTNNB1*, *PTGS2*, *TYMS* and *MTHFR* gene sequences are those defined by the specific oligonucleotide primers and probes corresponding to SEQ ID Nos:1-60, 64 and 65, as listed in TABLE II, or portions thereof. Preferably, the CpG islands are located within the promoter regions of the genes. Preferably, the *APC*, *ARF*, *CALCA*, *CDHI*, *CDKN2A*, *CDKN2B*, *ESR1*, *GSTP1*, *HIC1*, *MGMT*, *MLH1*, *MYOD1*, *RBI*, *TGFBR2*, *THBS1*, *TIMP3*, *CTNNB1*, *PTGS2*, and *TYMS* gene sequences correspond to any CpG island sequences associated with the sequences defined by the specific oligonucleotide primers and probes corresponding to SEQ ID Nos:1-54, 58-60, 64 and 65, as listed in TABLE II, or portions thereof, wherein the associated CpG island sequences are those contiguous sequences of genomic DNA that encompass at least one nucleotide of the sequences defined by the specific oligonucleotide primers and probes corresponding to SEQ ID Nos:1-54, 58-60, 64 and 65, and satisfy the criteria of having both a frequency of CpG dinucleotides corresponding to an Observed/Expected Ratio >0.6, and a GC Content >0.5.

Preferably, the genomic CpG sequences are located within at least one gene sequence selected from the group consisting of *APC*, *CDKN2A*, *MYOD1*, *CALCA*, *ESR1*, *MGMT* and *TIMP3*, and combinations thereof. Preferably, the genomic CpG sequences located within at least one gene sequence selected from the group consisting of *APC*, *CDKN2A*, *MYOD1*, *CALCA*, *ESR1*, *MGMT* and *TIMP3*, correspond to genomic CpG sequences of CpG islands. Preferably, the *APC*, *CDKN2A*, *MYOD1*, *CALCA*, *ESR1*, *MGMT* and *TIMP3* gene sequences are those defined by the specific oligonucleotide

WO 01/75172

PCT/US01/10658

primers and probes corresponding to SEQ ID NOs:19-21, SEQ ID NOs:1-3, SEQ ID NOs:7-9, SEQ ID NOs:10-12, SEQ ID NOs:4-6, SEQ ID NOs:16-18 and SEQ ID NOs:13-15, respectively, as listed in TABLE II. Preferably, the CpG islands are located within the promoter regions of the genes. Preferably, the *APC*, *CDKN2A*, *MYOD1*,
 5 *CALCA*, *ESR1*, *MGMT* and *TIMP3* gene sequences correspond to any CpG island sequences associated with the sequences defined by the specific oligonucleotide primers and probes corresponding to SEQ ID NOs:19-21, SEQ ID NOs:1-3, SEQ ID NOs:7-9, SEQ ID NOs:10-12, SEQ ID NOs:4-6, SEQ ID NOs:16-18 and SEQ ID NOs:13-15, respectively, as listed in TABLE II, or portions thereof, wherein the associated CpG island
 10 sequences are those contiguous sequences of genomic DNA that encompass at least one nucleotide of the sequences defined by the specific oligonucleotide primers and probes corresponding to SEQ ID NOs:19-21, SEQ ID NOs:1-3, SEQ ID NOs:7-9, SEQ ID NOs:10-12, SEQ ID NOs:4-6, SEQ ID NOs:16-18 and SEQ ID NOs:13-15, and satisfy the criteria of having both a frequency of CpG dinucleotides corresponding to an
 15 Observed/Expected Ratio >0.6, and a GC Content >0.5.

Preferably, the cancer or cancer-related condition is selected from the group consisting of gastrointestinal or esophageal adenocarcinoma, gastrointestinal or esophageal dysplasia, gastrointestinal or esophageal metaplasia, Barrett's intestinal tissue, pre-cancerous conditions in normal esophageal squamous mucosa, and combinations
 20 thereof. Preferably, the cancer is esophageal adenocarcinoma, and wherein making a diagnostic or prognostic prediction of the cancer, based upon the methylation state of the genomic CpG sequences provides for classification of the adenocarcinoma by grade or stage.

Preferably, the methylation assay used to determine the methylation state of
 25 genomic CpG sequences is selected from the group consisting of "MethylLightTM", MS-SNuPE, MSP, COBRA, MCA, and DMH, and combinations thereof.

Preferably, the methylation assay used to determine the methylation state of genomic CpG sequences is based, at least in part, on an array or microarray comprising CpG sequences located within at least one gene sequence selected from the group
 30 consisting of *APC*, *ARF*, *CALCA*, *CDH1*, *CDKN2A*, *CDKN2B*, *ESR1*, *GSTP1*, *HIC1*, *MGMT*, *MLH1*, *MYOD1*, *RB1*, *TGFBR2*, *THBS1*, *TIMP3*, *CTNNB1*, *PTGS2*, *TYMS* and *MTHFR*. Preferably, the *APC*, *ARF*, *CALCA*, *CDH1*, *CDKN2A*, *CDKN2B*, *ESR1*, *GSTP1*, *HIC1*, *MGMT*, *MLH1*, *MYOD1*, *RB1*, *TGFBR2*, *THBS1*, *TIMP3*, *CTNNB1*, *PTGS2*, and *TYMS* gene sequences correspond to any CpG island sequences associated with the
 35 sequences defined by the specific oligonucleotide primers and probes corresponding to SEQ ID Nos:1-54, 58-60, 64 and 65, as listed in TABLE II, or portions thereof, wherein the associated CpG island sequences are those contiguous sequences of genomic DNA that encompass at least one nucleotide of the sequences defined by the specific oligonucleotide primers and probes corresponding to SEQ ID Nos:1-54, 58-60, 64 and 65, and satisfy the

WO 01/75172

PCT/US01/10658

criteria of having both a frequency of CpG dinucleotides corresponding to an Observed/Expected Ratio >0.6, and a GC Content >0.5. Preferably, the *APC*, *ARF*, *CALCA*, *CDHI*, *CDKN2A*, *CDKN2B*, *ESR1*, *GSTP1*, *HIC1*, *MGMT*, *MLH1*, *MYOD1*, *RBI*, *TGFBR2*, *THBS1*, *TIMP3*, *CTNBN1*, *PTGS2*, *TYMS* and *MTHFR* gene sequences are those defined by, or correspond to the specific oligonucleotide primers and probes corresponding to SEQ ID Nos:1-60, 64 and 65, as listed in TABLE II, or portions thereof.

Preferably, the methylation state of genomic CpG sequences that is determined is that of hypermethylation, hypomethylation or normal methylation.

The present invention also provides a kit useful for diagnosis or prognosis of cancer or cancer-related conditions, comprising a carrier means containing one or more containers comprising: (a) a container containing a probe or primer which hybridizes to any region of a sequence located within at least one gene sequence selected from the group consisting of *APC*, *ARF*, *CALCA*, *CDHI*, *CDKN2A*, *CDKN2B*, *ESR1*, *GSTP1*, *HIC1*, *MGMT*, *MLH1*, *MYOD1*, *RBI*, *TGFBR2*, *THBS1*, *TIMP3*, *CTNBN1*, *PTGS2*, *TYMS* and *MTHFR*; and (b) additional standard methylation assay reagents required to affect detection of methylated CpG-containing nucleic acid based, at least in part, on the probe or primer. Preferably, the additional standard methylation assay reagents are standard reagents for performing a methylation assay from the group consisting of MethyLight™, MS-SNuPE, MSP, COBRA, MCA and DMH, and combinations thereof. Preferably, the probe or primer comprises at least about 12 to 15 nucleotides of a sequence selected from the group consisting of SEQ ID Nos:1-60, 64 and 65, as listed in TABLE II.

The present invention further provides a kit useful for diagnosis or prognosis of cancer or cancer-related conditions, comprising a carrier means containing one or more containers comprising: (a) an array or microarray comprising sequences of at least about 12 to 15 nucleotides of a sequence selected from the group consisting of SEQ ID Nos:1-60, 64, 65, and any sequence located within a CpG island sequence associated with SEQ ID NOs:1-54, 58-60, 64 and 65.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows, according to the present invention, a quantitative methylation analysis of a panel of 20 genes from a screen of 84 tissue specimens from 31 patients with different stages of Barrett's esophagus ("IM"), dysplasia ("DYS") and/or associated esophageal adenocarcinoma ("T"). Methylation analysis was performed using the MethyLight™ assay (Eads et al., *Cancer Res.* 59:2302-2306, 1999; Eads et al., *Nucleic Acids Res.* 28:E32, 2000). The percentage of fully methylated molecules at a specific locus (PMR = Percent of Methylated Reference) was calculated by dividing the *GENE/ACTB* ratio of a sample by the *GENE/ACTB* ratio of *SssI*-treated sperm DNA and multiplying by 100. The resulting percentages were then dichotomized at 4% PMR to facilitate graphical representation and to reveal tissue-specific patterns (as described

WO 01/75172

PCT/US01/10658

herein). "N" indicates an analysis for which the control gene *ACTB* did not reach sufficient levels to allow the detection of a minimal value of 1 PMR for that methylation reaction in that particular sample.

Figure 2 shows the percent of samples methylated for each gene by tissue type.

- 5 The data was dichotomized at 4 PMR, with 4 PMR and higher designated as methylated, and below 4 PMR as unmethylated. The genes, according to the present invention, were grouped according to their respective epigenetic gene classes (A-G) as shown in Figure 1. The letter "n" equals the number of samples analyzed for each tissue.

- 10 Figure 3 shows a comparison of epigenetic profiles according to the present invention. The data was dichotomized at 4 PMR, with 4 PMR and higher designated as methylated, and below 4 PMR as unmethylated. Error bars represent the standard error of the mean. *Top panel:* Mean percent of genes methylated in each gene Class (A-F or ALL 19 CpG islands) by tissue type (N, normal esophagus; S, stomach; IM, intestinal metaplasia; DYS, dysplasia; T, adenocarcinoma). The error bars represent the standard error of the mean (SEM). *Bottom panel:* Statistical analysis of the difference in mean percent of genes methylated in different tissues by gene Class (A-F) or for all 19 CpG islands combined (ALL). The *p*-values were generated by a Fisher's Protected Least Significant Difference (PLSD) test, adapted for use with unequal sample numbers (SAS Statview™ software).

- 20 Figure 4 shows the relationship between Class A methylation frequency and tumor stage according to the present invention. The data was dichotomized at 4 PMR, with 4 PMR and higher designated as methylated, and below 4 PMR as unmethylated. *Upper panel:* Mean number of genes methylated for Class A with respect to tumor stage (I-IV) is shown (see Figure 1). The error bars represent the standard error of the mean (SEM). The letter "n" equals the number of samples analyzed in each tumor stage. *Lower panel:* Statistical analysis of the difference in mean number of Class A genes methylated by tumor stage. The *p*-values were generated by a Fisher's Protected Least Significant Difference (PLSD) test, adapted for use with unequal sample numbers (SAS Statview™ software).

- 30 Figure 5 shows, according to the present invention, the percent of two or more Class A genes methylated in intestinal metaplasia ("IM") tissues with ("Y"), or without ("N") associated dysplasia and/or adenocarcinoma. The data was dichotomized at 4 PMR, with 4 PMR and higher designated as methylated, and below 4 PMR as unmethylated. *Left panel:* Class A methylation in the IM data illustrated in Figure 1. *Right panel:* Class A methylation in the IM for a completely independent follow-up study of twenty different microdissected IM samples. The error bars represent the standard error of the mean (SEM). The letter "n" equals the number of samples analyzed in each tissue group.

Figure 6 shows, according to the present invention, methylation frequency distributions in the progression of esophageal adenocarcinoma. The data was

WO 01/75172

PCT/US01/10658

dichotomized at 4 PMR, with 4 PMR and higher designated as methylated, and below 4 PMR as unmethylated. The proportion of patients with zero to three (Class A), zero to nine (Classes A + D) and zero to fourteen CpG islands (Classes A + B + C + D) methylated in each tissue is shown. Class E and F CpG islands were not included since there was no variation in the frequency of methylation between the different tissue. The letter "n" equals the number of samples analyzed in each tissue.

DETAILED DESCRIPTION OF THE INVENTION

10

Definitions:

The term "EAC" refers to esophageal adenocarcinoma, but also encompasses different histological stages of esophageal adenocarcinoma corresponding to a multistep process whereby normal squamous mucosa undergoes metaplasia to specialized columnar epithelium (Intestinal Metaplasia (IM) or Barrett's esophagus), which then ultimately progresses to dysplasia and subsequent malignancy (Barrett et al., *Nat. Genet.* 22:106-109, 1999; Zhuang et al., *Cancer Res.* 56:1961-4, 1996);

The term "CIMP" refers to CpG island methylator phenotype, characterized by widespread aberrant hypermethylation changes affecting multiple loci in a single tumor. This is reflected in a bimodal distribution of the frequency of the number of genes methylated in a group of tumors (16). CIMP tumors are a distinct group of tumors that are defined by a high degree of concordant CpG island hypermethylation of genes exclusively methylated in cancer, or type C genes. CIMP is now thought to be a new, distinct, yet major pathway of tumorigenesis (Toyota et al., *Proc. Natl. Acad. Sci. USA* 96:8681-8686, 1999; Toyota et al., *Cancer Res.* 59:5438-5442, 1999) (see "Background," above);

The term "PMR" refers to percent of methylated reference, and is calculated as described herein under Example I;

"GC Content" refers, within a particular DNA sequence, to the [(number of C bases + number of G bases) / band length for each fragment];

"Observed/Expected Ratio" ("O/E Ratio") refers to the frequency of CpG dinucleotides within a particular DNA sequence, and corresponds to the [(number of CpG sites / (number of C bases X number of G bases)) X band length for each fragment];

"CpG Island" refers to a contiguous region of genomic DNA that satisfies the criteria of (1) having a frequency of CpG dinucleotides corresponding to an "Observed/Expected Ratio" >0.6, and (2) having a "GC Content" >0.5. CpG islands are typically, but not always, between about 0.2 to about 1 kb in length. A CpG island sequence associated with a particular SEQ ID NO sequence of the present invention is that contiguous sequence of genomic DNA that encompasses at least one nucleotide of the particular SEQ ID NO sequence, and satisfies the criteria of having both a frequency of

- CpG dinucleotides corresponding to an Observed/Expected Ratio >0.6), and a GC Content >0.5;
- “Methylation state” refers to the presence or absence of 5-methylcytosine (“5-mCyt”) at one or a plurality of CpG dinucleotides within a DNA sequence;
- 5 “Hypermethylation” refers to the methylation state corresponding to an *increased* presence of 5-mCyt at one or a plurality of CpG dinucleotides within a DNA sequence of a test DNA sample, relative to the amount of 5-mCyt found at corresponding CpG dinucleotides within a normal control DNA sample;
- 10 “Hypomethylation” refers to the methylation state corresponding to a *decreased* presence of 5-mCyt at one or a plurality of CpG dinucleotides within a DNA sequence of a test DNA sample, relative to the amount of 5-mCyt found at corresponding CpG dinucleotides within a normal control DNA sample;
- “Methylation assay” refers to any assay for determining the methylation state of a CpG dinucleotide within a sequence of DNA;
- 15 “MS.AP-PCR” (Methylation-Sensitive Arbitrarily-Primed Polymerase Chain Reaction) refers to the art-recognized technology that allows for a global scan of the genome using CG-rich primers to focus on the regions most likely to contain CpG dinucleotides, and described by Gonzalgo et al., *Cancer Research* 57:594-599, 1997;
- 20 “MethyLight” refers to the art-recognized fluorescence-based real-time PCR technique described by Eads et al., *Cancer Res.* 59:2302-2306, 1999;
- “Ms-SNuPE” (Methylation-sensitive Single Nucleotide Primer Extension) refers to the art-recognized assay described by Gonzalgo & Jones, *Nucleic Acids Res.* 25:2529-2531, 1997;
- 25 “MSP” (Methylation-specific PCR) refers to the art-recognized methylation assay described by Herman et al. *Proc. Natl. Acad. Sci. USA* 93:9821-9826, 1996, and by US Patent No. 5,786,146;
- “COBRA” (Combined Bisulfite Restriction Analysis) refers to the art-recognized methylation assay described by Xiong & Laird, *Nucleic Acids Res.* 25:2532-2534, 1997;
- 30 “MCA” (Methylated CpG Island Amplification) refers to the methylation assay described by Toyota et al., *Cancer Res.* 59:2307-12, 1999, and in WO 00/26401A1;
- “DMH” (Differential Methylation Hybridization) refers to the art-recognized methylation assay described in Huang et al., *Hum. Mol. Genet.*, 8:459-470, 1999, and in Yan et al., *Clin. Cancer Res.* 6:1432-38, 2000;
- 35 Genes and associated literature references:
- “APC” refers to the adenomatous polyposis coli gene (Eads et al., *Cancer Res.* 59:2302-2306, 1999; Hiltunen et al., *Int. J. Cancer.* 70:644-648, 1997);
- “ARF” refers to the P14 cell cycle regulator, tumor suppressor gene (Esteller et al., *Cancer Res.* 60:129-133, 2000; Robertson & Jones, *Mol. Cell. Biol.* 18:6457-6473, 1998);

WO 01/75172

PCT/US01/10658

- "*CALCA*" refers to the calcitonin gene (Melki et al., *Cancer Res.* 59:3730-3740, 1999; Hakkarainen et al., *Int. J. Cancer.* 69:471-474, 1996);
- "*CDHI*" refers to the E-cadherin gene (Melki et al., *Cancer Res.* 59:3730-3740, 1999; Ueki et al., *Cancer Res.* 60:1835-1839, 2000);
- 5 "5" "*CDKN2A*" refers to the P16 gene (Jones & Laird, *Nat. Genet.* 21:163-167, 1999; Melki et al., *Cancer Res.* 59:3730-3740, 1999; Baylin & Herman, *Trends Genet.* 16:168-174, 2000; Cameron et al., *Nat. Genet.* 21:103-107, 1999; Ueki et al., *Cancer Res.* 60:1835-1839, 2000);
- "*CDKN2B*" refers to the P15 gene (Melki et al., *Cancer Res.* 59:3730-3740, 1999; Cameron et al., *Nat. Genet.* 21:103-107, 1999);
- 10 "10" "*CTNNB1*" refers to the beta-catenin gene;
- "*ESR1*" refers to the estrogen receptor alpha gene (Jones & Laird, *Nat. Genet.* 21:163-167, 1999; Baylin & Herman, *Trends Genet.* 16:168-174, 2000);
- "*GSTP1*" refers to the glutathione S-transferase P1 gene (Melki et al., *Cancer Res.* 59:3730-3740, 1999; Tchou et al., *Int. J. Oncol.* 16:663-676, 2000);
- 15 "15" "*HIC1*" refers to the hypermethylated in cancer 1 gene (Melki et al., *Cancer Res.* 59:3730-3740, 1999; Wales et al., *Nat. Med.* 1:570-577, 1995);
- "*MGMT*" refers to the O6-methylguanine-DNA methyltransferase gene (Esteller et al., *Cancer Res.* 59:793-797, 1999);
- 20 "20" "*MLH1*" refers to the Mut L homologue 1 gene (Jones & Laird, *Nat. Genet.* 21:163-167, 1999; Baylin & Herman, *Trends Genet.* 16:168-174, 2000; Cameron et al., *Nat. Genet.* 21:103-107, 1999; Esteller et al., *Am. J. Pathol.* 155:1767-1772, 1999; Ueki et al., *Cancer Res.* 60:1835-1839, 2000);
- "*MTHFR*" refers to the methyl-tetrahydrofolate reductase gene (Pereira et al., *Oncol. Rep.* 6:597-599, 1999);
- 25 "25" "*MYOD1*" refers to the myogenic determinant 1 gene (Eads et al., *Cancer Res.* 59:2302-2306, 1999; Cheng et al., *Br. J. Cancer.* 75:396-402, 1997);
- "*PTGS2*" refers to the cyclooxygenase 2 gene (Zimmermann et al., *Cancer Res.* 59:198-204, 1999);
- 30 "30" "*RBI*" refers to the retinoblastoma gene (Stirzaker et al., *Cancer Res.* 57:2229-2237, 1997; Sakai et al., *Am. J. Hum. Genet.* 48:880-888, 1991);
- "*TGFBR2*" refers to the transforming growth factor beta receptor II gene (Kang et al., *Oncogene.* 18:7280-7286, 1999; Hougaard et al., *Br. J. Cancer.* 79:1005-1011, 1999);
- "*THBS1*" refers to the thrombospondin 1 gene (Ueki et al., *Cancer Res.* 60:1835-1839, 2000; Li et al., *Oncogene.* 18:3284-3289, 1999);
- 35 "35" "*TIMP3*" refers to the tissue inhibitor of metalloproteinase 3 gene (Cameron et al., *Nat. Genet.* 21:103-107, 1999; Ueki et al., *Cancer Res.* 60:1835-1839, 2000; Bachman et al., *Cancer Res.* 59:798-802, 1999);
- "*TYMS1*" refers to the thymidylate synthetase gene (Sakamoto et al., *In: L. Herrera*

WO 01/75172

PCT/US01/10658

(ed.) Familial adenomatous polyposis, pp. 315-324. New York: Alan R. Liss, 1990).

Overview

5 The present invention encompasses a broad, multi-gene approach that provides novel and therapeutically useful insight into concordant methylation behavior between and among genes. In particular embodiments, the present invention provides *novel epigenomic fingerprints* for the different histological stages of esophageal adenocarcinoma (EAC).

10 More specifically, the present invention combines the advantages of both targeted and comprehensive approaches by analyzing 20 different genes (*see* Table 1, below) using a quantitative, high-throughput methylation assay, "MethyLightTM" (Eads et al., *Cancer Res.* 59:2302-2306, 1999; Eads et al., *Cancer Res.* 60:5021-5026, 2000; Eads et al., *Nucleic Acids Res.* 28:E32, 2000), to (i) more extensively characterize the methylation changes in esophageal adenocarcinoma (EAC); to (ii) generate epigenomic fingerprints for the different histological stages of EAC; to (iii) identify epigenetic biomarkers useful in
15 disease diagnosis and prevention; and to (iv) determine if CIMP is a contributor to the tumorigenesis of esophageal adenocarcinoma tumors.

A total of 104 tissue specimens from 51 patients with different stages of Barrett's esophagus and/or associated adenocarcinoma were analyzed. Specifically, 84 of these
20 tissue specimens were screened with the full panel of 20 genes, revealing distinct classes of methylation patterns in the different types of tissue.

The most informative genes, for purposes of the present invention, were those with an intermediate frequency of significant hypermethylation (*i.e.*, those ranging from about 15% (*CDKN2A*) to about 60% (*MGMT*) of the samples). This group of genes could be
25 further subdivided into three classes, according to the (1) absence (*CDKN2A*, *ESR1* and *MYOD1*), or (2) presence (*CALCA*, *MGMT* and *TIMP3*) of methylation in normal esophageal mucosa and stomach, or (3) the infrequent methylation of normal esophageal mucosa accompanied by methylation in all normal stomach samples (*APC*).

The other genes were relatively less informative, since the frequency of
30 hypermethylation was below about 5% (*ARF*, *CDH1*, *CDKN2B*, *GSTP1*, *MLH1*, *PTGS2* and *THBS1*), completely absent (*CTNNB1*, *RBI*, *TGFBR2* and *TYMS1*) or ubiquitous (*HIC1* and *MTHFR*), regardless of tissue type.

Each class of gene undergoes unique epigenetic changes at different steps of disease progression of EAC, consistent with a step-wise loss of multiple protective barriers
35 against CpG island hypermethylation. The aberrant hypermethylation occurs at many different loci in the same tissues, consistent with an overall deregulation of methylation control in EAC tumorigenesis. However, there was no clear evidence for a distinct group of tumors with a CpG island methylator phenotype ("CIMP").

Additionally, normal and metaplastic tissues from patients with evidence of

WO 01/75172

PCT/US01/10658

associated dysplasia or cancer displayed a significantly higher incidence of hypermethylation than similar tissues from patients with no further progression of their disease. The fact that the samples from these two groups of patients were histologically indistinguishable, yet molecularly distinct, indicates, according to the present invention, 5 that the occurrence of such hypermethylation provides a novel and valuable clinical tool to identify patients with pre-malignant Barrett's, who are at risk for further progression.

TABLE I shows a list of gene names and functions analyzed by the MethyLight™ assay in EAC. The genes are listed in alphabetical order based on their designated HUGO (HUMAN Genome Organization) names. The genes are divided into three groups 10 according to whether or not they have CpG islands and are known to be methylated in other tumors. A brief description of the function of each gene is included.

WO 01/75172

PCT/US01/10658

TABLE I. List of Gene Names Analyzed by the MethylLight™ assay in EAC

GENE SYMBOL	NAME	TUMORS WITH HYPERMETHYLATION †	PROPOSED FUNCTION
<i>APC</i>	Adenomatous Polyposis Coli	Colon	WNT Signaling Pathway; Beta-Catenin Degradation; Tumor Suppressor
<i>ARF</i>	P14	Colon, Lymphoma	Cell Cycle Regulator; Tumor Suppressor
<i>CALCA</i>	Calcitonin	Breast, Colon, Leukemia	Regulates Calcium Levels via Adenylate Cyclase
<i>CDH1</i>	E-Cadherin	AML, Bladder, Breast, Colon, Gastric, Thyroid	Cell Adhesion
<i>CDKN2A</i>	P16	AML, Bladder, Colon, Gastric, Lymphoma, Melanoma	Cell Cycle Regulator, Tumor Suppressor
<i>CDKN2B</i>	P15	Colon, Hematological Malignancies	Cell Cycle Regulator, Tumor Suppressor
<i>ESR1</i>	Estrogen Receptor Alpha	AML, Colon, Breast, Lung, Leukemia	Hormone Receptor in Mammary Cells; Growth Suppressive in Colorectal Cancer Cells
<i>GSTP1</i>	Glutathione S-Transferase P1	Breast, Prostate, Hepatocellular	Protects Against Oxidant and Electrophilic Carcinogens
<i>HIC1</i>	Hypermethylated in Cancer 1	Brain, Breast, Colon, Renal, Leukemia, Lymphoma	Zinc Finger Transcription Factor; Potential Tumor Suppressor
<i>MGMT</i>	O6-Methylguanine-DNA Methyltransferase	Brain, Colon, Lymphoma, Non-Small Cell Lung Cancer	DNA Repair
<i>MLH1</i>	Mut L Homologue 1	Colon, Endometrial, Gastric	DNA Mismatch Repair
<i>MYO1</i>	Myogenic Determinant 1	Breast, Colon, Ovarian	Muscle Differentiation
<i>RBI</i>	Retinoblastoma 1	Retinoblastoma	Cell Cycle Regulator; Tumor Suppressor

WO 01/75172

PCT/US01/10658

GENE SYMBOL	NAME	TUMORS WITH HYPERMETHYLATION [†]	PROPOSED FUNCTION
<i>TGFBR2</i>	Transforming Growth Factor Beta Receptor II	Colon, Gastric, Small-Cell Lung Cancer	Serine-Threonine Kinase Receptor; Cell Signaling
<i>THBS1</i>	Thrombospondin 1	Colon, Glioblastomas	Angiogenesis Inhibitor
<i>TIMP3</i>	Tissue inhibitor of Metalloproteinase 3	Brain, Breast, Colon, Kidney, Lung, Pancreatic	Metastasis, Angiogenesis
CPG ISLAND: UNKNOWN METHYLATION STATUS			
<i>CTNNB1</i>	Beta-Catenin	Unknown	WNT Signaling Pathway (Regulated by APC); Cell Adhesion via E-Cadherin
<i>PTGS2</i>	Cyclooxygenase 2	Unknown	Prostaglandin Synthesis from Arachidonic Acid
<i>TMS</i>	Thymidylate Synthetase	Unknown	Nucleotide/DNA Synthesis
NON-CPG ISLAND: UNKNOWN METHYLATION STATUS			
<i>MTHFR</i>	Methyl-Tetrahydrofolate Reductase	Unknown	Folate Metabolism, Methionine Synthesis; Predisposing Polymorphism in Colon Cancer

TABLE 1 cont.

[†] See literature references relating to specific genes under "DEFINITIONS," herein above.

WO 01/75172

PCT/US01/10658

Diagnostic and Prognostic Assays for Cancer

The present invention provides for diagnostic and prognostic cancer assays based on determination of the methylation state of one or more of the disclosed 20 gene sequences (5 *APC, ARF, CALCA, CDH1, CDKN2A, CDKN2B, ESR1, GSTP1, HIC1, MGMT, MLH1, MYOD1, RB1, TGFBR2, THBS1, TIMP3, CTNNB1, PTGS2, TYMS* and *MTHFR*; see TABLES I and II, below; and see under "Definitions," above), or methylation-altered DNA sequence embodiments thereof. These 20 gene sequence regions are defined herein by the oligomeric primers and probes corresponding to SEQ ID NOS:1-60, 64 and 65 (see TABLE 10 II, below). SEQ ID NOS:61-63 correspond to the *ACTB* "control" gene region used in the present analysis (see EXAMPLE 1, below).

Additionally, 19 of these 20 gene sequence regions correspond to CpG islands or regions thereof (based on GC Content and O/E ratio); namely *APC, ARF, CALCA, CDH1, CDKN2A, CDKN2B, ESR1, GSTP1, HIC1, MGMT, MLH1, MYOD1, RB1, TGFBR2, THBS1, TIMP3, CTNNB1, PTGS2* and *TYMS* (see TABLE 1, below). Thus, based on the fact that the methylation state of a portion of a given CpG island is generally representative of the island as a whole, the present invention further encompasses the novel use of any sequences within the 19 complete CpG islands associated with these 19 gene sequence regions (defined herein by the primers and probes corresponding to SEQ ID NOS:1-60, 64 and 65 (see TABLE II, 15 below) in cancer prognostic and diagnostic applications), where a CpG island sequence associated with one of these 19 gene sequences is that contiguous sequence of genomic DNA that encompasses at least one nucleotide of one of these 19 gene sequences, and satisfies the criteria of having both a frequency of CpG dinucleotides corresponding to an Observed/Expected Ratio >0.6, and a GC Content >0.5.

25 Typically, such assays involve obtaining a tissue sample from a test tissue, performing a methylation assay on DNA derived from the tissue sample to determine the associated methylation state, and making a diagnosis or prognosis based thereon.

The methylation assay is used to determine the methylation state of one or a plurality of CpG dinucleotide within a DNA sequence of the DNA sample. According to the present invention, possible methylation states include *hypermethylation* and *hypomethylation*, relative to a normal state (*i.e.*, non-cancerous control state). Hypermethylation and hypomethylation refer to the methylation states corresponding to an *increased* or *decreased*, respectively, presence of 5-methylcytosine ("5-mCyt") at one or a plurality of CpG dinucleotides within a DNA sequence of the test sample, relative to the amount of 5-mCyt found at corresponding CpG dinucleotides within a normal control DNA sample.

35 A diagnosis or prognosis is based, at least in part, upon the determined methylation

WO 01/75172

PCT/US01/10658

state of the sample DNA sequence compared to control data obtained from normal, non-cancerous tissue.

Methylation Assay Procedures

5

Various methylation assay procedures are known in the art, and can be used in conjunction with the present invention. These assays allow for determination of the methylation state of one or a plurality of CpG dinucleotides within a DNA sequence (e.g., CpG islands). Such assays involve, among other techniques, DNA sequencing of bisulfite-treated DNA, PCR (for sequence-specific amplification), Southern blot analysis, use of methylation-sensitive restriction enzymes, etc.

10

For example, genomic sequencing has been simplified for analysis of DNA methylation patterns and 5-methylcytosine distribution by using bisulfite treatment (Frommer et al., *Proc. Natl. Acad. Sci. USA* 89:1827-1831, 1992). Additionally, restriction enzyme digestion of PCR products amplified from bisulfite-converted DNA is used, e.g., the method described by Sadri & Hornsby (*Nucl. Acids Res.* 24:5058-5059, 1996), or COBRA (Combined Bisulfite Restriction Analysis) (Xiong & Laird, *Nucleic Acids Res.* 25:2532-2534, 1997).

15

Preferably, assays such as "MethylLightTM" (a fluorescence-based real-time PCR technique) (Eads et al., *Cancer Res.* 59:2302-2306, 1999), Methylation-sensitive Single Nucleotide Primer Extension reactions ("Ms-SnuPE"; Gonzalzo & Jones, *Nucleic Acids Res.* 25:2529-2531, 1997), methylation-specific PCR ("MSP"; Herman et al., *Proc. Natl. Acad. Sci. USA* 93:9821-9826, 1996; US Patent No. 5,786,146), and methylated CpG island amplification ("MCA"; Toyota et al., *Cancer Res.* 59:2307-12, 1999) are used alone or in combination with other of these methods. Methylation assays that can be used in various

20

25

embodiments of the present invention include, but are not limited to, the following assays. *COBRA (Combined Bisulfite Restriction Analysis)*. COBRA analysis is a quantitative methylation assay useful for determining DNA methylation levels at specific gene loci in small amounts of genomic DNA (Xiong & Laird, *Nucleic Acids Res.* 25:2532-2534, 1997). Briefly, restriction enzyme digestion is used to reveal methylation-dependent sequence differences in PCR products of sodium bisulfite-treated DNA. Methylation-dependent sequence differences are first introduced into the genomic DNA by standard bisulfite treatment according to the procedure described by Frommer et al. (*Proc. Natl. Acad. Sci. USA* 89:1827-1831, 1992). PCR amplification of the bisulfite converted DNA is then performed using primers specific for the interested CpG islands, followed by restriction endonuclease digestion, gel electrophoresis, and detection using specific, labeled hybridization probes. Methylation levels in the original DNA sample are represented by the relative amounts of

30

35

WO 01/75172

PCT/US01/10658

digested and undigested PCR product in a linearly quantitative fashion across a wide spectrum of DNA methylation levels. Additionally, this technique can be reliably applied to DNA obtained from microdissected paraffin-embedded tissue samples. Typical reagents (*e.g.*, as might be found in a typical COBRA-based methylation kit) for COBRA analysis may include, but are not limited to: PCR primers for specific gene (or methylation-altered DNA sequence or CpG island); restriction enzyme and appropriate buffer; gene-hybridization oligo; control hybridization oligo; kinase labeling kit for oligo probe; and radioactive nucleotides (although other label schemes known in the art including, but not limited, to fluorescent and phosphorescent schemes can be used). Additionally, bisulfite conversion reagents may include: DNA denaturation buffer; sulfonation buffer; DNA recovery reagents or kit (*e.g.*, precipitation, ultrafiltration, affinity column); desulfonation buffer; and DNA recovery components.

Ms-SNuPE (Methylation-sensitive Single Nucleotide Primer Extension). The Ms-SNuPE technique is a quantitative method for assessing methylation differences at specific CpG sites based on bisulfite treatment of DNA, followed by single-nucleotide primer extension (Gonzalzo & Jones, *Nucleic Acids Res.* 25:2529-2531, 1997). Briefly, genomic DNA is reacted with sodium bisulfite to convert unmethylated cytosine to uracil while leaving 5-methylcytosine unchanged. Amplification of the desired target sequence is then performed using PCR primers specific for bisulfite-converted DNA, and the resulting product is isolated and used as a template for methylation analysis at the CpG site(s) of interest. Small amounts of DNA can be analyzed (*e.g.*, microdissected pathology sections), and it avoids utilization of restriction enzymes for determining the methylation status at CpG sites. Typical reagents (*e.g.*, as might be found in a typical Ms-SNuPE-based methylation kit) for Ms-SNuPE analysis may include, but are not limited to: PCR primers for specific gene (or methylation-altered DNA sequence or CpG island); optimized PCR buffers and deoxynucleotides; gel extraction kit; positive control primers; Ms-SNuPE primers for specific gene; reaction buffer (for the Ms-SNuPE reaction); and radioactive nucleotides. Additionally, bisulfite conversion reagents may include: DNA denaturation buffer; sulfonation buffer; DNA recovery reagents or kit (*e.g.*, precipitation, ultrafiltration, affinity column); desulfonation buffer; and DNA recovery components.

MSP (Methylation-specific PCR). MSP allows for assessing the methylation status of virtually any group of CpG sites within a CpG island, independent of the use of methylation-sensitive restriction enzymes (Herman et al. *Proc. Natl. Acad. Sci. USA* 93:9821-9826, 1996; US Patent No. 5,786,146). Briefly, DNA is modified by sodium bisulfite converting all unmethylated, but not methylated cytosines to uracil, and subsequently amplified with primers specific for methylated versus unmethylated DNA. MSP requires only small quantities of

WO 01/75172

PCT/US01/10658

DNA, is sensitive to 0.1% methylated alleles of a given CpG island locus, and can be performed on DNA extracted from paraffin-embedded samples. Typical reagents (e.g., as might be found in a typical MSP-based kit) for MSP analysis may include, but are not limited to: methylated and unmethylated PCR primers for specific gene (or methylation-altered DNA sequence or CpG island), optimized PCR buffers and deoxynucleotides, and specific probes.

5 *MCA (Methylated CpG Island Amplification)*. The MCA technique is a method that can be used to screen for altered methylation patterns in genomic DNA, and to isolate specific sequences associated with these changes (Toyota et al., *Cancer Res.* 59:2307-12, 1999). Briefly, restriction enzymes with different sensitivities to cytosine methylation in their
10 recognition sites are used to digest genomic DNAs from primary tumors, cell lines, and normal tissues prior to arbitrarily primed PCR amplification. Fragments that show differential methylation are cloned and sequenced after resolving the PCR products on high-resolution polyacrylamide gels. The cloned fragments are then used as probes for Southern analysis to confirm differential methylation of these regions. Typical reagents (e.g., as might be found in
15 a typical MCA-based kit) for MCA analysis may include, but are not limited to: PCR primers for arbitrary priming Genomic DNA; PCR buffers and nucleotides, restriction enzymes and appropriate buffers; gene-hybridization oligos or probes; control hybridization oligos or probes.

DMH (Differential Methylation Hybridization). DMH refers to the art-recognized, array-based methylation assay described in Huang et al., *Hum. Mol. Genet.*, 8:459-470, 1999, and in Yan et al., *Clin. Cancer Res.* 6:1432-38, 2000. DMH allows for a genome-wide screening of CpG island hypermethylation in cancer cell lines, and. Briefly, CpG island tags are arrayed on solid supports (e.g., nylon membranes, silicon, etc.), and probed with
20 "amplicons" representing a pool of methylated CpG DNA, from test (e.g., tumor) or reference samples. The differences in test and reference signal intensities on screened CpG island arrays reflect methylation alterations of corresponding sequences in the test DNA.

MethylLight™. In preferred embodiments, the MethylLight™ assay is used to determine the methylation status of one or more CpG sequences. The MethylLight™ assay is a high-throughput quantitative methylation assay that utilizes fluorescence-based real-time
30 PCR (TaqMan®) technology that requires no further manipulations after the PCR step (Eads et al., *Cancer Res.* 60:5021-5026, 2000; Eads et al., *Cancer Res.* 59:2302-2306, 1999; Eads et al., *Nucleic Acids Res.* 28:E32, 2000). Briefly, the MethylLight™ process begins with a mixed sample of genomic DNA that is converted, in a sodium bisulfite reaction, to a mixed pool of methylation-dependent sequence differences according to standard procedures (the
35 bisulfite process converts unmethylated cytosine residues to uracil). Fluorescence-based PCR is then performed either in an "unbiased" (with primers that do not overlap known CpG

WO 01/75172

PCT/US01/10658

methylation sites) PCR reaction, or in a "biased" (with PCR primers that overlap known CpG dinucleotides) reaction. Sequence discrimination can occur either at the level of the amplification process or at the level of the fluorescence detection process, or both.

5 The MethyLight™ assay may be used as a quantitative test for methylation patterns in the genomic DNA sample, wherein sequence discrimination occurs at the level of probe hybridization. In this quantitative version, the PCR reaction provides for unbiased amplification in the presence of a fluorescent probe that overlaps a particular putative methylation site. An unbiased control for the amount of input DNA is provided by a reaction in which neither the primers, nor the probe overlap any CpG dinucleotides. Alternatively, a
10 qualitative test for genomic methylation is achieved by probing of the biased PCR pool with either control oligonucleotides that do not "cover" known methylation sites (a fluorescence-based version of the "MSP" technique), or with oligonucleotides covering potential methylation sites.

The MethyLight™ process can be used with a "TaqMan®" probe in the amplification
15 process. For example, double-stranded genomic DNA is treated with sodium bisulfite and subjected to one of two sets of PCR reactions using TaqMan® probes; e.g., with either biased primers and TaqMan® probe, or unbiased primers and TaqMan® probe. The TaqMan® probe is dual-labeled with fluorescent "reporter" and "quencher" molecules, and is designed to be specific for a relatively high GC content region so that it melts out at about 10°C higher
20 temperature in the PCR cycle than the forward or reverse primers. This allows the TaqMan® probe to remain fully hybridized during the PCR annealing/extension step. As the Taq polymerase enzymatically synthesizes a new strand during PCR, it will eventually reach the annealed TaqMan® probe. The Taq polymerase 5' to 3' endonuclease activity will then displace the TaqMan® probe by digesting it to release the fluorescent reporter molecule for
25 quantitative detection of its now unquenched signal using a real-time fluorescent detection system.

Typical reagents (e.g., as might be found in a typical MethyLight™-based methylation kit) for MethyLight™ analysis may include, but are not limited to: PCR primers for specific gene (or methylation-altered DNA sequence or CpG island); TaqMan® probes;
30 optimized PCR buffers and deoxynucleotides; and Taq polymerase. A detailed description of four alternate process applications ("A" through "D") of the MethyLight™ assay follows below. Preferably, the quantitative MethyLight™ process application "B" is used.

MethyLight™-based detection of the methylated nucleic acid is relatively rapid and is based on amplification-mediated displacement of specific oligonucleotide probes. In a
35 preferred embodiment, amplification and detection, in fact, occur simultaneously as measured by fluorescence-based real-time quantitative PCR ("RT-PCR") using specific, dual-labeled

WO 01/75172

PCT/US01/10658

TaqMan® oligonucleotide probes, with no requirement for subsequent manipulation or analysis. The displaceable probes can be specifically designed to distinguish between methylated and unmethylated CpG sites present in the original, unmodified nucleic acid sample.

5 Like the technique of methylation-specific PCR ("MSP"; US Patent 5,786,146), MethyLight™ provides for significant advantages over previous PCR-based and other methods (e.g., Southern analyses) used for determining methylation patterns. MethyLight™ is substantially more sensitive than Southern analysis, and facilitates the detection of a low number (percentage) of methylated alleles in very small nucleic acid samples, as well as
10 paraffin-embedded samples. Moreover, in the case of genomic DNA, analysis is not limited to DNA sequences recognized by methylation-sensitive restriction endonucleases, thus allowing for fine mapping of methylation patterns across broader CpG-rich regions. MethyLight™ also eliminates any false-positive results, that otherwise might result from incomplete digestion by methylation-sensitive restriction enzymes, inherent in previous PCR-
15 based methylation methods.

MethyLight™ can be applied as a quantitative process for measuring methylation amounts, and is substantially more rapid than other methods. MethyLight™ does not require any post-PCR manipulation or processing. This not only greatly reduces the amount of labor involved in the analysis of bisulfite-treated DNA, but it also provides a means to avoid
20 handling of PCR products that could contaminate future reactions.

One process embodiment uses MethyLight™ for the unbiased amplification of all possible methylation states using primers that do not cover any CpG sequences in the original, unmodified DNA sequence. To the extent that all methylation patterns are amplified equally, quantitative information about DNA methylation patterns are then distilled from the resulting
25 PCR pool by any technique capable of detecting sequence differences (e.g., by fluorescence-based PCR).

MethyLight™ employs one or a series of CpG-specific TaqMan® probes, each corresponding to a particular methylation site in a given amplified DNA region, are constructed. This series of probes is then utilized in parallel amplification reactions, using
30 aliquots of a single, modified DNA sample, to simultaneously determine the complete methylation pattern present in the original unmodified sample of genomic DNA. This is accomplished in a fraction of the time and expense required for direct sequencing of the sample of genomic DNA, and are substantially more sensitive. Moreover, one embodiment of MethyLight™ provides for a quantitative assessment of such a methylation pattern.

35 The present invention, as described herein, may be practiced using a variety of methylation assays. For MethyLight™ embodiments, there are four process techniques and

WO 01/75172

PCT/US01/10658

associated diagnostic kits that a methylation-dependent nucleic acid modifying agent (*e.g.*, bisulfite), to both qualitatively and quantitatively determine CpG methylation status in nucleic acid samples (*e.g.*, genomic DNA samples). The four processes are described herein as processes "A," "B," "C" and "D." Overall, methylated-CpG sequence discrimination is designed to occur at the level of amplification, probe hybridization or at both levels. For example, applications C and D utilize "biased" primers that distinguish between modified unmethylated and methylated nucleic acid and provide methylated-CpG sequence discrimination at the PCR amplification level. Process B uses "unbiased" primers (that do not cover CpG methylation sites), to provide for unbiased amplification of modified nucleic acid, but rather utilize probes that distinguish between modified unmethylated and methylated nucleic acid to provide for quantitative methylated-CpG sequence discrimination at the detection level (*e.g.*, at the fluorescent (or luminescent) probe hybridization level only). Process A does not, in itself, provide for methylated-CpG sequence discrimination at either the amplification or detection levels, but supports and validates the other three applications by providing control reactions for input DNA.

MethylLight™ Process D. In a first MethylLight™ embodiment, the invention provides a method for qualitatively detecting a methylated CpG-containing nucleic acid, the method including: contacting a nucleic acid-containing sample with a modifying agent that modifies unmethylated cytosine to produce a converted nucleic acid; amplifying the converted nucleic acid by means of two oligonucleotide primers in the presence of a specific oligonucleotide hybridization probe, wherein both the primers and probe distinguish between modified unmethylated and methylated nucleic acid; and detecting the "methylated" nucleic acid based on amplification-mediated probe displacement.

The term "modifies" as used herein means the conversion of an unmethylated cytosine to another nucleotide by the modifying agent, said conversion distinguishing unmethylated from methylated cytosine in the original nucleic acid sample. Preferably, the agent modifies unmethylated cytosine to uracil. Preferably, the agent used for modifying unmethylated cytosine is sodium bisulfite, however, other equivalent modifying agents that selectively modify unmethylated cytosine, but not methylated cytosine, can be substituted in the method of the invention. Sodium-bisulfite readily reacts with the 5, 6-double bond of cytosine, but not with methylated cytosine, to produce a sulfonated cytosine intermediate that undergoes deamination under alkaline conditions to produce uracil. Because Taq polymerase recognizes uracil as thymine and 5-methylcytidine (^{m5}C) as cytidine, the sequential combination of sodium bisulfite treatment and PCR amplification results in the ultimate conversion of unmethylated cytosine residues to thymine (C → U → T) and methylated cytosine residues (^{m5}C) to cytosine (^{m5}C → T → C). Thus, sodium-bisulfite treatment of genomic DNA

WO 01/75172

PCT/US01/10658

creates methylation-dependent sequence differences by converting unmethylated cytosines to uracil, and upon PCR the resultant product contains cytosine only at positions where methylated cytosine occurs in the unmodified nucleic acid.

Oligonucleotide "primers," as used herein, means linear, single-stranded, oligomeric deoxyribonucleic or ribonucleic acid molecules capable of sequence-specific hybridization (annealing) with complementary strands of modified or unmodified nucleic acid. As used herein, the specific primers are preferably DNA. The primers of the invention embrace oligonucleotides of appropriate sequence and sufficient length so as to provide for specific and efficient initiation of polymerization (primer extension) during the amplification process. As used in the inventive processes, oligonucleotide primers typically contain 12-30 nucleotides or more, although may contain fewer nucleotides. Preferably, the primers contain from 18-30 nucleotides. The exact length will depend on multiple factors including temperature (during amplification), buffer, and nucleotide composition. Preferably, primers are single-stranded although double-stranded primers may be used if the strands are first separated. Primers may be prepared using any suitable method, such as conventional phosphotriester and phosphodiester methods or automated embodiments which are commonly known in the art.

As used in the inventive embodiments herein, the specific primers are preferably designed to be substantially complementary to each strand of the genomic locus of interest. Typically, one primer is complementary to the negative (-) strand of the locus (the "lower" strand of a horizontally situated double-stranded DNA molecule) and the other is complementary to the positive (+) strand ("upper" strand). As used in the embodiment of Application D, the primers are preferably designed to overlap potential sites of DNA methylation (CpG nucleotides) and specifically distinguish modified unmethylated from methylated DNA. Preferably, this sequence discrimination is based upon the differential annealing temperatures of perfectly matched, versus mismatched oligonucleotides. In the embodiment of Application D, primers are typically designed to overlap from one to several CpG sequences. Preferably, they are designed to overlap from 1 to 5 CpG sequences, and most preferably from 1 to 4 CpG sequences. By contrast, in a quantitative embodiment of the invention employed in the Examples of the present invention, the primers do not overlap any CpG sequences.

In the case of fully "unmethylated" (complementary to modified unmethylated nucleic acid strands) primer sets, the anti-sense primers contain adenosine residues ("A"s) in place of guanosine residues ("G"s) in the corresponding (-) strand sequence. These substituted As in the anti-sense primer will be complementary to the uracil and thymidine residues ("Us" and "Ts") in the corresponding (+) strand region resulting from bisulfite modification of

WO 01/75172

PCT/US01/10658

unmethylated C residues ("Cs") and subsequent amplification. The sense primers, in this case, are preferably designed to be complementary to anti-sense primer extension products, and contain Ts in place of unmethylated Cs in the corresponding (+) strand sequence. These substituted Ts in the sense primer will be complementary to the As, incorporated in the anti-sense primer extension products at positions complementary to modified Cs (Us) in the original (+) strand.

In the case of fully-methylated primers (complementary to methylated CpG-containing nucleic acid strands), the anti-sense primers will not contain As in place of Gs in the corresponding (-) strand sequence that are complementary to methylated Cs (*i.e.*, ^mCpG sequences) in the original (+) strand. Similarly, the sense primers in this case will not contain Ts in place of methylated Cs in the corresponding (+) strand ^mCpG sequences. However, Cs that are not in CpG sequences in regions covered by the fully-methylated primers, and are not methylated, will be represented in the fully-methylated primer set as described above for unmethylated primers.

Preferably, as employed in the embodiment of process D, the amplification process provides for amplifying bisulfite converted nucleic acid by means of two oligonucleotide primers in the presence of a specific oligonucleotide hybridization probe. Both the primers and probe distinguish between modified unmethylated and methylated nucleic acid. Moreover, detecting the "methylated" nucleic acid is based upon amplification-mediated probe fluorescence. In one embodiment, the fluorescence is generated by probe degradation by 5' to 3' exonuclease activity of the polymerase enzyme. In another embodiment, the fluorescence is generated by fluorescence energy transfer effects between two adjacent hybridizing probes (Lightcycler® technology) or between a hybridizing probe and a primer. In another embodiment, the fluorescence is generated by the primer itself (Sunrise® technology). Preferably, the amplification process is an enzymatic chain reaction that uses the oligonucleotide primers to produce exponential quantities of amplification product, from a target locus, relative to the number of reaction steps involved.

As describe above, one member of a primer set is complementary to the (-) strand, while the other is complementary to the (+) strand. The primers are chosen to bracket the area of interest to be amplified; that is, the "amplicon." Hybridization of the primers to denatured target nucleic acid followed by primer extension with a DNA polymerase and nucleotides, results in synthesis of new nucleic acid strands corresponding to the amplicon. Preferably, the DNA polymerase is *Taq* polymerase, as commonly used in the art. Although equivalent polymerases with a 5' to 3' nuclease activity can be substituted. Because the new amplicon sequences are also templates for the primers and polymerase, repeated cycles of denaturing, primer annealing, and extension results in exponential production of the amplicon. The

WO 01/75172

PCT/US01/10658

product of the chain reaction is a discrete nucleic acid duplex, corresponding to the amplicon sequence, with termini defined by the ends of the specific primers employed. Preferably the amplification method used is that of PCR (Mullis et al., *Cold Spring Harb. Symp. Quant. Biol.* 51:263-273; Gibbs, *Anal. Chem.* 62:1202-1214, 1990), or more preferably, automated

5 embodiments thereof which are commonly known in the art.

Preferably, methylation-dependent sequence differences are detected by methods based on fluorescence-based quantitative PCR (real-time quantitative PCR, Heid et al., *Genome Res.* 6:986-994, 1996; Gibson et al., *Genome Res.* 6:995-1001, 1996) (e.g., "TaqMan®," "Lightcycler®," and "Sunrise®" technologies). For the TaqMan® and

10 Lightcycler® technologies, the sequence discrimination can occur at either or both of two steps: (1) the amplification step, or (2) the fluorescence detection step. In the case of the "Sunrise®" technology, the amplification and fluorescent steps are the same. In the case of the FRET hybridization, probes format on the Lightcycler®, either or both of the FRET oligonucleotides can be used to distinguish the sequence difference. Most preferably the

15 amplification process, as employed in all inventive embodiments herein, is that of fluorescence-based Real Time Quantitative PCR (Heid et al., *Genome Res.* 6:986-994, 1996) employing a dual-labeled fluorescent oligonucleotide probe (TaqMan® PCR, using an ABI Prism 7700 Sequence Detection System, Perkin Elmer Applied Biosystems, Foster City, California).

The "TaqMan®" PCR reaction uses a pair of amplification primers along with a nonextendible interrogating oligonucleotide, called a TaqMan® probe, that is designed to hybridize to a GC-rich sequence located between the forward and reverse (*i.e.*, sense and anti-sense) primers. The TaqMan® probe further comprises a fluorescent "reporter moiety" and a "quencher moiety" covalently bound to linker moieties (e.g., phosphoramidites) attached to

20 nucleotides of the TaqMan® oligonucleotide. Examples of suitable reporter and quencher molecules are: the 5' fluorescent reporter dyes 6FAM ("FAM"; 2,7 dimethoxy-4,5-dichloro-6-carboxy-fluorescein), and TET (6-carboxy-4,7,2',7'-tetrachlorofluorescein); and the 3' quencher dye TAMRA (6-carboxytetramethylrhodamine) (Livak et al., *PCR Methods Appl.* 4:357-362, 1995; Gibson et al., *Genome Res.* 6:995-1001; and 1996; Heid et al., *Genome Res.* 6:986-994, 1996).

25

One process for designing appropriate TaqMan® probes involves utilizing a software facilitating tool, such as "Primer Express" that can determine the variables of CpG island location within GC-rich sequences to provide for at least a 10°C melting temperature difference (relative to the primer melting temperatures) due to either specific sequence

30 (tighter bonding of GC, relative to AT base pairs), or to primer length.

The TaqMan® probe may or may not cover known CpG methylation sites, depending

WO 01/75172

PCT/US01/10658

on the particular inventive process used. Preferably, in the embodiment of process D, the TaqMan® probe is designed to distinguish between modified unmethylated and methylated nucleic acid by overlapping from 1 to 5 CpG sequences. As described above for the fully unmethylated and fully methylated primer sets, TaqMan® probes may be designed to be
5 complementary to either unmodified nucleic acid, or, by appropriate base substitutions, to bisulfite-modified sequences that were either fully unmethylated or fully methylated in the original, unmodified nucleic acid sample.

Each oligonucleotide primer or probe in the TaqMan® PCR reaction can span anywhere from zero to many different CpG dinucleotides that each can result in two different
10 sequence variations following bisulfite treatment (^mCpG, or UpG). For instance, if an oligonucleotide spans 3 CpG dinucleotides, then the number of possible sequence variants arising in the genomic DNA is $2^3 = 8$ different sequences. If the forward and reverse primer each span 3 CpGs and the probe oligonucleotide (or both oligonucleotides together in the case of the FRET format) spans another 3, then the total number of sequence permutations
15 becomes $8 \times 8 \times 8 = 512$. In theory, one could design separate PCR reactions to quantitatively analyze the relative amounts of each of these 512 sequence variants. In practice, a substantial amount of qualitative methylation information can be derived from the analysis of a much smaller number of sequence variants. Thus, in its most simple form, the
20 fully unmethylated variants that represent the most extreme sequence variants in a hypothetical example. The ratio between these two reactions, or alternatively the ratio between the methylated reaction and a control reaction (process A), would provide a measure for the level of DNA methylation at this locus.

Detection of methylation in the MethyLight™ embodiment of process D, as in other
25 MethyLight™ embodiments herein, is based on amplification-mediated displacement of the probe. In theory, the process of probe displacement might be designed to leave the probe intact, or to result in probe digestion. Preferably, as used herein, displacement of the probe occurs by digestion of the probe during amplification. During the extension phase of the PCR cycle, the fluorescent hybridization probe is cleaved by the 5' to 3' nucleolytic activity of the
30 DNA polymerase. On cleavage of the probe, the reporter moiety emission is no longer transferred efficiently to the quenching moiety, resulting in an increase of the reporter moiety fluorescent-emission spectrum at 518 nm. The fluorescent intensity of the quenching moiety (e.g., TAMRA), changes very little over the course of the PCR amplification. Several factors may influence the efficiency of TaqMan® PCR reactions including: magnesium and salt
35 concentrations; reaction conditions (time and temperature); primer sequences; and PCR target size (i.e., amplicon size) and composition. Optimization of these factors to produce the

WO 01/75172

PCT/US01/10658

optimum fluorescence intensity for a given genomic locus is obvious to one skilled in the art of PCR, and preferred conditions are further illustrated in the "Examples" herein. The amplicon may range in size from 50 to 8,000 base pairs, or larger, but may be smaller. Typically, the amplicon is from 100 to 1000 base pairs, and preferably is from 100 to 500
5 base pairs. Preferably, the reactions are monitored in real time by performing PCR amplification using 96-well optical trays and caps, and using a sequence detector (ABI Prism) to allow measurement of the fluorescent spectra of all 96 wells of the thermal cycler continuously during the PCR amplification. Preferably, process D is run in combination with the process A to provide controls for the amount of input nucleic acid, and to normalize data
10 from tray to tray.

MethyLight™ Process C. The MethyLight™ process can be modified to avoid sequence discrimination at the PCR product detection level. Thus, in an additional qualitative process embodiment, just the primers are designed to cover CpG dinucleotides, and sequence discrimination occurs solely at the level of amplification. Preferably, the probe used in this
15 embodiment is still a TaqMan® probe, but is designed so as not to overlap any CpG sequences present in the original, unmodified nucleic acid. The embodiment of process C represents a high-throughput, fluorescence-based real-time version of MSP technology, wherein a substantial improvement has been attained by reducing the time required for detection of methylated CpG sequences. Preferably, the reactions are monitored in real time
20 by performing PCR amplification using 96-well optical trays and caps, and using a sequence detector (ABI Prism) to allow measurement of the fluorescent spectra of all 96 wells of the thermal cycler continuously during the PCR amplification. Preferably, process C is run in combination with process A (below) to provide controls for the amount of input nucleic acid, and to normalize data from tray to tray.

MethyLight™ Process B. In preferred embodiments of the present invention, the
25 MethyLight™ process can be also be modified to avoid sequence discrimination at the PCR amplification level. In a quantitative process B embodiment, just the probe is designed to cover CpG dinucleotides, and sequence discrimination occurs solely at the level of probe hybridization. Preferably, TaqMan® probes are used. In this version, sequence variants
30 resulting from the bisulfite conversion step are amplified with equal efficiency; as long as there is no inherent amplification bias (Warnecke et al., *Nucleic Acids Res.* 25:4422-4426, 1997). Design of separate probes for each of the different sequence variants associated with a particular methylation pattern (e.g., $2^3=8$ probes in the case of 3 CpGs) would allow a
35 quantitative determination of the relative prevalence of each sequence permutation in the mixed pool of PCR products. Preferably, the reactions are monitored in real time by performing PCR amplification using 96-well optical trays and caps, and using a sequence

WO 01/75172

PCT/US01/10658

detector (ABI Prism) to allow measurement of the fluorescent spectra of all 96 wells of the thermal cycler continuously during the PCR amplification. Preferably, process B is run in combination with process A, below to provide controls for the amount of input nucleic acid, and to normalize data from tray to tray.

5 *MethylLight™ Process A.* MethylLight™ process A does not, in itself, provide for methylated-CpG sequence discrimination at either the amplification or detection levels, but supports and validates the other three process applications by providing control reactions for the amount of input DNA, and to normalize data from tray to tray. Thus, if neither the primers, nor the probe overlie any CpG dinucleotides, then the reaction represents unbiased
10 amplification and measurement of amplification using fluorescent-based quantitative real-time PCR serves as a control for the amount of input DNA. Preferably, process A not only lacks CpG dinucleotides in the primers and probe(s), but also does not contain any CpGs within the amplicon at all to avoid any differential effects of the bisulfite treatment on the amplification process. Preferably, the amplicon for process A is a region of DNA that is not
15 frequently subject to copy number alterations, such as gene amplification or deletion.

Results obtained with the qualitative MethylLight™ version (process embodiment "B" of the technology) are described in the Examples below. Dozens of human tumor samples have been analyzed using this technology with excellent results.

20

Cancer Diagnostic and Prognostic Assays and Kits

Typically, diagnostic and/or prognostic assays of the present invention involve obtaining a tissue sample from a test tissue, performing a methylation assay on DNA derived
25 from the tissue sample to determine the associated methylation state, and making a diagnosis or prognosis based thereon.

In preferred embodiments, diagnostic and prognostic cancer assays are based on determination of the methylation state of one or more of the disclosed 20 gene sequences (*APC, ARF, CALCA, CDHI, CDKN2A, CDKN2B, ESR1, GSTP1, HIC1, MGMT, MLH1,*
30 *MYOD1, RB1, TGFBR2, THBS1, TIMP3, CTNNB1, PTGS2, TYMS* and *MTHFR*, or methylation-altered DNA sequence embodiments thereof), as defined herein by the oligomeric primers and probes corresponding to SEQ ID NOS:1-60, 64 and 65 (see TABLE II, below). SEQ ID NOS:61-63 correspond to the *ACTB* "control" gene region used in the present analysis (see EXAMPLE 1, below).

35 Additionally, other primers or probes corresponding to other sequence regions of the CpG islands associated with the *APC, ARF, CALCA, CDHI, CDKN2A, CDKN2B, ESR1,*

WO 01/75172

PCT/US01/10658

GSTP1, HIC1, MGMT, MLH1, MYOD1, RB1, TGFBR2, THBS1, TIMP3, CTNNB1, PTGS2 and *TYMS* sequence regions used herein may be used, based on the fact that the methylation state of a portion of a given CpG island is generally representative of the island as a whole.

Accordingly, the reagents required to perform one or more art-recognized methylation assays (including those described above) are combined with such primers and/or probes, or portions thereof, to determine the methylation state of CpG-containing nucleic acids.

For example, the MethyLight™, Ms-SNuPE, MCA, COBRA, and MSP methylation assays could be used alone or in combination, along with primers or probes comprising the sequences of SEQ ID NOS:1-65, or portions thereof, to determine the methylation state of a CpG dinucleotide within one or more of the 20 gene sequence regions corresponding to *APC, ARF, CALCA, CDHI, CDKN2A, CDKN2B, ESR1, GSTP1, HIC1, MGMT, MLH1, MYOD1, RB1, TGFBR2, THBS1, TIMP3, CTNNB1, PTGS2, TYMS* or *MTHFR*, or, in the case of 19 of these 20 sequence regions (*i.e.*, for all but *MTHFR*), to other CpG island sequences associated with these sequences, where such other CpG island sequences associated with these 19 gene sequences are those contiguous sequences of genomic DNA that encompasses at least one nucleotide of one of these 19 gene sequence regions, and satisfy the criteria of having both a frequency of CpG dinucleotides corresponding to an Observed/Expected Ratio >0.6, and a GC Content >0.5.

20

EXAMPLE 1

CpG Island Hypermethylation Increased with the Progression of EAC

This Example shows the results of an analysis of the methylation status of a panel of CpG islands associated with 19 different genes selected for their known involvement in carcinogenesis or because they have been shown to be methylated in other tumors (*see* Table 1, and under "Definitions," above), and of one non-CpG island sequence (*MTHFR* control sequence), for a total of 20 gene loci.

Quantitative methylation data of the 20 genes from a screen of 84 tissue specimens from 31 patients with different stages of Barrett's esophagus and/or associated adenocarcinoma showed a general increase in the frequency and in the quantitative level of CpG island hypermethylation at progressively advanced stages of disease. Accordingly, genes were grouped into distinct classes by their methylation behavior, based on both frequency and level of hypermethylation in various tissues (Figure 1).

35 Materials and Methods

Sample Collection and histopathologic examination. Multiple tissue samples (normal

WO 01/75172

PCT/US01/10658

esophagus (NE), normal stomach (S), intestinal metaplasia (IM), dysplasia (DYS) and/or adenocarcinoma (T)) from a total of 51 patients (range 39-86 years of age) with either adenocarcinoma or IM as the most advanced stage of disease were collected.

5 The initial set of samples analyzed included biopsies from 31 patients which were collected fresh and subdivided such that a part of each specimen was immediately frozen in liquid nitrogen and also embedded in paraffin for histopathologic examination by a pathologist (K.W.). Normal esophageal tissue was collected from every patient 10 cm or more away from the diseased areas. Frozen section examination of the frozen tissues was performed if the diagnosis was uncertain. The site of origin of the cancers was classified as
10 esophageal if the epicenter of the tumor was above the anatomic gastroesophageal junction, with the junction defined as the proximal margin of the gastric rugal folds. TNM staging was used to classify the stage of each adenocarcinoma.

A second set of samples were obtained for a follow-up study of 20 cases. Two groups of IM samples were collected: patients that had only IM as the most advanced stage of disease
15 (8 patients), and patients that had IM with associated dysplasia/adenocarcinoma located in another region of the esophagus (12 patients). H&E slides (5-micron sections) for each sample were prepared and examined by a pathologist (K.W.) to verify and localize the IM tissue. Cases that showed any signs of dysplasia or adenocarcinoma in the paraffin block used for analysis were excluded from this follow-up study. The IM tissues were carefully
20 microdissected away from other cell types from a 30-micron section adjacent to the 5-micron H&E section. All specimens were classified according to the highest grade histopathologic lesion present in that sample. Approval for this study was obtained from the Institutional Review Board of the University of Southern California Keck School of Medicine.

Nucleic Acid Isolation. Genomic DNA was isolated from the frozen tissue biopsies by
25 a simplified proteinase K digestion method (Laird et al., *Nucleic Acids Res.* 19:4293, 1991). The DNA from the paraffin tissues was extracted in lysis buffer (100 mM Tris-HCl, pH 8; 10 mM EDTA; and 1mg/ml Proteinase K) overnight at 50°C (Shibata et al., *Am. J. Pathol.* 141:539-543, 1992).

Sodium Bisulfite Conversion. Sodium bisulfite conversion of genomic DNA was
30 performed as previously described (Olek et al., *Nucleic Acids Res.* 24:5064-5066, 1996). The beads were incubated for 14 hours at 50°C to ensure complete conversion. Sodium bisulfite treatment converts unmethylated cytosines to uracil, while leaving methylated cytosine residues intact (Frommer et al., *Proc. Natl. Acad. Sci. USA* 89:1827-31, 1992).

MethyLight™ Analysis. After sodium bisulfite conversion, the methylation analysis
35 was performed by the fluorescence-based, real-time PCR assay MethyLight™, as described herein, and as previously described (Eads et al., *Cancer Res.* 60:5021-5026, 2000; Eads et al.,

WO 01/75172

PCT/US01/10658

Cancer Res. 59:2302-2306, 1999; Eads et al., *Nucleic Acids Res.* 28:E32, 2000). Two sets of primers and probes, designed specifically for bisulfite converted DNA, were used: a methylated set for the gene of interest and a reference set, beta-actin (*ACTB*) to normalize for input DNA. Specificity of the reactions for methylated DNA were confirmed separately using
 5 human sperm DNA (with very low levels of CpG island methylation) and *SssI* (New England Biolabs)-treated sperm DNA (heavily methylated) as previously described (Eads et al., *Cancer Res.* 60:5021-5026, 2000).

The percentage of fully methylated molecules at a specific locus was calculated by dividing the *GENE/ACTB* ratio of a sample by the *GENE/ACTB* ratio of *SssI*-treated sperm DNA and multiplying by 100. The abbreviation PMR (Percent of Methylated Reference) is used to indicate this measurement. The methylation analysis on the paraffin microdissected samples was performed following bisulfite treatment as described above by an investigator blind to the associated dysplasia status of the samples.

TABLE II lists the MethyLight™ primer and probe sequences (SEQ ID NOs:1-65),
 15 based on Genbank sequence data (except for SEQ ID NOs:64 and 65, see below), used in the present methylation analysis. Three oligos were used in every reaction: two locus-specific PCR primers flanking an oligonucleotide probe with a 5' fluorescent reporter dye (6FAM) and a 3' quencher dye (TAMRA) (Livak et al., *PCR Methods Appl.* 4:357-362, 1995). The Genbank accession number for each sequence is listed with the corresponding PCR amplicon
 20 location within that sequence. The %GC content, CpG observed/expected value and CpG:GpC ratio of 200 base pairs encompassing the MethyLight amplicon are indicated for each gene. The reaction type is designated "M" for methylation reaction and "C" for control reaction. The bisulfite treated DNA strand (top ("T") or bottom ("B")) and amplicon orientation (parallel ("P") or antiparallel ("A")) is also indicated. All primer and probe
 25 sequences are listed in the 5' to 3' direction. The numbers in brackets after each primer or probe sequence correspond to the associated SEQ ID NOs. The single asterisk (*) notes that there are two bases in our *CDKN2A* primers that differ from this GenBank sequence, since a preliminary high-throughput GenBank entry was the only available sequence at the time of applicants' primer design. The correct primers should be the following: forward,
 30 TGGAGTTTTCGGTTGATTGGIT (SEQ ID NO:64) and reverse,
 AACAACGCCGCACCTCCT (SEQ ID NO:65). The bases differing from the GenBank sequences are underlined. The double asterisk (**) indicates that the start site is not well defined.

WO 01/75172

PCT/US01/10658

TABLE II. MethyLight™ primer and probe sequences

HUGO Gene Name	Genbank Accession Number	Strand (Amplicon Location)	Amplicon Location (Genbank Numbering)	Amplicon Location (bp) Relative to Transcription Start	% GC Content	Obs/exp Cpg	Cpg:Gpc	Forward Primer Sequence (5'-3')	Probe Sequence (5'-3')	Reverse Primer Sequence (5'-3')
Class A' genes*										
<i>CDKN2A</i>	NM_000077	T	166-133bp	+9/+69	72	1.68	0.81	TGGAAATTTCC GTTGATTTGGTT [1]	6FAM- ACCGGACCCCGAACCGG G-TAMRA [3]	AACAAGGTCCG CACCTCCT [2]
<i>ESR1</i>	X62462	T	2784bp-2884bp	+14/+114	69	1.52	0.83	GGCGTTGGTTT TGGGATTG [4]	6FAM- CGATAAAACCGAACCGAC CCGACGA-TAMRA [6]	GCCGACACGGG AACTCTAA [5]
<i>MTOD1</i>	AF027148	T	9889bp-9962bp	-375/-302	61	1.12	0.74	GAGCGGCGGT AGTTAGCG [7]	6FAM-CTCCAACACCGGA CTACTATATCCGGGAAA- TAMRA [9]	TCCGACACGGC CTTTCC [8]
Class B' genes										
<i>CALCA</i>	X15943	T	1706bp-1806bp	-128/-28	60	0.72	0.53	GTTTGGAAAGT ATGAGGGTGAC G [10]	6FAM-ATTCCGGCCAATAC ACAACAACCAATAAACG- TAMRA [12]	TTCGGCGCGCTA TAAATCG [11]
<i>MGMT</i>	U95038	B	206bp-297bp	+392/+483	67	0.88	0.55	CTAAGGTATA CGAAAATCCGTA ACAACC [16]	6FAM-CCTTACCTCTAAAT ACCAACCCCAACCCG- TAMRA [18]	AGTATGAAGG TAGGAAGAATT CGG [17]
<i>TIMP3</i>	U33110	T	1051bp-1143bp	+1051/+1143	78	1.04	0.63	GGTCCGGAGGT TAAGGTTGTT [13]	6FAM- AACTCGCTCCGCCCGGA A-TAMRA [15]	CTCTCCAAAAT ACCGTACGGC [14]
Class C' genes										
<i>APC</i>	U02509	B	759bp-832bp	+7/+80	69	1.36	0.71	GAACCAAAACG CTCCCCAT [19]	6FAM- CCCGTCGAAAACCCGCC GATTA-TAMRA [21]	TTATATGTCGGT TACGTCCGTTTA TAT [20]

WO 01/75172

PCT/US01/10658

HUGO Gene Name	Accession Number	Strand	Amplicon Location (Genbank Numbering)	Amplicon Location (bp) Relative to Transcription Start	% GC Content	Obs/exp Cpg	Cpg:Gpc	Forward Primer Sequence (5'-3')	Probe Sequence (5'-3')	Reverse Primer Sequence (5'-3')
<i>Class D⁺ genes</i>										
<i>ARF</i>	AF082338	T	5447bp-5515bp	-203/-135	78	2	0.89	ACGGGGTITT CGGTAGTIT [22]	6FAM- CGACTCTAAACCCTACGC ACGCGAAA-TAMRA [24]	CCGAACCTCCA AAATCTCGA [23]
<i>CDH1</i>	L34545	T	842bp-911bp	-180/-111	69	1.2	0.68	AAITTTAAGTIT AGAGGGTATC GGGT [25]	6FAM- CGCCACCCCGACCTCGCA T-TAMRA [27]	TCCGCAAAAG AAAATAACGAC [26]
<i>CDKN2B</i>	S75756	T	350bp-430bp	+222/+302	64	1.44	0.82	AGTGGTTCG G [28]	6FAM- TTAACGACACTCTTCCCTT CTTCCCAAG-TAMRA [30]	CGAATAATCCA CGGTAAACCG [29]
<i>GSTP1</i>	M24485	T	1146bp-1245bp	-79/+21	75	2	0.93	GTCCGGTTCGT GATTTAGTAT G [31]	6FAM- TTCCGAACTTATAAAA- TAMRA [33]	AAACTACGAGG ACGAAACTCCA A [32]
<i>MLH1</i>	U26559	T	254bp-341bp	-662/-575	68	2.08	1.04	CGTTATATC GTTCGTAGTAT TCGTGTIT [35]	6FAM- CGCGACGTCAAAACGCCA CTACG-TAMRA [36]	CTATCGCGCGCT CATCGT [34]
<i>PTGS2</i>	AF044206	T	6779bp-6924bp	-362/-217	62	0.8	0.71	CGGAAAGCGTTC GGGTAAG [37]	6FAM- TTTCCGCCAAATATCTTT CTTCTTCGCA-TAMRA [39]	AAATCCACGGCC CCAAAC [38]
<i>THBS1</i>	U04835	B	1642bp-1716bp	-636/-562	75	2.08	1	CGACGGACCAA CCTAACCG [40]	6FAM- ACGCGCGCTCACCTCCC T-TAMRA [42]	GTTTTGAAGTTGG TTTATCGTTCGT T [41]
<i>Class E⁺ genes</i>										
<i>CTNWB1</i>	X89448	T	583bp-664bp	-413/-332	75	1.92	1	CGAAAGGCGC GTCCGAGT [43]	6FAM- CGCGGTTCCTCCGAAACCG -TAMRA [45]	TCCCCTATCCCA AAACCCG [44]

HUGO Gene Name	Genbank Accession Number	Strand	Amplikon Location (Genbank Numbering)	Amplikon Location (bp) Relative to Transcription Start	% GC Content	Obs/exp Cpg	Cpg:Gpc	Forward Primer Sequence (5'-3')	Probe Sequence (5'-3')	Reverse Primer Sequence (5'-3')
RB1	L11	T	178bp	-279/	68	2.24	1.27	TTAGTTGCGGT	6FAM-TCACGTCGCGGAAACTCC	ACTAAACGGCGG
	910		1900bp	-160				ATCGATTAGCG [46]	CGA-TAMRA [48]	CGTCCAA [47]
TGFB2	U52	T	256bp	-58/	75	2.08	0.81	GCGGGAGCG	6FAM-CAGGACGACGCCITCCC	CAAACCCGCT
	240		333bp	+19				TAGTTAGG [49]	GAA-TAMRA [51]	ACTCGTCAT [50]
TTMS	D00	T	1052bp	+301/	75	2.16	0.9	CGGCGTTAGGA	6FAM-CCGAAATACGGACA	TCTCAAACTATA
	517		1128bp	+377				AGGACGAT [52]	TAMRA [54]	ACGGGCTTACA
"Class F" genes										
HIC-1	L41	T	562bp	-55/	71	1.12	0.5	GTTAGGGGTT	6FAM-CAACATCGTCTAC	CGGAAGGCCTC
	919		662bp	+45**				AGGGGTC [58]	TAMRA [60]	CATCGTAT [59]
"Class G" genes										
MTHFR	AF1	T	123bp	-72/	57	0.56	0.54	TGGTAGTGAGA	6FAM-TCCTATACCGCTCAAAAT	CGCCTCATCTTC
	7		212bp	+45				GTTTTAAAGAT	CCAAACCCG-TAMRA [57]	TCCCGA [56]
"Control" gene										
ACTB	Y00	T	390bp	-1,599/	58	0.16	0.14	TGGTGATGGAG	6FAM-ACCACCCCAAC	AACCAATAAA
	474		522bp	-1,467**				GAGGTTTAGTA	TAMRA [63]	CCTACTCTCC
								AGI [61]		TTAA [62]

TABLE II cont. MethylLight™ primer and probe sequences. Numbers in brackets correspond to SEQ ID NOs: 1-63.

* Gene "Classes" are defined according to the present invention.

WO 01/75172

PCT/US01/10658

5 *Statistics.* The PMR values obtained by MethylLight™ (see above) were “dichotomized” at 4 PMR for statistical purposes as described previously (Eads et al., *Cancer Res.* 60:5021-5026, 2000). Dichotomization facilitates graphical representation, and moderates the quantitative impact of gene loci with different levels of hypermethylation, resulting in a more reliable cross-gene comparison of hypermethylation frequencies. Specifically, dichotomization equalizes the quantitative impact of methylated genes within each class (see “Epigenetic gene classes,” below), simplifying cross-gene comparisons of methylation frequencies.

10 A dichotomization point of 4 PMR was selected because it gave the best discrimination between normal and malignant tissues, across the board for all CpG islands (Eads et al., *Cancer Res.* 60:5021-5026, 2000). However, the precise dichotomization point does not significantly affect the statistics or alter the conclusions, and other dichotomization points are within the scope of the present invention (see below).

15 Accordingly, samples containing 4 PMR or higher were designated as methylated and given a value of 1, while samples containing less than 4 PMR were designated as unmethylated and given a value of 0. The cumulative value of genes methylated in each class (see “Epigenetic gene classes” A-G, herein below), or for all 19 genes was then used as a continuous variable in a Fisher’s Protected Least Significant Difference test, adapted for use with unequal sample sizes (SAS Statview software) to obtain *p*-values. The different parameters such as tissue type, presence of associated dysplasia, tumor stage, etc., were used as the nominal variables. The IM samples in the above-mentioned “follow-up” study of hypermethylation in IM, and the presence of associated dysplasia and/or carcinoma, were further dichotomized at 1 or fewer, versus two or more Class A genes methylated. A Fisher’s exact test was then used to determine statistical significance.

25

Results

CpG Island Hypermethylation and the Progression of EAC. The methylation status of a panel of CpG islands associated with 19 different genes and of one non-CpG island sequence for a total of 20 gene loci, was analyzed by the quantitative, high-throughput MethylLight™ assay (Eads et al., *Cancer Res.* 59:2302-2306, 1999; Eads et al., *Nucleic Acids Res.* 28:E32, 2000). The efficiencies of the methylation reactions were controlled for in each analysis by including unmethylated control DNA and methylated control DNA (Eads et al., *Cancer Res.* 60:5021-5026, 2000). The 20 genes were selected for their known involvement in carcinogenesis or because they have been shown to be methylated in other tumors (see Table 1, and under “Definitions,” above). We included a region located in the *MTHFR* gene as a “non-CpG island” control for a single copy sequence that does not satisfy the criteria (see

35

WO 01/75172

PCT/US01/10658

"Definitions," above) of a CpG island. CpG dinucleotides outside of an island are presumably normally methylated, unlike CpG dinucleotides within CpG islands.

Figure 1 illustrates the quantitative methylation data of the 20 genes from our screen of 84 tissue specimens from 31 patients with different stages of Barrett's esophagus and/or associated adenocarcinoma. Methylation analysis was performed using the MethyLight assay (Eads et al., *Cancer Res.* 59:2302-2306, 1999; Eads et al., *Nucleic Acids Res.* 28:E32, 2000). The percentage of fully methylated molecules at a specific locus (PMR = Percent of Methylated Reference) was calculated by dividing the *GENE/ACTB* ratio of a sample by the *GENE/ACTB* ratio of *SssI*-treated sperm DNA and multiplying by 100. The resulting percentages were then dichotomized at 4% PMR to facilitate graphical representation and to reveal tissue-specific patterns. The various squares, each having one of four possible shading intensity levels (see bottom axis of Figure 1), designate samples with less than 4 PMR, 4 - 20 PMR, 21 - 50 PMR and more than 51 PMR, where progressively increasing shading intensity levels correspond to progressively higher PMR values. The tissue types are shown on the left. The TNM tumor staging is designated by "1", "2", "3" and "4". The occurrence of distally located dysplasia and/or adenocarcinoma in the patient is indicated at the right of the figure by "YES" if present and "NO" if absent. "N" indicates an analysis for which the control gene *ACTB* did not reach sufficient levels to allow the detection of a minimal value of 1 PMR for that methylation reaction in that particular sample.

There was a general increase in the frequency and in the quantitative level of CpG island hypermethylation at progressively advanced stages of disease. However, the propensity for aberrant methylation of the genes was not uniform. Genes differed both in their frequency and in their levels of hypermethylation in various tissues.

Therefore, according to the present invention, genes can be grouped into classes based on their methylation behavior (Classes A-G, as shown at the right of Figure 1). This allowed for a visual assessment of concordant methylation of the different genes during various stages of tumorigenesis. A rationale for each of the gene classes is presented in the following section.

Epigenetic Gene Classes. The analysis of combined behavior of genes with different levels of DNA methylation would, without appropriate data treatment, be expected to lead to a bias of the group behavior towards genes with quantitatively high levels of DNA methylation. For instance, the mean values for gene "Class B" for most of the tumor samples would be driven primarily by the *TIMP3* values, since this gene tended to have higher levels of methylation than the other two genes in this group (see Figure 1).

Therefore, the methylation values used to generate Figure 1 were collapsed into a

WO 01/75172

PCT/US01/10658

binary variable with a dichotomization point of 4 PMR to equalize the quantitative impact of methylated genes within each epigenetic class. Samples containing 4 PMR or higher were designated as methylated and given a value of 1, while samples containing less than 4 PMR were designated as unmethylated and given a value of 0 (see "Statistics" above, under
5 "Materials and Methods"). This dichotomization moderates the effect of highly methylated genes, simplifies cross-gene comparisons of methylation frequencies, as shown in Figure 2, and allowed the calculation of class averages of methylation frequencies as shown in Figure 3 (below).

Figure 2 shows the percent of samples methylated for each gene by tissue type. The
10 data was dichotomized at 4 PMR, with 4 PMR and higher designated as methylated, and below 4 PMR as unmethylated. The genes, according to the present invention, were grouped according to their respective epigenetic gene classes (A-G) as shown in Figure 1. The letter "n" equals the number of samples analyzed for each tissue.

The suitability of the 4 PMR dichotomization point was based on its ability to
15 discriminate between the different tissue types, as shown in Figures 1-3 (see also Klump et al., *Gastroenterology*. 115:1381-1386, 1998). Other dichotomization point values are within the scope of the present invention, where such dichotomization point values moderate the statistical effects of highly methylated genes, simplify cross-gene comparisons of methylation frequencies, and facilitate calculation of class averages of methylation frequencies. For
20 instance, there is still a statistically significant difference in the mean percent of genes methylated (out of 19 genes) between the normal esophageal mucosa and the IM ($p = 0.0003$), DYS ($p < 0.0001$) and T ($p < 0.0001$) tissues when the data is dichotomized at 10 PMR.

Additionally, all of the statistically significant findings of the NE and IM methylation
25 frequency with or without associated dysplasia (see Example 3, below) remain significant at a dichotomization point of 10 PMR, instead of 4 PMR. It is important to note that 4 PMR is not comparable to a 4% methylation level of a single CpG dinucleotide. Rather, it indicates that in this sample, 4% of the DNA molecules had *complete* methylation at all CpG dinucleotides covered by the three MethyLight™ primers (usually about 8 CpGs). The nature of the MethyLight™ assay is such that it is oblivious to all other methylation patterns that may be
30 present (Eads et al., *Nucleic Acids Res.* 28:E32, 2000).

Therefore, 4 PMR is likely to represent a higher mean level of methylation than 4%. The extensively methylated molecules that are assayed by MethyLight™ are likely to represent alleles that have been completely silenced by CpG island hypermethylation, although this was not investigated herein.

35 Of the panel of 20 genes, the most informative genes were those with an intermediate frequency of hypermethylation (ranging from 15% (*CDKN2A*) to 60% (*MGMT*) of the sample

WO 01/75172

PCT/US01/10658

values above the 4 PMR methylation cutoff). This group was further subdivided into three epigenetic gene classes according to the absence (Class "A") or presence (Class "B") of methylation in normal esophageal mucosa and stomach, or the infrequent methylation of normal esophageal mucosa accompanied by methylation in all normal stomach samples (Class "C"). The other genes were less informative, since the incidence of hypermethylation was either very infrequent (Class "D"), completely absent (Class "E"), or ubiquitous (Classes "F" and "G") regardless of tissue type (Figures 1, 2 and 3).

Epigenetic gene Class A comprises the genes *CDKN2A*, *ESR1* and *MYOD1* (Figures 1, 2 and 3). There was a statistically significant difference in the methylation frequency of *ESR1* ($p = 0.0001$) and *MYOD1* ($p = 0.0038$) of normal esophagus (NE), as compared to IM tissue, but not for *CDKN2A* ($p = 0.097$). The frequency of *CDKN2A* methylation increased significantly in the more advanced stages of the adenocarcinoma (T) ($p < 0.0001$).

Epigenetic gene Class B comprises the genes *CALCA*, *MGMT* and *TIMP3*. In contrast to Class A, this class exhibited methylation in the normal esophageal mucosa (NE) and stomach (S) tissue (Figures 1 and 2). Only *TIMP3* showed a significant difference in methylation frequency between the NE and IM values ($p = 0.0074$).

Epigenetic gene Class C comprises the gene *APC* which was, in contrast to genes of Classes A and B, methylated in all normal stomach samples (Figures 1 and 2). This confirms previous documentation of *APC* methylation in normal stomach tissue (Eads et al., *Cancer Res.* 60:5021-5026, 2000). The mechanism which protects *APC* from methylation in the normal esophageal tissues (NE) but not in normal stomach tissues (S) is not clear.

Epigenetic gene Class D comprises the genes *ARF*, *CDHI*, *CDKN2B*, *GSTP1*, *MLH1*, *PTGS2* and *THBS1*, which were infrequently methylated (Figures 1 and 2). There was a slight increase in the frequency of this class of genes in adenocarcinoma (T), but this did not approach statistical significance (Figure 3). Interestingly, with the exception of *PTGS2*, which has not yet been investigated in other systems, the remaining Class D genes are frequently hypermethylated in other tumor types (Table 2).

Epigenetic gene Class E comprises the *CTNNB1*, *RBI*, *TGFBR2* and *TYMS1* genes, which were unmethylated at each stage in the progression of EAC. Similar to most Class D genes, *RBI* and *TGFBR2* have been found to be hypermethylated in other tumors types (see Table 1, and literature references under "DEFINITIONS" herein above). It should be noted that all samples scored positive for DNA input as measured by the control gene (*ACTB*). Therefore, the lack of detectable DNA methylation cannot be attributed to a lack of input DNA. The control reaction was sufficient in each sample, so that a level as low as 1 PMR for a given test gene could be detected. The integrity and specificity of all methylation reactions was confirmed using *in vitro* methylated human DNA.

WO 01/75172

PCT/US01/10658

- The epigenetic Class F comprises the *HIC1* gene, which was completely methylated, regardless of tissue type (Figures 1 and 2). *HIC1* is commonly methylated in other types of cancers (Jones & Laird, *Nat. Genet.* 21:163-167, 1999; Baylin & Herman, *Trends Genet.* 16:168-174, 2000), and has been shown to be methylated in normal breast ductal tissue and bone marrow samples of breast cancer and AML patients, respectively (Melki et al., *Cancer Res.* 59:3730-3740, 1999; Fujii et al., *Oncogene.* 16:2159-2164, 1998). Nevertheless, the finding of ubiquitous methylation of a CpG island in normal tissues was unexpected. Therefore, the validity of the *HIC1* MethyLight™ results was confirmed using a different technique (HpaII-PCR) (Singer-Sam et al., *Nucleic Acids Res.* 18:687, 1990).
- Epigenetic Class G comprises the non-CpG island *MTHFR* gene, used herein as a control. Interestingly, the ubiquitous *HIC1* methylation pattern is similar to the non-CpG island *MTHFR* control (Class G), however the percentage of methylated molecules was quantitatively higher for *HIC1* (Figure 1).
- Epigenetic Profiles of EAC Progression.* Each tissue type showed a unique epigenetic profile or fingerprint that changed during disease progression (Figure 3, upper panel). Figure 3 shows a comparison of epigenetic profiles according to the present invention. The data was dichotomized at 4 PMR, with 4 PMR and higher designated as methylated, and below 4 PMR as unmethylated. Error bars represent the standard error of the mean. *Upper panel:* Mean percent of genes methylated in each gene Class (A-F or ALL 19 CpG islands) by tissue type (N, normal esophagus; S, stomach; IM, intestinal metaplasia; DYS, dysplasia; T, adenocarcinoma). The error bars represent the standard error of the mean (SEM). *Lower panel:* Statistical analysis of the difference in mean percent of genes methylated in different tissues by gene Class (A-F) or for all 19 CpG islands combined (ALL). The *p*-values were generated by a Fisher's Protected Least Significant Difference (PLSD) test, adapted for use with unequal sample numbers (SAS Statview™ software).
- Classes A, B and C were methylated at a significantly higher frequency in IM tissue than in normal esophageal mucosa (NE) (Figure 3, upper and lower panels). Furthermore, the transition from IM to dysplasia (DYS) or malignancy (T) was associated with an additional increase in Class A methylation (Figure 3, upper and lower panels). The lack of a significant difference between dysplasia and adenocarcinoma for any of the gene classes or when all 19 genes are combined (Figure 3, upper and lower panels) suggests that most of these abnormal epigenetic alterations occur early in the progression of EAC.
- In summary of this Example.* According to the present invention, quantitative methylation data of 20 genes (Tables I and II, above) from a screen of 84 tissue specimens

WO 01/75172

PCT/US01/10658

from 31 patients with different stages of Barrett's esophagus and/or associated adenocarcinoma showed a general increase in the frequency and in the quantitative level of CpG island hypermethylation at progressively advanced stages of disease (Figures 1-3, above).

5 Additionally, genes were grouped into novel epigenetic classes based on their methylation behavior (Classes A-G, as shown herein in Figures 1-3) during tumor progression. This allowed for graphical representation of concordant methylation of the different genes during various stages of tumorigenesis, which can be readily appreciated by means of a simple visual assessment.

10 Each tissue type showed a unique epigenetic profile or fingerprint that changed during disease progression (Figure 3, upper panel). Classes A, B and C were methylated at a significantly higher frequency in IM tissue than in normal esophageal mucosa (NE) (Figure 3, upper and lower panels). Furthermore, the transition from IM to dysplasia (DYS) or malignancy (T) was associated with an additional increase in Class A methylation (Figure 3, upper and lower panels).

EXAMPLE 2

Hypermethylation was Reflective of EAC Tumor Grade and Stage

20 This Example examines whether the grade or stage of an esophageal adenocarcinoma correlates with a higher frequency of CpG island hypermethylation. According to the present invention, for EAC, epigenetic Class A gene methylation is significantly higher in stage II, III and IV tumors relative to less advanced stage I tumors (Figure 4).

Materials and Methods

25 *TNM staging.* The American Joint Committee on Cancer ("AJCC") has designated staging by TNM classification (Tumor; Lymph Node metastasis, distant Metastasis). TNM staging was used to classify the stage of each esophageal adenocarcinoma from the tissues of Example 1.

30 *Methylation and statistical analysis.* Methylation and statistical analysis was as described herein under Example 1.

Results

35 *Methylation of epigenetic Class A genes increases with tumor stage.* Moderately differentiated tumors have significantly less frequent Class A methylation compared to poorly differentiated tumors ($p = 0.045$). Additionally, Figure 4 (upper and lower panels) shows that there is a significantly higher mean number of Class A genes methylated in stage II, III and

WO 01/75172

PCT/US01/10658

IV tumors relative to less advanced stage I tumors. The differences between stage I tumors and stage II, III and IV tumors did not reach statistical significance for any of the other epigenetic gene classes.

5 Figure 4 shows the relationship between Class A methylation frequency and tumor stage according to the present invention. The data was dichotomized at 4 PMR, with 4 PMR and higher designated as methylated, and below 4 PMR as unmethylated. *Upper panel:* Mean number of genes methylated for Class A with respect to tumor stage (I-IV) is shown (see Figure 1). The error bars represent the standard error of the mean (SEM). The letter "n" equals the number of samples analyzed in each tumor stage. *Lower panel:* Statistical analysis
10 of the difference in mean number of Class A genes methylated by tumor stage. The *p*-values were generated by a Fisher's Protected Least Significant Difference (PLSD) test, adapted for use with unequal sample numbers (SAS Statview™ software).

In summary for this Example. According to the present invention, in addition to the
15 epigenetic profiles or fingerprints (comprising the gene classes disclosed herein) that can be used to assess oncogenic progression, the mean number of methylated Class A genes can be used to assess the relative stages of EAC tumors.

EXAMPLE 3

20 Methylation of Premalignant Tissues With or Without Associated Dysplasia

This Example shows that the frequency of Class B methylation in the normal esophagus (NE) was found to be significantly higher in patients with associated dysplasia/tumor ($p = 0.0037$) (Figure 1). Additionally, Class A methylation was found to be more frequent in IM samples from patients with concurrent dysplasia or cancer, than in IM
25 samples from patients without any evidence of further progression ($p < 0.0001$) (Figures 1 and 5). That is, there was a significant positive association between hypermethylation of epigenetic Class A genes in IM tissue, and the presence of associated dysplasia or cancer (Figure 5).

30 Materials and Methods

Histopathology. Histopathological classification was as described under "Materials and Methods," Example 1 above.

Methylation and statistical analysis. Methylation and statistical analysis was as described herein under Example 1.

35

Results

WO 01/75172

PCT/US01/10658

Methylation of Premalignant Tissues with or without Associated Dysplasia. The occurrence, according to the present invention, of CpG island hypermethylation in some cases of IM for Class A and some cases of normal esophageal mucosa for Class B raised the question whether these methylation events represent normal methylation patterns in these
5 non-dysplastic tissues, or whether they reflect methylation changes that predispose cells to further progression. In the latter case, one would expect to find a higher frequency of such CpG island hypermethylation in these tissues in patients who have already undergone further disease progression. Therefore, the frequency of such CpG island hypermethylation was compared between tissues (of the present study) with or without associated dysplasia.

10 In the initial study, patients were divided based on whether or not they had Barrett's esophagus (IM) as their most advanced stage of disease (Figure 1, "NO") or whether they had associated dysplasia and/or adenocarcinoma present in a different region of the esophagus (Figure 1, "YES"). The frequency of Class B methylation in the normal esophagus (NE) was indeed found to be significantly higher in patients with associated dysplasia/tumor ($p = 0.0037$) (Figure 1). Additionally, Class A methylation was found to be more frequent in IM
15 samples from patients with concurrent dysplasia or cancer, than in IM samples from patients without any evidence of further progression ($p < 0.0001$) (Figure 1).

A potential criticism of this analysis is that the same set of samples was used to delineate the class of genes, as was used to test the association with a clinical parameter.
20 Therefore, a follow-up study of 20 additional cases of IM was performed entirely independent of the first data set.

In the follow-up study of 20 cases, two groups of IM samples were collected: patients that had only IM as the most advanced stage of disease (8 patients), and patients that had IM with associated dysplasia/adenocarcinoma located in another region of the esophagus (12
25 patients). H&E slides (5-micron sections) for each sample were prepared and examined by a pathologist (K.W.) to verify and localize the IM tissue. Cases that showed any signs of dysplasia or adenocarcinoma in the paraffin block used for analysis were excluded from this follow-up study. The IM tissues were carefully microdissected away from other cell types from a 30-micron section adjacent to the 5-micron H&E section. All specimens were
30 classified according to the highest grade histopathologic lesion present in that sample.

The initial study had revealed that all IM samples associated with further disease progression ("YES") had at least two Class A genes methylated, while all IM samples without associated dysplasia or adenocarcinoma ("NO") did not show any methylation of Class A genes (Figure 1, under "Barrett's (IM)"). Therefore, a state of having two or more Class A
35 genes methylated was defined as an indicator of increased risk for the presence of associated dysplasia or cancer.

WO 01/75172

PCT/US01/10658

The data from our first series gave a p -value of 0.0048 in a Fisher's exact test of this association (Figure 5, left panel). The follow-up series of 20 independent cases gave a p -value of 0.018 (Figure 5, right panel).

Figure 5 shows the percent of two or more Class A genes methylated in intestinal metaplasia ("IM") tissues with ("Y"), or without ("N") associated dysplasia and/or adenocarcinoma. The data was dichotomized at 4 PMR, with 4 PMR and higher designated as methylated, and below 4 PMR as unmethylated. *Left panel:* Class A methylation in the IM data illustrated in Figure 1. *Right panel:* Class A methylation in the IM for a completely independent follow-up study of twenty different microdissected IM samples. The error bars represent the standard error of the mean (SEM). The letter "n" equals the number of samples analyzed in each tissue group.

Therefore, the positive association between hypermethylation of Class A genes and the presence of associated dysplasia or cancer is significant. It should be noted that the IM samples without associated dysplasia in this follow-up study (Figure 5, right panel) showed a low frequency of samples with at least two genes methylated, which is in contrast to the absence of methylation in the first study (Figure 1, and Figure 5, left panel). This may be attributed to the fact that the samples in the second series were microdissected from paraffin sections. Therefore, there is a lower background of unmethylated stromal cells in the sample. In this case, the methylation signal is not as diluted by other normal cells and consequently the ratio of methylated molecules to total DNA may rise above the 4 PMR threshold. Alternatively, dysplastic or malignant tissue may have been missed during the endoscopic survey in some of the cases scored as free of further disease progression due to the sampling limitations of endoscopy. This is a well-documented problem in the detection of esophageal adenocarcinoma (Peters et al., *J. Thorac. Cardiovasc. Surg.* 108:813-821, 1994).

EXAMPLE 4

No Clear Evidence of CpG Island Methylator Phenotype ("CIMP") for EAC

This Example shows that, for the present study of EAC, there was no clear evidence of a separate group of CIMP tumors, as has been previously defined for colorectal and gastric cancer (Toyota et al., *Proc. Natl. Acad. Sci. U.S.A.* 96:8681-8686, 1999; Toyota et al., *Cancer Res.* 59:5438-5442, 1999). However, CpG island hypermethylation in EAC did occur across multiple loci in a given sample. Furthermore, the number of loci hypermethylated in a single sample increased as the disease progressed through different histological stages (Figure 6). The bimodal distributions seen in IM tissues (Figure 6) can be fully attributed to the concurrent association with dysplasia or cancer described herein above.

WO 01/75172

PCT/US01/10658

Materials and Methods

Histopathology. Histopathological classification was as described under "Materials and Methods," Example I above.

Methylation and statistical analysis. Methylation and statistical analysis was as described herein under Example 1.

Results

CIMP Analysis. It has previously been reported that a subset of colorectal and gastric tumors display a CpG island methylator phenotype ("CIMP"), characterized by widespread, aberrant hypermethylation changes affecting multiple loci in a single tumor (Toyota et al., *Proc. Natl. Acad. Sci. USA* 96:8681-8686, 1999; Toyota et al., *Cancer Res.* 59:5438-5442, 1999). This is reflected in a bimodal distribution of the frequency of the number of genes methylated in a group of tumors (Toyota et al., *Proc. Natl. Acad. Sci. USA* 96:8681-8686, 1999). CIMP tumors are a distinct group of tumors that are defined by a high degree of concordant CpG island hypermethylation of genes exclusively methylated in cancer, or "type-C" genes. CIMP is currently thought to be a new, distinct, yet major pathway of tumorigenesis (Toyota et al., *Proc. Natl. Acad. Sci. USA* 96:8681-8686, 1999; Toyota et al., *Cancer Res.* 59:5438-5442, 1999).

Therefore the question of whether esophageal adenocarcinoma tumors exhibit a CpG island methylator phenotype (CIMP) was investigated.

Class A genes of the present invention most closely exemplify the "type-C" genes, because they lack methylation in the normal tissues. The distribution of the number of Class A genes methylated was examined for EAC (Figure 6).

Figure 6 shows, according to the present invention, methylation frequency distributions in the progression of esophageal adenocarcinoma. The data was dichotomized at 4 PMR, with 4 PMR and higher designated as methylated, and below 4 PMR as unmethylated. The proportion of patients with zero to three (Class A), zero to nine (Classes A + D) and zero to fourteen CpG islands (Classes A + B + C + D) methylated in each tissue is shown. Class E and F CpG islands were not included since there was no variation in the frequency of methylation between the different tissue. The letter "n" equals the number of samples analyzed in each tissue.

However, the frequency of genes methylated in the adenocarcinoma tissue did not show the expected bimodal distribution of CIMP (Figure 6) (Toyota et al., *Cancer Res.* 59:5438-5442, 1999). Similar results were observed when Class D genes, which also exhibit type C methylation, were included along with Class A (Figure 6, middle panel) and when Classes A, B, C and D genes were combined (Figure 6, right panel). Classes E and F genes

WO 01/75172

PCT/US01/10658

were not included since they did not exhibit any methylation variation between the different tissue types.

There was a single sample with 10 out of 14 Class A-D genes methylated (Figure 1, Case #3 and Figure 6). However, this sample only stands out when Class B genes, which are methylated in normal esophageal mucosa and therefore do not satisfy the definition of "type-C" genes that constitute the CIMF phenotype, are included.

Therefore, there was no clear evidence of a separate group of CIMF tumors in the present study of esophageal adenocarcinoma, as has been previously defined for colorectal and gastric cancer.

However CpG island hypermethylation in EAC did occur across multiple loci in a given sample. Furthermore, the number of loci hypermethylated in a single sample increased as the disease progressed through different histological stages (Figure 6). The bimodal distributions seen in IM tissues (Figure 6) can be fully attributed to the concurrent association with dysplasia or cancer described herein above.

15

EXAMPLE 5

Array- and Microarray-based Applications

Microarray-based embodiments are within the scope of the present invention. For example, one such array-based embodiment uses differential methylation hybridization ("DMH") (Huang et al., *Hum. Mol. Genet.*, 8:459-470, 1999; Yan et al., *Clin. Cancer Res.* 6:1432-38, 2000). DMH is applied to screen paired test and normal samples and to determine whether patterns (see "Epigenetic patterns," herein under Example 1) of specific epigenetic alterations correlate with pathological parameters in the tissue samples analyzed.

"Amplicons" (*Id.*), representing a pool of methylated CpG DNA derived from these samples, are used as hybridization probes in an array panel containing the CpG island tags of the present invention.

Accordingly, one or more of the CpG island sequences associated with 19 of the 20 disclosed gene sequences (*i.e.*, *APC*, *ARF*, *CALCA*, *CDH1*, *CDKN2A*, *CDKN2B*, *ESR1*, *GSTP1*, *HIC1*, *MGMT*, *MLH1*, *MYOD1*, *RBI*, *TGFBR2*, *THBS1*, *TIMP3*, *CTNNB1*, *PTGS2* and *TYMS* (see TABLES I and II, above; and see under "Definitions," above), or methylation-altered DNA sequence embodiments thereof, can be used as CpG island tags in an array or microarray-based assay embodiment. These 19 gene sequence regions are defined herein by the oligomeric primers and probes corresponding to SEQ ID NOs:1-54, 58-60, 64 and 65 (see TABLE II, above; SEQ ID NOs:61-63 correspond to the *ACTB* "control" gene region used in the present analysis (see EXAMPLE 1, below)). Associated CpG island sequences are (based on the fact that the methylation state of a portion of a given CpG island

WO 01/75172

PCT/US01/10658

is generally representative of the island as a whole) those contiguous sequences of genomic DNA that encompass at least one nucleotide of the sequences defined by these specific oligonucleotide primers and probes, and satisfy the criteria of having both a frequency of CpG dinucleotides corresponding to an Observed/Expected Ratio >0.6, and a GC Content >0.5.

5 These CpG island tags are then arrayed on solid supports (e.g., nylon membranes, silicon, etc.), and probed with amplicons representing a pool of methylated CpG DNA, from test (e.g., tumor) or reference samples. The differences in test and reference signal intensities on screened CpG island arrays reflect methylation alterations of corresponding sequences in the test DNA.

10 Comparison of the resulting data with the epigenetic patterns disclosed herein allows for a diagnostic or prognostic determination.

Therefore, according to this embodiment, pattern analysis (see working Examples 1-4, below) in a subset of CpG island tags, affixed to a solid support to form an array or microarray, is used to follow progression during various stages of cancer progression (e.g., 15 gastrointestinal and esophageal dysplasia, gastrointestinal and esophageal metaplasia, Barrett's esophagus, and pre-cancerous conditions in normal esophageal squamous mucosa), and can be used to determine histological grades or stages of tumors, such as esophageal adenocarcinoma.

Other array or microarray embodiments of the present invention will be obvious to 20 those of ordinary skill in the relevant art. Such embodiments include, but are not limited to those wherein the specific primers and/or probes for *APC*, *ARF*, *CALCA*, *CDH1*, *CDKN2A*, *CDKN2B*, *ESR1*, *GSTP1*, *HIC1*, *MGMT*, *MLH1*, *MYOD1*, *RB1*, *TGFBR2*, *THBS1*, *TIMP3*, *CTNNB1*, *PTGS2* and *TYMS* (see TABLES I and II, above; and see under "Definitions," above), corresponding to SEQ ID NOs:1-54, 58-60, 64 and 65 (see TABLE II, above; SEQ ID 25 NOs:61-63 correspond to the *ACTB* "control" gene region used in the present analysis (see EXAMPLE 1, above)) are arrayed on solid supports.

DISCUSSION

30 There is a need in the art for novel and more sensitive methods of cancer detection, chemoprediction and prognostics. There is a need in the art to define novel coordinate patterns of CpG island methylation changes (i.e., novel epigenetic patterns) at multiple loci during progression of a disease, such as cancer. There is a need in the art to determine tumor- 35 type-specific, and patient-specific epigenetic patterns or fingerprints. There is a need in the art to provide biomarkers or probes, such as EAC-specific biomarkers or probes, that can be used in diagnostic and/or prognostic methods for the treatment of cancer. There is a need in

WO 01/75172

PCT/US01/10658

the art to determine whether esophageal adenocarcinoma displays a CIMP. There is a need in the art for novel methods for determining the stage of a tumor. The present invention addresses these needs.

5 A high-throughput, fluorescence-based methylation assay (MethyLight™) was used herein to examine and define novel hypermethylation patterns of 19 CpG islands and one non-CpG island during the progression of esophageal adenocarcinoma ("EAC"). The genes were thereby segregated into six classes of epigenetic patterns in the various tissue types. This is the most comprehensive methylation survey yet performed on a system having so many distinct histological stages of disease progression. Furthermore, the present analysis of
10 abnormal DNA hypermethylation offers a significant advantage over other approaches, such as gene expression analysis, in that it has greater sensitivity in the presence of contaminating normal cells, a common limiting factor.

DNA hypermethylation, as disclosed herein, is an early epigenetic alteration in the multi-step progression of EAC. The premalignant intestinal metaplasia ("IM," or Barrett's
15 esophagus) is already significantly more methylated than the normal tissue (normal squamous mucosa). The present invention, in certain embodiments, provides the novel finding of frequent hypermethylation of five additional genes in this tumor system: *MYOD1*, *MGMT*, *CALCA*, *TIMP3*, and *HIC1*.

The methylation observed for *MGMT*, *TIMP3*, and *HIC1* in normal tissues may be attributed to the particular region of the gene in which we analyzed methylation levels (Stoger
20 et al., *Cell*. 73:61-71, 1993; Larsen et al., *Hum. Mol. Genet.* 2:775-80, 1993; Jones, P. A., *Trends Genet.* 15:34-37, 1999). These three genes were analyzed at CpG islands located at or downstream of the transcription start site (TABLE 2). However, this does not account for the *CALCA* methylation we observed, because we analyzed the promoter region of this gene.
25 Low levels of *CALCA* methylation has been previously reported in normal bone marrow samples of AML patients (Melki et al., *Cancer Res.* 59:3730-3740, 1999), suggesting that this locus may have a higher propensity to be methylated in normal tissues of cancer patients.

It is of particular interest to note that dysplastic tissues are more frequently methylated than stage I tumors for both Class A ($p < 0.0001$) and B ($p = 0.0174$) (Figure 1). This is
30 similar to the finding of genetic abnormalities (LOH, deletions and mutations) present in Barrett's esophagus with high grade dysplasia but not present in the adjacent invasive EAC (Barrett et al., *Nat. Genet.* 22:106-109, 1999). Because stage II-IV tumors appear to be methylated at Class A genes at a similar frequency as dysplasia, this suggests that stage I tumors may actually evolve from a different origin than the dysplastic tissue and higher staged tumors, or may diverge after dysplasia independently from stage II-IV tumors during
35 clonal expansion. Alternatively, but less likely, stage I tumors could undergo a transient

WO 01/75172

PCT/US01/10658

reversal of hypermethylation. Tumor development in Barrett's esophagus is proposed to evolve clonally through the linear multistep pathway of metaplasia-dysplasia-tumor (Zhuang et al., *Cancer Res.* 56:1961-4, 1996). However, the occurrence of genetic and, according to the present invention, epigenetic alterations in a non-linear order, indicates that the clonal evolution of EAC is more complex than originally predicted (Barrett et al., *Nat. Genet.* 22:106-109, 1999). A similar observation has been described for different stages of bladder tumors (Salem et al., *Cancer Res.* 60:2473-2476, 2000).

There was, under the present analysis, no clear evidence, aside from one tumor with 10 genes methylated, for a separate cluster of tumors with extensive concordant methylation, indicative of a CpG island methylator phenotype ("CIMP"). Similar results were obtained even when only "type-C" genes, as defined for CIMP (methylated in cancer, not methylated in normal tissues; Toyota et al., *Proc. Natl. Acad. Sci. USA* 96:8681-8686, 1999; Toyota et al., *Cancer Res.* 59:5438-5442, 1999), were examined. Interestingly, the "type-C" genes in EAC differ from those described for colorectal cancer (*Id.*). For example, *ESR1* is classified as a "type-A" (defined as methylated in aging normal tissues) rather than a "type-C" gene in colorectal cancer, because it is frequently methylated in the normal colonic epithelium of aging individuals (*Id.*). However, in esophageal adenocarcinoma, *ESR1* clearly behaves as a "type-C" gene. This may be attributed to the difference in the technology used to measure hypermethylation, or more likely may be due to differences in tissue types.

According to the present invention, there is a tissue-specific and tumor-specific propensity for particular genes to become hypermethylated. For instance, *APC* is hypermethylated in normal stomach, but not in normal esophageal mucosa. The tumor-specificity of hypermethylation is illustrated by the lack of detectable methylation of the two Class E genes *TGFB2* and *RBI*, which are frequently hypermethylated in gastric and lung tumors, and retinoblastoma tumors, respectively (Stirzaker et al., *Cancer Res.* 57:2229-2237, 1997; Kang et al., *Oncogene* 18:7280-7286, 1999; Hougaard et al., *Br. J. Cancer* 79:1005-1011, 1999).

The tumor-specificity of CpG island hypermethylation suggests that there may be tissue-specific trans-acting factors that modulate methylation changes of these CpG islands during tumorigenesis and which differ between esophageal adenocarcinomas and other tumor types. Alternatively, there may be a lack of selective advantage to the silencing of these genes in esophageal adenocarcinomas by DNA methylation. There are two scenarios in which this would be the case. One is if the gene in question has been inactivated by a different, genetic mechanism, rendering hypermethylation of no further selective advantage. The other is if the gene does not play a role in tumor suppression in this particular tumor system.

Although alterations in DNA methylation changes are common events in

WO 01/75172

PCT/US01/10658

tumorigenesis, the underlying mechanism is unclear. Abnormal methylation, at least in colorectal tumors, is not due to a mere upregulation of the DNA methyltransferase genes, suggesting that other major players are involved (Eads et al., *Cancer Res.* 59:2302-2306, 1999). The present invention provides some first glimpses into the process underlying these

5 abnormal methylation changes.

According to the present invention, different, functionally unrelated, genes can behave in distinct classes with respect to their methylation changes within various tissues of EAC progression. The CpG island hypermethylation does not appear to be a random, stochastic process (although there is a stochastic component), but rather a step-wise process that

10 involves multiple, distinct groups of alterations. This is consistent with the existence of several different mechanisms that protect against CpG island hypermethylation. In this scenario, the concerted changes seen at different CpG islands would be the result of the loss of a different type of protective element at different stages of disease progression. This finding does not appear to be dependent on the location of the CpG island relative to the gene, since both promoter and internal CpG islands were observed in all gene classes. The

15 structural features of these CpG islands were also examined under the present analysis by analyzing the %GC content, the observed/expected CpG ratio and the CpG:GpC ratio and found no association with gene class (TABLE 2).

According to the present invention, the IM or NE samples themselves, with or without associated dysplasia or cancer, were histologically indistinguishable, yet molecularly distinct. NE and IM samples derived from individuals with concurrent distally located dysplasia or malignancy show a statistically higher incidence of CpG island hypermethylation. These findings were confirmed herein in the IM tissues in a completely independent study. This

20 provides strong support for the use of epigenetic markers, particularly Class A and B genes, as disease screening tools and as predictive markers for the progression of more advanced staged

25 disease.

The methylation profiles of the present invention provide methods and compositions for the early detection of cancer. Such a molecular diagnostic approach using normal and/or premalignant tissues to identify patients with cancer or at elevated risk for developing cancer

30 provides an opportunity for early intervention. Furthermore, a benefit of using CpG island hypermethylation as a diagnostic or prognostic marker is that it can easily be detected in a field of normal cell contamination as a gain of signal, unlike loss of gene expression (*e.g.*, LOH and deletion analysis), which is difficult to resolve in a sample with contaminating normal cells.

WO 01/75172

PCT/US01/10658

SUMMARY

According to the present invention, the 19 CpG islands (TABLES I and II) studied segregate into six classes of epigenetic patterns in the various tissue types. Each class
5 undergoes unique epigenetic changes at different steps of disease progression of EAC. The methylation profiles provide methods and compositions for the early detection of cancer.

WO 01/75172

PCT/US01/10658

I claim:

1. A method for diagnosing cancer or cancer-related conditions from tissue samples, comprising:
 - 5 (a) obtaining a tissue sample from a test tissue or region to be diagnosed;
 - (b) performing a methylation assay of the tissue sample, wherein the methylation assay determines the methylation state of genomic CpG sequences, wherein the genomic CpG sequences are located within at least one gene sequence selected from the group consisting of *APC*, *ARF*, *CALCA*, *CDHI*, *CDKN2A*, *CDKN2B*, *ESR1*, *GSTP1*, *HIC1*, *MGMT*, *MLH1*, *MYOD1*, *RB1*, *TGFBR2*, *THBS1*, *TIMP3*, *CTNNB1*, *PTGS2*, *TYMS* and *MTHFR*, and combinations thereof; and
 - 10 (c) making a diagnostic or prognostic prediction of the cancer based, at least in part, upon the methylation state of the genomic CpG sequences.
2. The method of claim 1, wherein the genomic CpG sequences located within at least one gene sequence selected from the group consisting of *APC*, *ARF*, *CALCA*, *CDHI*, *CDKN2A*, *CDKN2B*, *ESR1*, *GSTP1*, *HIC1*, *MGMT*, *MLH1*, *MYOD1*, *RB1*, *TGFBR2*, *THBS1*, *TIMP3*, *CTNNB1*, *PTGS2* and *TYMS*, correspond to genomic CpG sequences of CpG islands.
 3. The method of claim 1, wherein the *APC*, *ARF*, *CALCA*, *CDHI*, *CDKN2A*, *CDKN2B*, *ESR1*, *GSTP1*, *HIC1*, *MGMT*, *MLH1*, *MYOD1*, *RB1*, *TGFBR2*, *THBS1*, *TIMP3*, *CTNNB1*, *PTGS2*, *TYMS* and *MTHFR* gene sequences are those defined by the specific oligonucleotide primers and probes corresponding to SEQ ID Nos:1-60, 64 and 65, as listed in TABLE II, or portions thereof.
 4. The method of claim 2 wherein the CpG islands are located within the promoter regions of one or more of the *APC*, *ARF*, *CALCA*, *CDHI*, *CDKN2A*, *CDKN2B*, *ESR1*, *GSTP1*, *HIC1*, *MGMT*, *MLH1*, *MYOD1*, *RB1*, *TGFBR2*, *THBS1*, *TIMP3*, *CTNNB1*, *PTGS2* and *TYMS* genes.
 5. The method of claim 2, wherein the *APC*, *ARF*, *CALCA*, *CDHI*, *CDKN2A*, *CDKN2B*, *ESR1*, *GSTP1*, *HIC1*, *MGMT*, *MLH1*, *MYOD1*, *RB1*, *TGFBR2*, *THBS1*, *TIMP3*, *CTNNB1*, *PTGS2*, and *TYMS* gene sequences correspond to any CpG island sequences associated with the sequences defined by the specific oligonucleotide primers and probes corresponding to SEQ ID NOs:1-54, 58-60, 64 and 65, as listed in TABLE II, or portions thereof, and wherein the associated CpG island sequences are those contiguous sequences of genomic DNA that encompass at least one nucleotide of the sequences defined by the specific oligonucleotide primers and probes corresponding to SEQ ID NOs:1-54, 58-60, 64 and 65, and satisfy the criteria of having both a frequency of CpG dinucleotides corresponding to an Observed/Expected Ratio >0.6, and a GC Content >0.5.

WO 01/75172

PCT/US01/10658

6. The method of claim 1, wherein the genomic CpG sequences are located within at least one gene sequence selected from the group consisting of *APC*, *CDKN2A*, *MYODI*, *CALCA*, *ESRI*, *MGMT* and *TIMP3*, and combinations thereof.
7. The method of claim 6, wherein the genomic CpG sequences located within at least one gene sequence selected from the group consisting of *APC*, *CDKN2A*, *MYODI*, *CALCA*, *ESRI*, *MGMT* and *TIMP3*, correspond to genomic CpG sequences of CpG islands.
8. The method of claim 6, wherein the *APC*, *CDKN2A*, *MYODI*, *CALCA*, *ESRI*, *MGMT* and *TIMP3* gene sequences are those defined by the specific oligonucleotide primers and probes corresponding to SEQ ID NOs:19-21, SEQ ID NOs:1-3, SEQ ID NOs:7-9, SEQ ID NOs:10-12, SEQ ID NOs:4-6, SEQ ID NOs:16-18 and SEQ ID NOs:13-15, respectively, as listed in TABLE II.
9. The method of claim 7 wherein the CpG islands are located within the promoter regions of one or more of the *APC*, *CDKN2A*, *MYODI*, *CALCA*, *ESRI*, *MGMT* and *TIMP3* genes.
10. The method of claim 7 wherein the *APC*, *CDKN2A*, *MYODI*, *CALCA*, *ESRI*, *MGMT* and *TIMP3* gene sequences correspond to any CpG island sequences associated with the sequences defined by the specific oligonucleotide primers and probes corresponding to SEQ ID NOs:19-21, SEQ ID NOs:1-3, SEQ ID NOs:7-9, SEQ ID NOs:10-12, SEQ ID NOs:4-6, SEQ ID NOs:16-18 and SEQ ID NOs:13-15, respectively, as listed in TABLE II, or portions thereof, and wherein the associated CpG island sequences are those contiguous sequences of genomic DNA that encompass at least one nucleotide of the sequences defined by the specific oligonucleotide primers and probes corresponding to SEQ ID NOs:19-21, SEQ ID NOs:1-3, SEQ ID NOs:7-9, SEQ ID NOs:10-12, SEQ ID NOs:4-6, SEQ ID NOs:16-18 and SEQ ID NOs:13-15, and satisfy the criteria of having both a frequency of CpG dinucleotides corresponding to an Observed/Expected Ratio >0.6, and a GC Content >0.5.
11. The method of claim 1, wherein the cancer or cancer-related condition is selected from the group consisting of gastrointestinal or esophageal adenocarcinoma, gastrointestinal or esophageal dysplasia, gastrointestinal or esophageal metaplasia, Barrett's intestinal tissue, pre-cancerous conditions in normal esophageal squamous mucosa, and combinations thereof.
12. The method of claim 11, wherein the cancer is esophageal adenocarcinoma, and wherein making a diagnostic or prognostic prediction of the cancer, based upon the methylation state of the genomic CpG sequences provides for classification of the adenocarcinoma by grade or stage.
13. The method of claim 6, wherein the cancer or cancer-related condition is selected from the group consisting of gastrointestinal or esophageal adenocarcinoma,

WO 01/75172

PCT/US01/10658

gastrointestinal or esophageal dysplasia, gastrointestinal or esophageal metaplasia, Barrett's intestinal tissue, pre-cancerous conditions in normal esophageal squamous mucosa, and combinations thereof.

14. The method of claim 13, wherein the cancer is esophageal adenocarcinoma, and wherein making a diagnostic or prognostic prediction of the cancer, based upon the methylation state of the genomic CpG sequences provides for classification of the adenocarcinoma by grade or stage.

15. The method of claim 1, wherein the methylation assay used to determine the methylation state of genomic CpG sequences is selected from the group consisting of MethylLight™, MS-SNuPE, MSP, COBRA, MCA, and DMH, and combinations thereof.

16. The method of claim 6, wherein the methylation assay used to determine the methylation state of genomic CpG sequences is selected from the group consisting of MethylLight™, MS-SNuPE, MSP, COBRA, MCA and DMH, and combinations thereof.

17. The method of claim 1, wherein the methylation assay used to determine the methylation state of genomic CpG sequences is based, at least in part, on an array or microarray comprising CpG-containing sequences located within at least one gene sequence selected from the group consisting of *APC*, *ARF*, *CALCA*, *CDHI*, *CDKN2A*, *CDKN2B*, *ESR1*, *GSTP1*, *HIC1*, *MGMT*, *MLH1*, *MYOD1*, *RBI*, *TGFBR2*, *THBS1*, *TIMP3*, *CTNNB1*, *PTGS2*, *TYMS* and *MTHFR*.

18. The method of claim 17, wherein the *APC*, *ARF*, *CALCA*, *CDHI*, *CDKN2A*, *CDKN2B*, *ESR1*, *GSTP1*, *HIC1*, *MGMT*, *MLH1*, *MYOD1*, *RBI*, *TGFBR2*, *THBS1*, *TIMP3*, *CTNNB1*, *PTGS2*, and *TYMS* gene sequences correspond to any CpG island sequences associated with the sequences defined by the specific oligonucleotide primers and probes corresponding to SEQ ID NOs:1-54, 58-60, 64 and 65, as listed in TABLE II, or portions thereof, and wherein the associated CpG island sequences are those contiguous sequences of genomic DNA that encompass at least one nucleotide of the sequences defined by the specific oligonucleotide primers and probes corresponding to SEQ ID NOs:1-54, 58-60, 64 and 65, and satisfy the criteria of having both a frequency of CpG dinucleotides corresponding to an Observed/Expected Ratio >0.6, and a GC Content >0.5.

19. The method of claim 17, wherein the *APC*, *ARF*, *CALCA*, *CDHI*, *CDKN2A*, *CDKN2B*, *ESR1*, *GSTP1*, *HIC1*, *MGMT*, *MLH1*, *MYOD1*, *RBI*, *TGFBR2*, *THBS1*, *TIMP3*, *CTNNB1*, *PTGS2*, *TYMS* and *MTHFR* gene sequences are those defined by, or correspond to the specific oligonucleotide primers and probes corresponding to SEQ ID NOs:1-60, 64 and 65, as listed in TABLE II, or portions thereof.

20. The method of claim 1 wherein the methylation state of genomic CpG sequences that is determined is that of hypermethylation, hypomethylation or normal

WO 01/75172

PCT/US01/10658

methylation.

21. A kit useful for diagnosis or prognosis of cancer or cancer-related conditions, comprising a carrier means containing one or more containers comprising:
- 5 (a) a container containing a probe or primer which hybridizes to any region of a sequence located within at least one gene sequence selected from the group consisting of *APC*, *ARF*, *CALCA*, *CDHI*, *CDKN2A*, *CDKN2B*, *ESR1*, *GSTP1*, *HIC1*, *MGMT*, *MLH1*, *MYOD1*, *RBI*, *TGFBR2*, *THBS1*, *TIMP3*, *CTNNB1*, *PTGS2*, *TYMS* and *MTHFR*; and
- (b) additional standard methylation assay reagents required to affect detection of methylated CpG-containing nucleic acid based, at least in part, on the probe or primer.
- 10 22. The kit of claim 21, wherein the additional standard methylation assay reagents are standard reagents for performing a methylation assay from the group consisting of MethyLight™, MS-SNuPE, MSP, COBRA, MCA and DMH, and combinations thereof.
23. The kit of claim 21, wherein the probe or primer comprises at least about 12 to 15 nucleotides of a sequence selected from the group consisting of SEQ ID NOs:1-60, 64 and 65, as listed in TABLE II.
- 15 24. A kit useful for diagnosis or prognosis of cancer or cancer-related conditions, comprising a carrier means containing one or more containers comprising:
- (a) an array or micorarray comprising sequences of at least about 12 to 15 nucleotides of a sequence selected from the group consisting of SEQ ID NOs:1-60, 64, 65, 20 and any sequence located within a CpG island sequence associated with SEQ ID NOs:1-54, 58-60, 64 and 65.

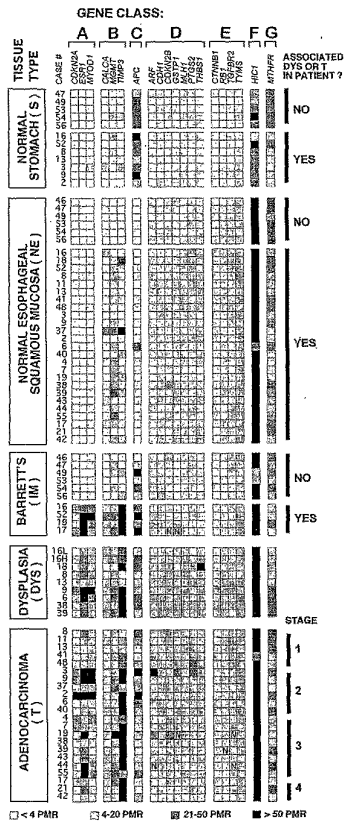


Fig. 1

WO 01/75172

PCT/US01/10658

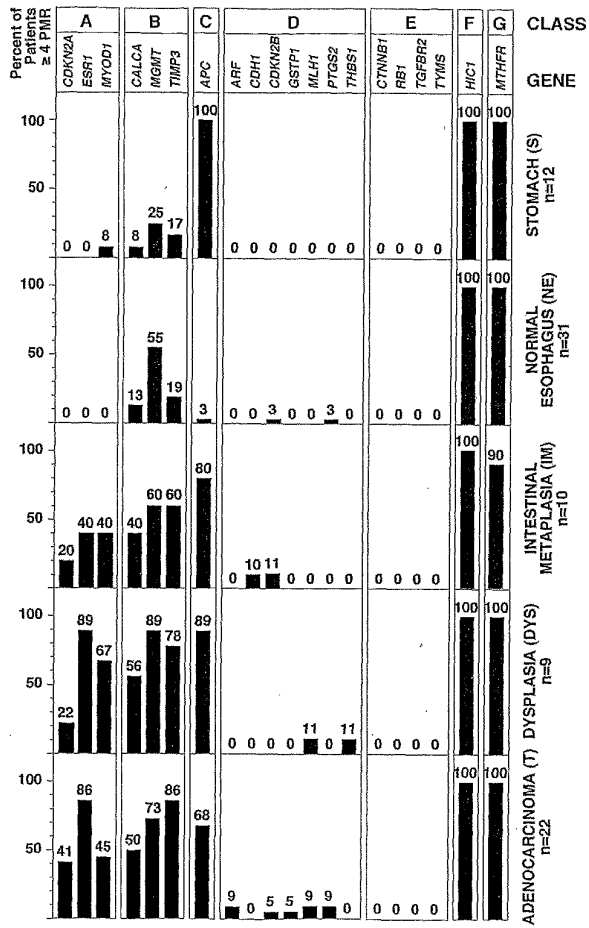
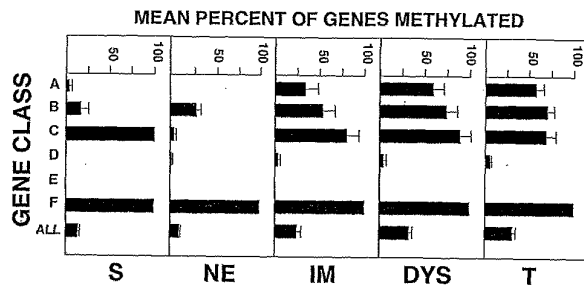


Fig. 2

WO 01/75172

PCT/US01/10658



TISSUE COMPARISON

p-values	NE → IM	IM → DYS	IM → T	DYS → T
A	0.0006	0.031	0.015	0.87
B	0.027	0.13	0.15	0.71
C	<0.0001	0.55	0.34	0.11
D	0.46	0.92	0.40	0.48
E	—	—	—	—
F	—	—	—	—
ALL	<0.0001	0.044	0.042	0.70

Fig. 3

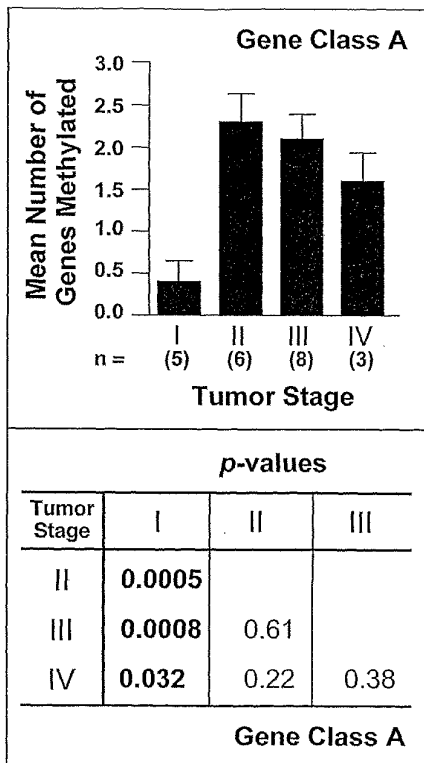


Fig. 4

WO 01/75172

PCT/US01/10658

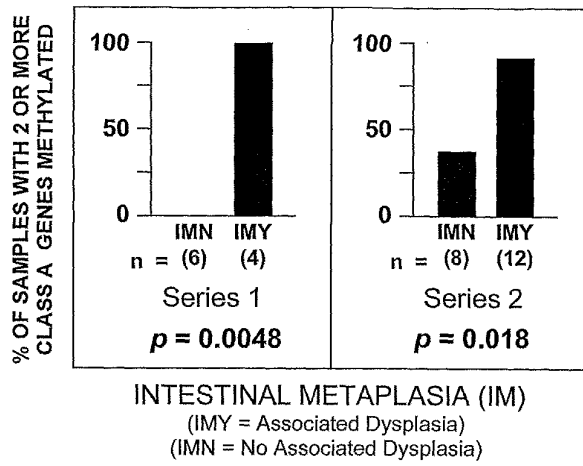


Fig. 5

WO 01/75172

PCT/US01/10658

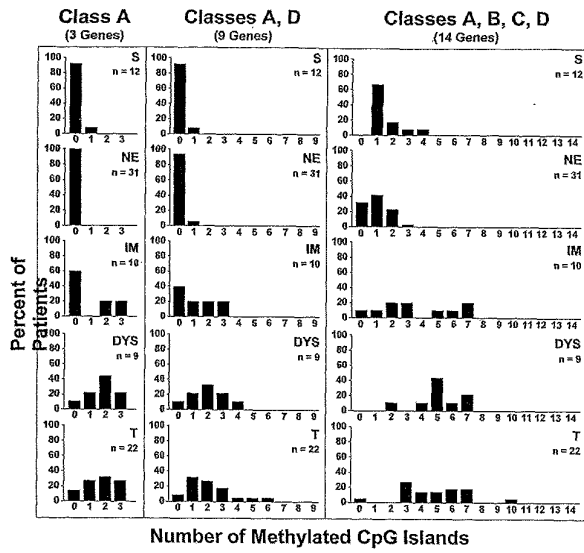


Fig. 6

WO 01/75172

PCT/US01/10658

SEQUENCE LISTING

<110> LAIRD, Peter
EADS, Cindy

<120> EPIGENETIC SEQUENCES FOR ESOPHAGEAL ADENOCARCINOMA

<130> 47675-12

<140> 60/193,839

<141> 2000-03-31

<160> 65

<170> PatentIn version 3.0

<210> 1

<211> 22

<212> DNA

<213> Homo sapiens

<400> 1

tggaatttc ggttgattgg tt

22

<210> 2

<211> 19

<212> DNA

<213> Homo sapiens

<400> 2

aacaacgtcc gcacctcct

19

<210> 3

<211> 18

<212> DNA

<213> Homo sapiens

WO 01/75172

PCT/US01/10658

<400> 3
accgacccc gaaccgcg

18

<210> 4
<211> 19
<212> DNA
<213> Homo sapiens

<400> 4
ggcggttcggtt ttgggattg

19

<210> 5
<211> 19
<212> DNA
<213> Homo sapiens

<400> 5
gccgacacgc gaactctaa

19

<210> 6
<211> 24
<212> DNA
<213> Homo sapiens

<400> 6
cgataaaacc gaacgacccg acga

24

<210> 7
<211> 18
<212> DNA
<213> Homo sapiens

WO 01/75172

PCT/US01/10658

<400> 7
gagcgcgcgt agttagcg

18

<210> 8
<211> 17
<212> DNA
<213> Homo sapiens

<400> 8
tccgacacgc cctttcc

17

<210> 9
<211> 30
<212> DNA
<213> Homo sapiens

<400> 9
ctccaacacc cgactactat atccgcgaaa

30

<210> 10
<211> 23
<212> DNA
<213> Homo sapiens

<400> 10
gttttggaag tatgagggtg acg

23

<210> 11
<211> 19
<212> DNA

WO 01/75172

PCT/US01/10658

<213> Homo sapiens

<400> 11
ttcccgccgc tataaatcg

19

<210> 12

<211> 30

<212> DNA

<213> Homo sapiens

<400> 12
attccgccaa tacacaacaa ccaataaacg

30

<210> 13

<211> 21

<212> DNA

<213> Homo sapiens

<400> 13
gcgtcggagg ttaaggttgt t

21

<210> 14

<211> 22

<212> DNA

<213> Homo sapiens

<400> 14
ctctcaaaa ttaccgtacg cg

22

<210> 15

<211> 19

WO 01/75172

PCT/US01/10658

<212> DNA
<213> Homo sapiens

<400> 15
aactcgctcg cccgccgaa

19

<210> 16
<211> 28
<212> DNA
<213> Homo sapiens

<400> 16
ctaactgtata acgaaaatcg taacaacc

28

<210> 17
<211> 25
<212> DNA
<213> Homo sapiens

<400> 17
agtatgaagg gtaggaagaa ttcgg

25

<210> 18
<211> 30
<212> DNA
<213> Homo sapiens

<400> 18
ccttacctct aaataccaac cccaaaccg

30

<210> 19

WO 01/75172

PCT/US01/10658

<211> 19
<212> DNA
<213> Homo sapiens

<400> 19
gaacccaaaac gctccccat

19

<210> 20
<211> 27
<212> DNA
<213> Homo sapiens

<400> 20
ttatatgtcg gttacgtgcg tttatat

27

<210> 21
<211> 22
<212> DNA
<213> Homo sapiens

<400> 21
cccgtcgaaa acccgccgat ta

22

<210> 22
<211> 19
<212> DNA
<213> Homo sapiens

<400> 22
acgggcggtt tcggtagtt

19

WO 01/75172

PCT/US01/10658

<210> 23
<211> 20
<212> DNA
<213> Homo sapiens

<400> 23
ccgaacctcc aaaatctcga

20

<210> 24
<211> 26
<212> DNA
<213> Homo sapiens

<400> 24
cgactctaaa ccctacgcac gcgaaa

26

<210> 25
<211> 26
<212> DNA
<213> Homo sapiens

<400> 25
aatttaggt tagagggtta tcgct

26

<210> 26
<211> 22
<212> DNA
<213> Homo sapiens

<400> 26
tccccaaaac gaaactaacg ac

22

WO 01/75172

PCT/US01/10658

<210> 27
<211> 19
<212> DNA
<213> Homo sapiens

<400> 27
cgcccacccg acctcgcat

19

<210> 28
<211> 20
<212> DNA
<213> Homo sapiens

<400> 28
aggaaggaga gagtgcgtcg

20

<210> 29
<211> 21
<212> DNA
<213> Homo sapiens

<400> 29
cgaataatcc accgttaacc g

21

<210> 30
<211> 29
<212> DNA
<213> Homo sapiens

<400> 30
ttaacgacac tcttccttc tttcccacg

29

WO 01/75172

PCT/US01/10658

<210> 31
<211> 23
<212> DNA
<213> Homo sapiens

<400> 31
gtcggcgtcg tgatttagta ttg
23

<210> 32
<211> 23
<212> DNA
<213> Homo sapiens

<400> 32
aaactacgac gacgaaactc caa
23

<210> 33
<211> 29
<212> DNA
<213> Homo sapiens

<400> 33
aaacctcgcg acctccgaac cttataaaa
29

<210> 34
<211> 18
<212> DNA
<213> Homo sapiens

<400> 34
ctatcgccgc ctcacgt

WO 01/75172

PCT/US01/10658

18

<210> 35
<211> 30
<212> DNA
<213> Homo sapiens

<400> 35
cgttatatat cgttcgtagt attcgtgttt

30

<210> 36
<211> 22
<212> DNA
<213> Homo sapiens

<400> 36
cgcgagctca aacgccacta cg

22

<210> 37
<211> 19
<212> DNA
<213> Homo sapiens

<400> 37
cggaagcggt cgggtaaag

19

<210> 38
<211> 18
<212> DNA
<213> Homo sapiens

<400> 38
aattccaccg ccccaaac

WO 01/75172

PCT/US01/10658

18

<210> 39
<211> 29
<212> DNA
<213> Homo sapiens

<400> 39
tttccgcaaa atatcttttc ttcttegeca

29

<210> 40
<211> 18
<212> DNA
<213> Homo sapiens

<400> 40
cgacgcacca acotaccg

18

<210> 41
<211> 25
<212> DNA
<213> Homo sapiens

<400> 41
gttttgagtt ggttttaoct tcggt

25

<210> 42
<211> 19
<212> DNA
<213> Homo sapiens

<400> 42

WO 01/75172

PCT/US01/10658

acgccgcgct cacctccct

19

<210> 43
<211> 17
<212> DNA
<213> Homo sapiens

<400> 43
ggaaaggcgc gtcgagt

17

<210> 44
<211> 18
<212> DNA
<213> Homo sapiens

<400> 44
tcccctatcc caaaccg

18

<210> 45
<211> 18
<212> DNA
<213> Homo sapiens

<400> 45
cgcggtttc ccgaaccg

18

<210> 46
<211> 22
<212> DNA
<213> Homo sapiens

WO 01/75172

PCT/US01/10658

<400> 46
ttagttcgcg tatcgattag cg

22

<210> 47
<211> 18
<212> DNA
<213> Homo sapiens

<400> 47
actaaacgcc gcgtccaa

18

<210> 48
<211> 21
<212> DNA
<213> Homo sapiens

<400> 48
tcacgtccgc gaaactcccg a

21

<210> 49
<211> 18
<212> DNA
<213> Homo sapiens

<400> 49
gcgcggagcg tagttagg

18

<210> 50
<211> 20
<212> DNA
<213> Homo sapiens

WO 01/75172

PCT/US01/10658

<400> 50
caaaccocgc tactcgtcat

20

<210> 51
<211> 21
<212> DNA
<213> Homo sapiens

<400> 51
caacgaacgac gccttccga a

21

<210> 52
<211> 19
<212> DNA
<213> Homo sapiens

<400> 52
cggcgtagg aaggacgat

19

<210> 53
<211> 24
<212> DNA
<213> Homo sapiens

<400> 53
tctcaacta taacgcgcct acat

24

<210> 54
<211> 29
<212> DNA

WO 01/75172

PCT/US01/10658

<213> Homo sapiens

<400> 54

ccgaataccg acaaaatacc gataccggt

29

<210> 55

<211> 29

<212> DNA

<213> Homo sapiens

<400> 55

tggtagtgag agttttaag atagttcga

29

<210> 56

<211> 18

<212> DNA

<213> Homo sapiens

<400> 56

agcctcatct tctcccga

18

<210> 57

<211> 27

<212> DNA

<213> Homo sapiens

<400> 57

tctcataccg ctcaaatcc aaaccg

27

<210> 58

<211> 19

WO 01/75172

PCT/US01/10658

<212> DNA
<213> Homo sapiens

<400> 58
gttaggcggt tagggcgtc

19

<210> 59
<211> 19
<212> DNA
<213> Homo sapiens

<400> 59
ccgaacgect ccacgtat

19

<210> 60
<211> 31
<212> DNA
<213> Homo sapiens

<400> 60
caacatcgtc taccacaacac actctcctac g

31

<210> 61
<211> 25
<212> DNA
<213> Homo sapiens

<400> 61
tggtgatgga ggaggtttag taagt

25

<210> 62

WO 01/75172

PCT/US01/10658

<211> 27
<212> DNA
<213> Homo sapiens

<400> 62
aaccaataaa acctactcct cccttaa

27

<210> 63
<211> 30
<212> DNA
<213> Homo sapiens

<400> 63
accaccaccc aacacacaat aacaaacaca

30

<210> 64
<211> 22
<212> DNA
<213> Homo sapiens

<400> 64
tggagtttcc ggttgattgg tt

22

<210> 65
<211> 19
<212> DNA
<213> Homo sapiens

<400> 65
aacaacgccc gcacctcct

19

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		Intern: application No. PCT/US01/10688																		
A. CLASSIFICATION OF SUBJECT MATTER IPC(7) : C12Q 1/68 US CL : 435/6 According to International Patent Classification (IPC) or to both national classification and IPC																				
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) U.S. : 435/6: 435/29.01, 24.25 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) Please See Extra Sheet.																				
C. DOCUMENTS CONSIDERED TO BE RELEVANT																				
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.																		
Y,P	KAWAKAMI, K et al. Hypermethylated APC DNA in plasma and prognosis of patients with esophageal adenocarcinoma. Journal of the National Cancer Institute. 15 November 2000. vol. 92., No. 22, pages 1805-1811, see entire document.	1-24																		
Y,P	EADS, C.A. et al. Fields of Aberrant CpG Island Hypermethylation in Barrett's Esophagus and associated Adenocarcinoma. Cancer Research. 15 September 2000. Vol. 50, pages 5021-5026, see entire document.	1-24																		
Y	SUZUKI, H. et al. Intragenic mutations of CDKN2B and CDKN2A in Primary human esophageal cancers. Human Molecular Genetics. 1995. Vol. 4. no. 10, pages 1883-1887, see entire document.	1-24																		
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.																				
<table border="0" style="width: 100%;"> <tr> <td style="width: 33%;">* Special categories of cited documents</td> <td style="width: 33%;">** Later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</td> <td style="width: 33%;"></td> </tr> <tr> <td>*A* Document defining the general state of the art which is not considered to be of particular relevance</td> <td>**X* Document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</td> <td></td> </tr> <tr> <td>**B* Earlier document published on or after the international filing date</td> <td>**Y* Document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</td> <td></td> </tr> <tr> <td>*L* Document which may throw doubt on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</td> <td>**Z* Document member of the same patent family</td> <td></td> </tr> <tr> <td>*O* Document referring to an oral disclosure, use, exhibition or other means</td> <td></td> <td></td> </tr> <tr> <td>**P* Document published prior to the international filing date but later than the priority date claimed</td> <td></td> <td></td> </tr> </table>			* Special categories of cited documents	** 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		*A* Document defining the general state of the art which is not considered to be of particular relevance	**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		**B* Earlier document published on or after the international filing date	**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		*L* Document which may throw doubt on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	**Z* Document member of the same patent family		*O* Document referring to an oral disclosure, use, exhibition or other means			**P* Document published prior to the international filing date but later than the priority date claimed		
* Special categories of cited documents	** 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																			
A Document defining the general state of the art which is not considered to be of particular relevance	**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																			
**B* Earlier document published on or after the international filing date	**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																			
L Document which may throw doubt on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	**Z* Document member of the same patent family																			
O Document referring to an oral disclosure, use, exhibition or other means																				
**P* Document published prior to the international filing date but later than the priority date claimed																				
Date of the actual completion of the international search 24 JULY 2001		Date of mailing of the international search report 24 AUG 2001																		
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box FCT Washington, D.C. 20231 Facsimile No. (703) 505-8320		Authorized officer LISA ARTHUR <i>Lisa Collins for</i> Telephone No. (703) 506-8986																		
Form PCT/ISA/210 (second sheet) (July 1998)*																				

INTERNATIONAL SEARCH REPORT		International application No. PCT/US01/10668
C (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CODY, D.T. et al. Differential DNA methylation of the p16 INK4A/CDKN2A promoter in human oral cancer cells and normal human oral keratinocytes. ORAL ONCOLOGY. 1999, Vol. 35, pages 516-522, see entire document.	1-10, 13, 15-24
Y	GRAFF, J.R. et al. Distinct patterns of E-Cadherin CpG island methylation in papillary, follicular, Hurtle's cell and poorly differentiated thyroid carcinoma. Cancer Research. 15 May 1998, Vol. 58, pages 2063-2066, see entire document.	1-10, 15-24
Y	IWASE, H. et al. DNA methylation analysis at distal and proximal promoter regions of the oestrogen receptor gene in breast cancers. British Journal of Cancer. 1999, Vol. 80, No. 12, pages 1982-1986, see entire document.	1-10, 15-24
Y	MILLAR, D. S. et al. Detailed methylation analysis of the glutathione S-transferase(GSTP1) gene in prostate cancer. ONCOGENE. 1999, Vol. 18, pages 1313-1324, see entire document.	1-10, 15-24
Y	JHAVERI, M.S. et al. Methylation-mediated regulation of the glutathione S-transferase Pi gene in human breast cancer cells. Gene (1998) Vol. 210, pages 1-7, see entire document.	1-10 and 15-24
Y	BARRETT, M.T. et al. Evolution of neoplastic cell lineages in Barrett oesophagus. Nature Genetics. May 1999, Vol. 22, pages 106-109, see entire document.	1-24
Y	HERMAN, J.G. et al. Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma. Proceedings of the National Academy of Sciences. June 1998, vol. 95, pages 6870-6875, see entire document.	1-11, 13, 15-24
Y	ESTELLER, M. et al. Inactivation of the DNA Repair gene O6-methylguanine-DNA methyltransferase by promoter hypermethylation is a common event in primary human neoplasia. Cancer Research. 15 February 1999, Vol. 59, pages 793-797, see entire document.	1-24
Y	IACOPETTA, B.J. et al. Hypermethylation of the Myf-3 gene in human colorectal cancer. Anticancer Research. 1997, Vol. 17, pages 429-432, see entire document.	1-11, 13, 15-24
Y	ESTELLER, M. et al. Hypermethylation-associated inactivation of p14ARF is independent of p16INK4A methylation and p53 mutational status. Cancer Research. 01 January 2000, Vol. 60, pages 128-133, see entire document.	1-24

INTERNATIONAL SEARCH REPORT		International application No. PCT/US01/10668
C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y,P	NAKAMURA, M. et al. Promoter hypermethylation of the RBI gene in glioblastomas. Laboratory Investigations January 2001. Vol. 81. no. 1, page 77-88, see entire document.	1-10, 15-24
Y	MEKI, J.R. et al. Cancer-specific region of hypermethylation identified within the HIC1 putative tumour suppressor gene in acute myeloid leukaemia. Leukemia. 1999. Vol. 13, pages 877-888, see entire document.	1-10, 15-24

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US01/10058

B. FIELDS SEARCHED

Electronic data bases consulted (Name of data base and where practicable terms used):

WEST, MEDLINE, BIOSIS, EMBASE, CAPLUS

search terms: hypermethylation, cpg islands, apc, arf/calca, cdh1, cdka2a, cdks2b, ecr1, gsp1, hie1, ngnt, mlh1, myd1,
rbl, tgrbr2, thbs1, timp2, cinnabim ptg22, tyms, mtbf

フロントページの続き

(81)指定国 AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT,BE,CH,CY,DE,DK,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GW,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BY,BZ,CA,CH,CN,CO,CR,CU,CZ,DE,DK,DM,DZ,EE,ES,FI,GB,GD,GE,GH,GM,HR,HU,ID,IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,MG,MK,MN,MW,MX,MZ,NO,NZ,PL,PT,RO,RU,SD,SE,SG,SI,SK,SL,TJ,TM,TR,TT,TZ,UA,UG,UZ,VN,YU,ZA,ZW

(74)代理人 100081330

弁理士 樋口 外治

(72)発明者 レアード, ピーター

アメリカ合衆国,カリフォルニア 91010,サウス パサデナ,フォレスト アベニュー 649

(72)発明者 イーズ, シンディー

アメリカ合衆国,カリフォルニア 91010,サウス パサデナ プロスペクト ストリート 635 #ビー15

Fターム(参考) 4B024 AA12 DA02 EA04 HA12

4B029 AA07 AA23 FA12

4B063 QA19 QQ43 QR32 QR55 QR62 QS34