

(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2006-525811

(P2006-525811A)

(43) 公表日 平成18年11月16日(2006.11.16)

| (51) Int. Cl. | F I | テーマコード (参考) |
|-------------------------------------|-----------------------|-------------|
| C 1 2 N 15/09 (2006.01) | C 1 2 N 15/00 Z N A A | 4 B O 2 4 |
| C 1 2 Q 1/68 (2006.01) | C 1 2 Q 1/68 A | 4 B O 6 3 |
| C 1 2 N 1/15 (2006.01) | C 1 2 N 1/15 | 4 B O 6 5 |
| C 1 2 N 1/19 (2006.01) | C 1 2 N 1/19 | 4 C O 8 4 |
| C 1 2 N 5/10 (2006.01) | C 1 2 N 5/00 A | 4 C O 8 6 |
| 審査請求 未請求 予備審査請求 未請求 (全 98 頁) 最終頁に続く | | |

| | | | |
|---------------|------------------------------|----------|---|
| (21) 出願番号 | 特願2006-514886 (P2006-514886) | (71) 出願人 | 505424125 ロゼッタ インフーマティクス エルエルシー |
| (86) (22) 出願日 | 平成16年5月17日 (2004.5.17) | | |
| (85) 翻訳文提出日 | 平成18年1月13日 (2006.1.13) | | |
| (86) 国際出願番号 | PCT/US2004/015439 | | アメリカ合衆国 98109 ワシントン州, シアトル, テリー アベニュー ノース 401 |
| (87) 国際公開番号 | W02005/018534 | (74) 代理人 | 100091096 弁理士 平木 祐輔 |
| (87) 国際公開日 | 平成17年3月3日 (2005.3.3) | (74) 代理人 | 100096183 弁理士 石井 貞次 |
| (31) 優先権主張番号 | 60/471, 392 | (74) 代理人 | 100118773 弁理士 藤田 節 |
| (32) 優先日 | 平成15年5月16日 (2003.5.16) | (74) 代理人 | 100111741 弁理士 田中 夏夫 |
| (33) 優先権主張国 | 米国 (US) | | |
| (31) 優先権主張番号 | 60/515, 223 | | |
| (32) 優先日 | 平成15年10月27日 (2003.10.27) | | |
| (33) 優先権主張国 | 米国 (US) | | |

最終頁に続く

(54) 【発明の名称】 RNA干渉の方法と組成物

(57) 【要約】

本発明は、RNA干渉による遺伝子サイレンシングのための方法と組成物を提供する。特に本発明また、その標的遺伝子と部分的配列相同性を有する小干渉性RNA (siRNA) を使用する、遺伝子サイレンシングまたはRNAノックダウンのための方法を提供する。本発明はまた、遺伝子を標的とする複数の異なるsiRNAに対する共通のおよび/または鑑別的応答を同定する方法を提供する。本発明はまた、siRNAの2つの鎖の相対活性を評価する方法を提供する。本発明はさらに、遺伝子サイレンシングのためのsiRNAを設計する方法を提供する。本発明はさらに、疾患の治療のための治療薬としてsiRNAを使用する方法を提供する。

。

【特許請求の範囲】

【請求項 1】

真核細胞を小干渉性RNA (siRNA) の分子に付すことを含むRNA干渉により該細胞中の標的遺伝子をサイレンシングさせる方法であって、該siRNAは、該標的遺伝子の転写体の配列と同一の、センス鎖またはアンチセンス鎖の少なくとも11ヌクレオチドの連続的ヌクレオチド配列を含むが、該siRNAは、該転写体中のどの配列とも完全長のセンス鎖またはアンチセンス鎖配列の同一性を持たず、該連続的ヌクレオチド配列は該siRNAの中央領域にある方法。

【請求項 2】

siRNAは、標的遺伝子の転写体の配列と同一の16ヌクレオチドより大きいセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まない、請求項 1 に記載の方法。 10

【請求項 3】

siRNAは、標的遺伝子の転写体の配列と同一の15ヌクレオチドより大きいセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まない、請求項 2 に記載の方法。

【請求項 4】

該siRNAは、該標的遺伝子の該転写体の配列と同一の14ヌクレオチドより大きいセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まない、請求項 2 に記載の方法。

【請求項 5】

siRNAは、標的遺伝子の転写体の配列と同一の13ヌクレオチドより大きいセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まない、請求項 2 に記載の方法。 20

【請求項 6】

siRNAは、標的遺伝子の転写体の配列と同一の12ヌクレオチドより大きいセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まない、請求項 2 に記載の方法。

【請求項 7】

真核細胞を小干渉性RNA (siRNA) の分子に付すことを含むRNA干渉により該細胞中の標的遺伝子をサイレンシングさせる方法であって、該siRNAは、該標的遺伝子の転写体の配列と同一の、センス鎖の少なくとも9ヌクレオチドの連続的ヌクレオチド配列を含むが、該siRNAは、該転写体中のどの配列とも完全長のセンス鎖またはセンス鎖配列の同一性を持たず、該連続的ヌクレオチド配列は該siRNA鎖の3'末端にある方法。

【請求項 8】

siRNAは、標的遺伝子の転写体の配列と同一の16ヌクレオチドより大きいセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まない、請求項 7 に記載の方法。 30

【請求項 9】

siRNAは、標的遺伝子の転写体の配列と同一の12ヌクレオチドより大きいセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まない、請求項 8 に記載の方法。

【請求項 10】

siRNAは、標的遺伝子の転写体の配列と同一の11ヌクレオチドより大きいセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まない、請求項 8 に記載の方法。

【請求項 11】

siRNAは、標的遺伝子の転写体の配列と同一の10ヌクレオチドより大きいセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まない、請求項 8 に記載の方法。 40

【請求項 12】

RNA干渉により真核細胞中の複数の異なる遺伝子をサイレンシングさせる方法であって、複数の異なる遺伝子のそれぞれの転写体配列は、該複数の異なる遺伝子に共通する9~18ヌクレオチドのヌクレオチド配列を含み、前記細胞を小干渉性RNA (siRNA) の分子に付すことを含み、前記siRNAは、(i)前記共通配列中の配列と同一の少なくとも11ヌクレオチドのセンス鎖またはアンチセンス鎖の中央の連続的ヌクレオチド配列および/または(ii)前記共通配列中の配列と同一の少なくとも9ヌクレオチドの3'センス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含む方法。

【請求項 13】

中央の連続的ヌクレオチド配列は11~15ヌクレオチドの長さである、請求項12に記載の方法。

【請求項14】

中央の連続的ヌクレオチド配列は14~15ヌクレオチドの長さである、請求項13に記載の方法。

【請求項15】

中央の連続的ヌクレオチド配列は13ヌクレオチドの長さである、請求項13に記載の方法。

【請求項16】

中央の連続的ヌクレオチド配列は12ヌクレオチドの長さである、請求項13に記載の方法。 10

【請求項17】

中央の連続的ヌクレオチド配列は11ヌクレオチドの長さである、請求項13に記載の方法。

【請求項18】

RNA干渉により真核細胞中の第1遺伝子をサイレンシングするが第2遺伝子はサイレンシングしない方法であって、(i)第1遺伝子の転写体配列と同一の少なくとも11ヌクレオチドのセンス鎖またはアンチセンス鎖の中央の連続的ヌクレオチド配列、または(ii)該第1遺伝子の転写体配列と同一の少なくとも9ヌクレオチドの3'センス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含むsiRNA分子に該真核細胞を付すことを含み、該siRNAは、前記第2遺伝子の前記転写体の配列と同一の10ヌクレオチドを超えるセンス鎖またはアンチセンス鎖の中央の連続的ヌクレオチド配列を含まず、かつ前記siRNAは、前記第2遺伝子の前記転写体の配列と同一の8ヌクレオチドを超える3'センス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まない方法。 20

【請求項19】

siRNA分子は、第2遺伝子の転写体の配列と同一の8ヌクレオチドの長さを超える連続的ヌクレオチド配列を含まない、請求項18に記載の方法。

【請求項20】

(i)の中央の連続的ヌクレオチド配列は11~15ヌクレオチドの長さである、請求項18に記載の方法。 30

【請求項21】

(i)の中央の連続的ヌクレオチド配列は14~15ヌクレオチドの長さである、請求項20に記載の方法。

【請求項22】

(i)の中央の連続的ヌクレオチド配列は13ヌクレオチドの長さである、請求項20に記載の方法。

【請求項23】

(i)の中央の連続的ヌクレオチド配列は12ヌクレオチドの長さである、請求項20に記載の方法。

【請求項24】

(i)の中央の連続的ヌクレオチド配列は11ヌクレオチドの長さである、請求項20に記載の方法。 40

【請求項25】

(ii)の連続的ヌクレオチド配列は9~15ヌクレオチドの長さである、請求項18に記載の方法。

【請求項26】

(ii)の連続的ヌクレオチド配列は9~12ヌクレオチドの長さである、請求項25に記載の方法。

【請求項27】

(ii)の該連続的ヌクレオチド配列は10ヌクレオチドの長さである、請求項26に記載の 50

方法。

【請求項 2 8】

(ii)の連続的ヌクレオチド配列は9ヌクレオチドの長さである、請求項 2 6 に記載の方法。

【請求項 2 9】

RNA干渉により真核細胞中の第1遺伝子をサイレンシングするが第2遺伝子はサイレンシングしないための小干渉性RNAを設計する方法であって、(i)該第1遺伝子の転写体配列と同一の少なくとも11ヌクレオチドのセンス鎖またはアンチセンス鎖の中央の連続的ヌクレオチド配列、または(ii)該第1遺伝子の転写体配列と同一の少なくとも9ヌクレオチドの3'センス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含むsiRNA分子を同定することを含み、該siRNAは、前記第2遺伝子の前記転写体の配列と同一の10ヌクレオチドを超えるセンス鎖またはアンチセンス鎖の中央の連続的ヌクレオチド配列を含まず、かつ前記siRNAは、前記第2遺伝子の前記転写体の配列と同一の8ヌクレオチドを超える3'センス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まない方法。

10

【請求項 3 0】

siRNA分子は、第2遺伝子の転写体の配列と同一の8ヌクレオチドの長さを超えるセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まない、請求項 2 9 に記載の方法。

【請求項 3 1】

(i)の中央の連続的ヌクレオチド配列は11~15ヌクレオチドの長さである、請求項 2 9 に記載の方法。

20

【請求項 3 2】

(i)の中央の連続的ヌクレオチド配列は14~15ヌクレオチドの長さである、請求項 3 1 に記載の方法。

【請求項 3 3】

(i)の中央の連続的ヌクレオチド配列は13ヌクレオチドの長さである、請求項 3 1 に記載の方法。

【請求項 3 4】

(i)の中央の連続的ヌクレオチド配列は12ヌクレオチドの長さである、請求項 3 1 に記載の方法。

30

【請求項 3 5】

(i)の中央の連続的ヌクレオチド配列は11ヌクレオチドの長さである、請求項 3 1 に記載の方法。

【請求項 3 6】

(ii)の連続的ヌクレオチド配列は9~15ヌクレオチドの長さである、請求項 2 9 に記載の方法。

【請求項 3 7】

(ii)の連続的ヌクレオチド配列は9~12ヌクレオチドの長さである、請求項 3 6 に記載の方法。

【請求項 3 8】

(ii)の連続的ヌクレオチド配列は10ヌクレオチドの長さである、請求項 3 7 に記載の方法。

40

【請求項 3 9】

(ii)の連続的ヌクレオチド配列は9ヌクレオチドの長さである、請求項 3 7 に記載の方法。

【請求項 4 0】

真核細胞をsiRNAに付した後の選択した時点で該真核細胞の発現プロファイルを測定することを含む、該真核細胞に対するsiRNAの作用を測定する方法であって、各発現プロファイルは複数の異なる遺伝子の測定された転写体レベルを含む方法。

【請求項 4 1】

50

- 複数の異なる遺伝子は5個の異なる遺伝子を含む、請求項40に記載の方法。
- 【請求項42】
 複数の異なる遺伝子は10個の異なる遺伝子を含む、請求項41に記載の方法。
- 【請求項43】
 複数の異なる遺伝子は100個の異なる遺伝子を含む、請求項42に記載の方法。
- 【請求項44】
 複数の異なる遺伝子は1,000個の異なる遺伝子を含む、請求項43に記載の方法。
- 【請求項45】
 該複数の異なる遺伝子は10,000個の異なる遺伝子を含む、請求項44に記載の方法。
- 【請求項46】 10
 複数の異なる遺伝子は25,000個の異なる遺伝子を含む、請求項45に記載の方法。
- 【請求項47】
 真核細胞に対するsiRNAの作用を測定する方法であって、
 (a)真核細胞をsiRNAに付した後の複数の異なる各時点で該真核細胞の発現プロファイル
 を測定し、ここで各発現プロファイルは複数の異なる遺伝子の測定された転写体レベルを
 含み；そして
 (b)該測定された転写体レベルの動力学的挙動に基づいて、該複数の遺伝子を異なる動
 力学群に分類する、ことを含む方法。
- 【請求項48】
 複数の異なる遺伝子は5個の異なる遺伝子を含む、請求項47に記載の方法。 20
- 【請求項49】
 複数の異なる遺伝子は10個の異なる遺伝子を含む、請求項48に記載の方法。
- 【請求項50】
 複数の異なる遺伝子は100個の異なる遺伝子を含む、請求項49に記載の方法。
- 【請求項51】
 複数の異なる遺伝子は1,000個の異なる遺伝子を含む、請求項50に記載の方法。
- 【請求項52】
 複数の異なる遺伝子は10,000個の異なる遺伝子を含む、請求項51に記載の方法。
- 【請求項53】
 複数の異なる遺伝子は25,000個の異なる遺伝子を含む、請求項52に記載の方法。 30
- 【請求項54】
 真核細胞中の標的遺伝子をサイレンシングするように設計される小干渉性RNA (siRNA)
 によりサイレンシングされる該真核細胞中の1つ以上の遺伝子を同定する方法であって、
 (a)前記真核細胞を前記siRNAに付した後の複数の異なる各時点で前記真核細胞の発現プ
 ロファイル測定し、各発現プロファイルは複数の異なる遺伝子の測定された転写体レベ
 ルを含み；そして
 (b)その転写体レベルが、標的遺伝子にコードされるタンパク質レベルの低下より実質
 的に速く低下する、前記複数の異なる遺伝子の中の1つ以上の遺伝子を同定する、こと
 を含む方法。
- 【請求項55】 40
 標的遺伝子によりコードされるタンパク質のレベルがその摂動されないレベルの約50%
 まで低下する前に、1つ以上の遺伝子の転写体レベルは少なくとも50%低下する、請求項
 54に記載の方法。
- 【請求項56】
 タンパク質レベルの低下は、複数の異なる各時点で該タンパク質の量を測定すること
 を含む方法により測定される、請求項54に記載の方法。
- 【請求項57】
 複数の異なる遺伝子は5個の異なる遺伝子を含む、請求項54に記載の方法。
- 【請求項58】
 複数の異なる遺伝子は10個の異なる遺伝子を含む、請求項57に記載の方法。 50

- 【請求項 59】
複数の異なる遺伝子は100個の異なる遺伝子を含む、請求項 58 に記載の方法。
- 【請求項 60】
複数の異なる遺伝子は1,000個の異なる遺伝子を含む、請求項 59 に記載の方法。
- 【請求項 61】
複数の異なる遺伝子は10,000個の異なる遺伝子を含む、請求項 60 に記載の方法。
- 【請求項 62】
複数の異なる遺伝子は25,000個の異なる遺伝子を含む、請求項 61 に記載の方法。
- 【請求項 63】
複数の異なる小干渉性RNA分子 (siRNA) により直接サイレンシングされる真核細胞中の 1 つ以上の遺伝子を同定する方法であって、各 siRNA は、真核細胞中の同じ標的遺伝子をサイレンシングするように設計され、該方法は、
(a) 複数の異なる各 siRNA について前記真核細胞の発現プロファイルを測定し、ここで各発現プロファイルは複数の異なる遺伝子の測定された転写体レベルを含み；そして
(b) その転写体レベルが前記複数の siRNA により一般的に影響を受ける、前記複数の異なる遺伝子の中の 1 つ以上の遺伝子を同定する、ことを含む方法。
- 【請求項 64】
工程 (b) は、異なる siRNA の発現プロファイルの中の複数の遺伝子を集めることを含む方法により行われる、請求項 63 に記載の方法。
- 【請求項 65】
各発現プロファイルは、該細胞を各 siRNA に付した後に同時に測定される、請求項 64 に記載の方法。
- 【請求項 66】
測定時点は、標的遺伝子によりコードされるタンパク質レベルのその摂動されないレベルの約 50% 低下の時間スケールより実質的に短い、請求項 65 に記載の方法。
- 【請求項 67】
複数の異なる遺伝子は5個の異なる遺伝子を含む、請求項 63 に記載の方法。
- 【請求項 68】
複数の異なる遺伝子は10個の異なる遺伝子を含む、請求項 67 に記載の方法。
- 【請求項 69】
複数の異なる遺伝子は100個の異なる遺伝子を含む、請求項 68 に記載の方法。
- 【請求項 70】
複数の異なる遺伝子は1,000個の異なる遺伝子を含む、請求項 69 に記載の方法。
- 【請求項 71】
複数の異なる遺伝子は10,000個の異なる遺伝子を含む、請求項 70 に記載の方法。
- 【請求項 72】
複数の異なる遺伝子は25,000個の異なる遺伝子を含む、請求項 71 に記載の方法。
- 【請求項 73】
真核細胞中の 1 つ以上の候補遺伝子を同定する方法であって、1 つ以上の候補遺伝子の発現レベルの変化が真核細胞中の表現型の特徴になり、
(a) 真核細胞への小干渉性RNA (siRNA) の導入に関連する真核細胞の表現型特徴を同定し、ここで該 siRNA は、真核細胞中の標的遺伝子をサイレンシングするように設計され、
；
(b) siRNA の導入後の複数の異なる各時点で真核細胞の発現プロファイルを測定し、ここで各発現プロファイルは、複数の異なる遺伝子の測定された転写体レベルを含み；
(c) その転写体レベルが、該標的遺伝子によりコードされるタンパク質のレベルの低下より実質的に速く低下する、該複数の遺伝子から 1 つ以上の遺伝子を同定し；そして
(d) 該 1 つ以上の遺伝子を 1 つ以上の候補遺伝子として同定する、ことを含む方法。
- 【請求項 74】
標的遺伝子によりコードされるタンパク質のレベルがその摂動されないレベルの約 50%

10

20

30

40

50

まで低下する前に、1つ以上の遺伝子の転写体レベルは少なくとも50%低下する、請求項73に記載の方法。

【請求項75】

タンパク質レベルの低下は、複数の異なる各時点で該タンパク質の量を測定することを含む、請求項73に記載の方法。

【請求項76】

複数の異なる遺伝子は5個の異なる遺伝子を含む、請求項73に記載の方法。

【請求項77】

複数の異なる遺伝子は10個の異なる遺伝子を含む、請求項76に記載の方法。

【請求項78】

複数の異なる遺伝子は100個の異なる遺伝子を含む、請求項77に記載の方法。

10

【請求項79】

複数の異なる遺伝子は1,000個の異なる遺伝子を含む、請求項78に記載の方法。

【請求項80】

複数の異なる遺伝子は10,000個の異なる遺伝子を含む、請求項79に記載の方法。

【請求項81】

複数の異なる遺伝子は25,000個の異なる遺伝子を含む、請求項80に記載の方法。

【請求項82】

1つ以上の候補遺伝子の発現レベルの変化が真核細胞中の表現型の特徴になる、真核細胞中の1つ以上の候補遺伝子を同定する方法であって、

20

(a)複数の異なる各小干渉性RNA分子(sirRNA)のそれぞれの導入に通常関連する真核細胞の表現型特徴を同定し、ここで該sirRNAは、真核細胞中の標的遺伝子をサイレンシングするように設計され;

(b)複数の異なる各小干渉性RNA分子(sirRNA)のそれぞれの発現プロファイルを測定し、ここで各sirRNAは前記真核細胞中の同じ標的遺伝子をサイレンシングするように設計され、かつ各発現プロファイルは、複数の異なる遺伝子の測定された転写体レベルを含み;

(c)その転写体レベルが、前記複数のsirRNAにより通常影響を受ける、前記複数の遺伝子から1つ以上の遺伝子を同定し;そして

(d)前記1つ以上の遺伝子を1つ以上の候補遺伝子として同定する、ことを含む方法。

【請求項83】

工程(c)は、異なるsirRNAの発現プロファイルの中に複数の遺伝子を集めることを含む方法により行われる、請求項82に記載の方法。

30

【請求項84】

発現プロファイルは、各sirRNAの導入後に同じ時点で測定される、請求項82に記載の方法。

【請求項85】

測定時点は、標的遺伝子によりコードされるタンパク質レベルのその摂動されないレベルの約50%低下の時間スケールより実質的に短い、請求項84に記載の方法。

【請求項86】

複数の異なる遺伝子は5個の異なる遺伝子を含む、請求項82に記載の方法。

40

【請求項87】

複数の異なる遺伝子は10個の異なる遺伝子を含む、請求項86に記載の方法。

【請求項88】

複数の異なる遺伝子は100個の異なる遺伝子を含む、請求項87に記載の方法。

【請求項89】

複数の異なる遺伝子は1,000個の異なる遺伝子を含む、請求項88に記載の方法。

【請求項90】

複数の異なる遺伝子は10,000個の異なる遺伝子を含む、請求項89に記載の方法。

【請求項91】

複数の異なる遺伝子は25,000個の異なる遺伝子を含む、請求項90に記載の方法。

50

【請求項 9 2】

真核細胞中で第1タンパク質により制御されるが第2タンパク質には制御されない1つ以上の遺伝子を同定する方法であって、

(a)(i)該第1タンパク質をコードする第1遺伝子の転写体配列と同一の少なくとも11ヌクレオチドのセンス鎖またはアンチセンス鎖の中央の連続的ヌクレオチド配列、または(ii)該第1遺伝子の転写体配列と同一の少なくとも9ヌクレオチドの3'センス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含む、siRNAの分子に前記真核細胞を付し、ここで該siRNAは、第2タンパク質をコードする第2遺伝子の転写体配列と同一の10ヌクレオチドを超えるセンス鎖またはアンチセンス鎖の中央の連続的ヌクレオチド配列を含まず、かつsiRNAは、第2遺伝子の転写体配列と同一の8ヌクレオチドを超える3'センス鎖またはアンチセンス鎖の連続的ヌクレオチド配列とを含まず；

10

(b)前記真核細胞を前記siRNAに付した後のある期間後に、真核細胞の発現プロファイルを測定し、ここで各発現プロファイルは、複数の異なる遺伝子の測定された転写体レベルを含み、前記期間は、siRNAに付されていない細胞中の第1タンパク質のレベルと比較して、第1タンパク質のレベルの50%低下が観察される時間より長く；そして

(c)第1タンパク質によりサイレンシングされるが第2タンパク質にはサイレンシングされない遺伝子としてのsiRNAに付されていない細胞中の転写体レベルから、その転写体レベルが変化している1つ以上の遺伝子を同定する、ことを含む方法。

【請求項 9 3】

siRNA分子は、第2遺伝子の任意の配列と同一の、8ヌクレオチドを超える長さのセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まない、請求項 9 2 に記載の方法

20

【請求項 9 4】

中央の連続的ヌクレオチド配列は11~15ヌクレオチドの長さである、請求項 9 2 に記載の方法。

【請求項 9 5】

中央の連続的ヌクレオチド配列は14~15ヌクレオチドの長さである、請求項 9 4 に記載の方法。

【請求項 9 6】

中央の連続的ヌクレオチド配列は13ヌクレオチドの長さである、請求項 9 4 に記載の方法。

30

【請求項 9 7】

中央の連続的ヌクレオチド配列は12ヌクレオチドの長さである、請求項 9 4 に記載の方法。

【請求項 9 8】

中央の連続的ヌクレオチド配列は11ヌクレオチドの長さである、請求項 9 4 に記載の方法。

【請求項 9 9】

(i)の連続的ヌクレオチド配列は9~15ヌクレオチドの長さである、請求項 9 2 に記載の方法。

40

【請求項 1 0 0】

(ii)の連続的ヌクレオチド配列は9~12ヌクレオチドの長さである、請求項 9 9 に記載の方法。

【請求項 1 0 1】

(i)の連続的ヌクレオチド配列は10ヌクレオチドの長さである、請求項 1 0 0 に記載の方法。

【請求項 1 0 2】

(ii)の連続的ヌクレオチド配列は9ヌクレオチドの長さである、請求項 1 0 0 に記載の方法。

【請求項 1 0 3】

50

複数の異なる遺伝子は100個の異なる遺伝子を含む、請求項 9 2、9 5、または 1 0 2 に記載の方法。

【請求項 1 0 4】

複数の異なる遺伝子は1,000個の異なる遺伝子を含む、請求項 1 0 3 に記載の方法。

【請求項 1 0 5】

複数の異なる遺伝子は10,000個の異なる遺伝子を含む、請求項 1 0 4 に記載の方法。

【請求項 1 0 6】

複数の異なる遺伝子は25,000個の異なる遺伝子を含む、請求項 1 0 5 に記載の方法。

【請求項 1 0 7】

小干渉性RNA (siRNA) を含む真核細胞であって、該 siRNA は、該真核細胞中の遺伝子の転写体配列と同一の少なくとも11ヌクレオチドのセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含むが、該真核細胞のゲノム中の他の遺伝子の転写体の任意の配列と完全長のセンス鎖またはアンチセンス鎖の配列同一性は無く、前記連続的ヌクレオチド配列は前記 siRNA 分子の中央領域にある真核細胞。

10

【請求項 1 0 8】

siRNA は、標的遺伝子の転写体配列と同一の16ヌクレオチドより大きいセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まない、請求項 1 0 7 に記載の真核細胞。

【請求項 1 0 9】

siRNA は、標的遺伝子の転写体配列と同一の15ヌクレオチドより大きいセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まない、請求項 1 0 8 に記載の真核細胞。

20

【請求項 1 1 0】

siRNA は、標的遺伝子の転写体配列と同一の14ヌクレオチドより大きいセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まない、請求項 1 0 8 に記載の真核細胞。

【請求項 1 1 1】

siRNA は、標的遺伝子の転写体配列と同一の13ヌクレオチドより大きいセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まない、請求項 1 0 8 に記載の真核細胞。

【請求項 1 1 2】

siRNA は、標的遺伝子の転写体配列と同一の12ヌクレオチドより大きいセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まない、請求項 1 0 8 に記載の真核細胞。

【請求項 1 1 3】

小干渉性RNA (siRNA) を含む真核細胞であって、該 siRNA は、該真核細胞中の遺伝子の転写体配列と同一の少なくとも11ヌクレオチドのセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含むが、前記真核細胞のゲノム中の他の遺伝子の転写体の任意の配列と完全長のセンス鎖またはアンチセンス鎖の配列同一性は無く、前記連続的ヌクレオチド配列は前記 siRNA 分子の3'末端にある真核細胞。

30

【請求項 1 1 4】

siRNA は、標的遺伝子の転写体配列と同一の15ヌクレオチドより大きいセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まない、請求項 1 1 3 に記載の真核細胞。

【請求項 1 1 5】

siRNA は、標的遺伝子の転写体配列と同一の12ヌクレオチドより大きいセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まない、請求項 1 1 4 に記載の真核細胞。

40

【請求項 1 1 6】

siRNA は、標的遺伝子の転写体配列と同一の11ヌクレオチドより大きいセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まない、請求項 1 1 4 に記載の真核細胞。

【請求項 1 1 7】

siRNA は、標的遺伝子の転写体配列と同一の9ヌクレオチドより大きいセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まない、請求項 1 1 4 に記載の真核細胞。

【請求項 1 1 8】

真核細胞はヒト細胞である、請求項 1、7、1 2、1 8、2 9、4 0、4 7、5 4、6 3、7 3、8 2 および 9 2 のいずれか 1 項に記載の方法。

50

【請求項 1 1 9】

真核細胞はヒト細胞である、請求項 1 0 7 または 1 1 3 に記載の真核細胞。

【請求項 1 2 0】

真核細胞中の標的遺伝子をサイレンシングするように設計された第1のsiRNAと第2のsiRNAにより鑑別的にサイレンシングされる遺伝子を同定する方法であって、

(a)前記真核細胞の第1の発現プロファイルと前記真核細胞の第2の発現プロファイルとを比較し、ここで第1の発現プロファイルは、前記細胞への第1siRNAの導入後の選択された時点で測定され、第2の発現プロファイルは前記細胞への第2siRNAの導入後の前記の選択された時点で測定され、前記発現プロファイルは複数の異なる遺伝子の測定された転写体レベルを含み；そして

10

(b)転写体レベルが該第1および第2siRNAにより鑑別的に影響を受ける、前記標的遺伝子以外の遺伝子を同定する、ことを含む方法。

【請求項 1 2 1】

選択された時点は、標的遺伝子によりコードされるタンパク質レベルのその摂動されないレベルの50%低下の時間スケールより実質的に短く、同定される遺伝子は、第1および第2siRNAにより直接鑑別的にサイレンシングされる、請求項 1 2 0 に記載の方法。

【請求項 1 2 2】

選択された時点は、標的遺伝子によりコードされるタンパク質レベルのその摂動されないレベルからの前記標的遺伝子によりコードされるタンパク質のレベルの50%低下の時間スケールより実質的に長く、同定される遺伝子は、直接サイレンシングされた遺伝子の鑑別的サイレンシングの結果として、第1および第2siRNAにより鑑別的にサイレンシングされる、請求項 1 2 0 に記載の方法。

20

【請求項 1 2 3】

複数の異なる遺伝子は5個の異なる遺伝子を含む、請求項 1 2 0 ~ 1 2 2 のいずれか 1 項に記載の方法。

【請求項 1 2 4】

複数の異なる遺伝子は10個の異なる遺伝子を含む、請求項 1 2 3 に記載の方法。

【請求項 1 2 5】

複数の異なる遺伝子は100個の異なる遺伝子を含む、請求項 1 2 4 に記載の方法。

【請求項 1 2 6】

複数の異なる遺伝子は1,000個の異なる遺伝子を含む、請求項 1 2 5 に記載の方法。

30

【請求項 1 2 7】

複数の異なる遺伝子は10,000個の異なる遺伝子を含む、請求項 1 2 6 に記載の方法。

【請求項 1 2 8】

複数の異なる遺伝子は25,000個の異なる遺伝子を含む、請求項 1 2 7 に記載の方法。

【請求項 1 2 9】

哺乳動物に治療的に十分な量のsiRNAを投与することを含む哺乳動物の疾患または好ましくない症状の治療方法であって、該siRNAは、その発現が前記疾患または好ましくない症状を引き起こす遺伝子を標的とし、前記siRNAは、(i)前記遺伝子の転写体中の配列と同一の少なくとも11ヌクレオチドであり18ヌクレオチド以下のセンス鎖またはアンチセンス鎖の中央の連続的ヌクレオチド配列および/または(ii)前記転写体中の配列と同一の少なくとも9ヌクレオチドであり18ヌクレオチド以下の3'センス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含む治療方法。

40

【請求項 1 3 0】

siRNAは、遺伝子の転写体配列と同一の16ヌクレオチドより大きいセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まない、請求項 1 2 9 に記載の方法。

【請求項 1 3 1】

siRNAは、遺伝子の転写体配列と同一の15ヌクレオチドより大きいセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まない、請求項 1 3 0 に記載の方法。

【請求項 1 3 2】

50

siRNAは、遺伝子の転写体配列と同一の14ヌクレオチドより大きいセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まない、請求項130に記載の方法。

【請求項133】

siRNAは、遺伝子の転写体配列と同一の13ヌクレオチドより大きいセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まない、請求項130に記載の方法。

【請求項134】

siRNAは、遺伝子の転写体配列と同一の12ヌクレオチドより大きいセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まない、請求項130に記載の方法。

【請求項135】

哺乳動物はヒトである、請求項130～134のいずれか1項に記載の方法。

10

【請求項136】

真核細胞中の遺伝子をサイレンシングするために複数の異なるsiRNAから1つ以上の小干渉性RNA (siRNA) を選択する方法であって、各異なるsiRNAは遺伝子の転写体中の異なる配列を標的とするように設計され、

(a)各siRNAについて応答プロファイルを測定し、ここで該応答プロファイルは、複数の遺伝子の発現レベルの測定値を含み；そして

(b)その応答プロファイルに基づいて1つ以上のsiRNAを選択する、ことを含む方法。

【請求項137】

複数の異なるsiRNAは、その配列が標的遺伝子の一部のまたは全てのコード配列をうめるsiRNAを含む、請求項136に記載の方法。

20

【請求項138】

複数の異なるsiRNAは、その配列が10塩基の間隔で標的遺伝子の一部のまたは全てのコード配列をうめるsiRNAを含む、請求項136に記載の方法。

【請求項139】

複数の異なるsiRNAは、その配列が5塩基の間隔で標的遺伝子の一部のまたは全てのコード配列をうめるsiRNAを含む、請求項136に記載の方法。

【請求項140】

複数の異なるsiRNAは、その配列が1塩基の間隔で標的遺伝子の一部のまたは全てのコード配列をうめるsiRNAを含む、請求項136に記載の方法。

【請求項141】

1つ以上のsiRNAは、各siRNAの応答プロファイルを所望の応答プロファイルと比較し、その応答プロファイルが所望の応答プロファイルに一致する1つ以上のsiRNAを選択することにより選択される、請求項136～140のいずれか1項に記載の方法。

30

【請求項142】

真核細胞中の複数の異なる遺伝子を標的とする小干渉性RNA (siRNA) を設計する方法であって、複数の異なる遺伝子は9～18ヌクレオチドの共通のヌクレオチド配列を共有し、(i)前記共通配列中の配列と同一の少なくとも11ヌクレオチドのセンス鎖またはアンチセンス鎖の中央の連続的ヌクレオチド配列および/または(ii)該共通配列中の配列と同一の少なくとも9ヌクレオチドの3'センス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含む、siRNAを選択することを含む方法。

40

【請求項143】

中央の連続的ヌクレオチド配列は11～15ヌクレオチドの長さである、請求項142に記載の方法。

【請求項144】

中央の連続的ヌクレオチド配列は14～15ヌクレオチドの長さである、請求項143に記載の方法。

【請求項145】

中央の連続的ヌクレオチド配列は13ヌクレオチドの長さである、請求項143に記載の方法。

【請求項146】

50

中央の連続的ヌクレオチド配列は12ヌクレオチドの長さである、請求項143に記載の方法。

【請求項147】

中央の連続的ヌクレオチド配列は11ヌクレオチドの長さである、請求項143に記載の方法。

【請求項148】

標的遺伝子をサイレンシングするように設計された複数の異なるsiRNAを該真核細胞に導入することを含む、RNA干渉により真核細胞中の標的遺伝子をサイレンシングする方法。

【請求項149】

複数の異なるsiRNAは少なくとも3個の異なるsiRNAからなる、請求項148に記載の方法。

【請求項150】

複数のsiRNAの総濃度は、標的遺伝子をサイレンシングするための最適の濃度であり、該最適濃度は、その濃度をさらに上昇させてもサイレンシングレベルが実質的に上昇しない濃度である、請求項149に記載の方法。

【請求項151】

最適濃度は、その濃度をさらに上昇させても標的遺伝子のサイレンシングレベルが5%より大きく上昇しない濃度である、請求項150に記載の方法。

【請求項152】

複数中の異なるsiRNAの数と該複数中の各siRNAの濃度は、前記複数のsiRNAが任意のオフターゲット遺伝子の30%未満のサイレンシングを引き起こすものである、請求項151に記載の方法。

【請求項153】

複数の異なるsiRNAは等しい比率で各siRNAを含む、請求項151に記載の方法。

【請求項154】

複数の異なるsiRNAは5%未満だけ互いに異なる比率で各siRNAを含む、請求項151に記載の方法。

【請求項155】

複数の異なるsiRNAのいずれも、複数のsiRNA中で総siRNA濃度の90%より大きい濃度を構成しない、請求項151に記載の方法。

【請求項156】

複数のsiRNA中の各siRNAは、他のsiRNAの非存在下または標的遺伝子をサイレンシングするように設計された他のsiRNAの非存在下で使用される時、少なくとも30%のサイレンシングを達成するのに有効なsiRNAの濃度より低い濃度を有する、請求項149に記載の方法。

【請求項157】

各siRNAは、他のsiRNAの非存在下または標的遺伝子をサイレンシングするように設計された他のsiRNAの非存在下で使用される時、30%未満のサイレンシングを引き起こす濃度を有する、請求項156に記載の方法。

【請求項158】

1つ以上の遺伝子のそれぞれの転写体配列との第1鎖の配列類似性と第2鎖の配列類似性とを比較することを含む、siRNAの第1鎖と第2鎖の相対的遺伝子サイレンシング活性を評価する方法。

【請求項159】

請求項158に記載の方法であって、比較は、

(a) 1つ以上の遺伝子のそれぞれについて、該遺伝子の転写体配列とsiRNAの第1鎖をアラインメントさせた場合の配列同一性の最も長い連続区間の長さを測定し；

(b) 1つ以上の遺伝子のそれぞれについて、該遺伝子の転写体配列とsiRNAの第2鎖をアラインメントさせた場合の配列同一性の最も長い連続区間の長さを測定し；そして

10

20

30

40

50

(c)トータルの第1鎖の同一性長さとトータルの第2鎖の同一性長さとを比較し、ここで、トータルの第1鎖の同一性長さは、1つ以上の遺伝子について工程(a)で測定した長さを合計し、トータルの第2鎖の同一性長さは、1つ以上の遺伝子について工程(b)で測定した長さを合計することにより得られる、方法により行われる方法。

【請求項160】

請求項158に記載の方法であって、遺伝子の転写体配列への第1鎖の配列類似性は、(i)該遺伝子の転写体配列と第1鎖をアラインメントさせた場合の配列同一性の最も長い連続区間の長さで示され、遺伝子の転写体配列への第2鎖の配列類似性は、(ii)前記遺伝子の転写体配列と第2鎖をアラインメントさせた場合の配列同一性の最も長い連続区間の長さで示される方法。

10

【請求項161】

比較は、遺伝子について(i)と(ii)との差を測定することを含む、請求項160に記載の方法。

【請求項162】

請求項161に記載の方法であって、差は、(i)と(ii)の短い方より過剰にある(i)と(ii)の長い方のヌクレオチドの数であり、配列類似性の比較は、

(a)該1つ以上の遺伝子について、(i)の総ヌクレオチド数から(ii)の総ヌクレオチド数を引いた数を決定し；

(b)該1つ以上の遺伝子について、(ii)の総ヌクレオチド数から(i)の総ヌクレオチド数を引いた数を決定し；そして

20

(c)もし工程(a)で決定した総ヌクレオチド数が工程(b)で決定した総ヌクレオチド数より大きい場合は、siRNAを、第1鎖の遺伝子サイレンシング活性より高い第2鎖のサイレンシング活性を有するとして確定し、またはもし工程(b)で決定した総ヌクレオチド数が工程(a)で決定した総ヌクレオチド数より大きい場合は、第2鎖の遺伝子サイレンシング活性より高い第1鎖のサイレンシング活性を有するとして確定することを含む方法。

【請求項163】

請求項158に記載の方法であって、比較は、

(a)siRNAの第1鎖の配列と同一の少なくとも7塩基の配列の連続区間を有し、かつ該siRNAの第1鎖の3'末端の3塩基内で終止する、1つ以上の遺伝子中の各遺伝子を同定し；

(b)siRNAの第2鎖の配列と同一の少なくとも7塩基の配列の連続区間を有し、かつ該siRNAの第2鎖の3'末端の3塩基内で終止する、1つ以上の遺伝子中の各遺伝子を同定し；そして

30

(c)工程(a)で同定された遺伝子の数と工程(b)で同定された遺伝子の数とを比較する、ことにより行われ、

ここでもし工程(a)で同定された遺伝子の数が工程(b)で同定された遺伝子の数より大きい場合は、siRNAは第1鎖のサイレンシング活性より高い第2鎖のサイレンシング活性を有すると確定され、またはもし工程(b)で同定された遺伝子の数が工程(a)で同定された遺伝子の数より大きい場合は、siRNAは第2鎖のサイレンシング活性より高い第1鎖のサイレンシング活性を有すると確定される方法。

【請求項164】

40

工程(c)における比較は、以下の式に従ってSLRスコア算出することにより行われる請求項163に記載の方法：

$$SLR = \log \left(\frac{\text{第1鎖同一遺伝子の数}}{\text{第2鎖同一遺伝子の数}} \right)$$

(ここで、第1鎖同一遺伝子の数は、工程(a)で同定された遺伝子の数であり、第2鎖同一遺伝子の数は、工程(b)で同定された遺伝子の数である)。

【請求項165】

1つ以上の遺伝子は少なくとも5個の異なる遺伝子である、請求項158～164のいずれか1項に記載の方法。

【請求項166】

1つ以上の遺伝子は、siRNAにより直接サイレンシングされる遺伝子からなる、請求項

50

165に記載の方法。

【請求項167】

siRNAにより直接サイレンシングされる遺伝子は、該siRNAにより同じ動力学でダウンレギュレートされる遺伝子である、請求項166に記載の方法。

【請求項168】

1つ以上の遺伝子は、siRNAにより直接サイレンシングされる遺伝子と、siRNAにより間接的にサイレンシングされる遺伝子とを含む、請求項165に記載の方法。

【請求項169】

1つ以上の遺伝子は、細胞へのsiRNAの導入後のある時点で検出すると、その転写体レベルがsiRNAにより低下させられる遺伝子からなる、請求項165に記載の方法。

10

【請求項170】

ある時点は、細胞へのsiRNAの導入後12時間である、請求項169に記載の方法。

【請求項171】

ある時点は、細胞へのsiRNAの導入後24時間である、請求項169に記載の方法。

【請求項172】

工程(c)の比較は、下記式に従ってSLRスコアを計算することにより行われる、請求項162に記載の方法：

$SLR = \log(1 \text{ つ以上の遺伝子中の第2鎖より過剰の第1鎖中の総ヌクレオチド数} / 1 \text{ つ以上の遺伝子中の第2鎖より過剰の第2鎖中の総ヌクレオチド数})$ 。

【発明の詳細な説明】

20

【技術分野】

【0001】

本出願は、35 U.S.C. & 119(e)に従って米国仮特許出願第60/515,223号(2003年10月27日出願)および米国仮特許出願第60/471,392号(2003年5月16日出願)(これらは参照することによりその全体が本明細書に組み込まれる)の利益を請求する。

【0002】

1. 発明の分野

本発明は、RNA干渉(RNA interference)を使用する遺伝子サイレンシング(gene silencing)のための方法と組成物に関する。本発明また、遺伝子をサイレンス(silence)させるように設計された複数の小干渉性RNAに対する共通のおよび/または鑑別的応答を同定する方法にも関する。本発明はさらに、遺伝子サイレンシングのための小干渉性RNAを設計する方法に関する。本発明はさらに、治療薬としてsiRNAを使用する方法に関する。

30

【背景技術】

【0003】

2. 発明の背景

RNA干渉(RNAi)は、哺乳動物細胞中の遺伝子発現を抑制するための強力な方法であり、科学会で大きな刺激を与えている(Couzin, 2002, Science 298:2296-2297; McManusら、2002, Nat. Rev. Genet. 3, 737-747; Hannon, G.J. 2002, Nature 418, 244-251; Padisonら、2002, Cancer Cell 2, 17-23)。RNA干渉は、進化を通して線虫(C. elegans)からヒトまで保存されており、RNAウイルスの侵入から細胞を防御する機能をすると考えられている。細胞はdsRNAウイルスで感染されると、dsRNAはRNaseIII型酵素(ダイサー(Dicer)と呼ぶ)により認識され切断の標的となる。ダイサー(Dicer)酵素はRNAを「さいの目に切って」、完全に対になった19ntのリボヌクレオチドからなるが各鎖の3'末端上に2つの対になっていないヌクレオチドを有する、21ntの短い2本鎖(siRNAまたは短干渉性RNAと呼ぶ)にする。これらの短い2本鎖は、リスク(RISC)と呼ぶマルチタンパク質複合体と結合し、この複合体を、siRNAと配列類似性を有するmRNA転写体に向ける。その結果、リスク(RISC)複合体中に存在するヌクレアーゼはmRNA転写体を切断し、こうして遺伝子産物の発現を止める。ウイルス感染の場合、この機構によりウイルス転写体が破壊され、ウイルス合成が妨害されるであろう。siRNAは2本鎖であるため、いずれかの鎖はリスク(RISC)に結合して、配列類似性を有する転写体のサイレンシングを指令する。

40

50

【0004】

特異的遺伝子サイレンシングは、遺伝子機能を解明し、薬剤標的を特定し、より特異的な治療薬を開発するためにヒトゲノムデータを利用できる可能性がある。これらの応用の多くは、その目的の標的に対してsiRNAが高度に特異性を有すると仮定する。siRNA配列と部分的な同一性を有する転写体とのクロスハイブリダイゼーションは、標的遺伝子以外に目的ではない転写体のサイレンシングを反映する表現型を誘発するかも知れない。これは、表現型に含まれる遺伝子の同定を複雑にするであろう。無数の文献の報告が、siRNAの微妙な特異性を記載し、siRNA配列とのほとんど完全な同一性の必要性を示唆している (Elbashirら、2001, *EMBO J.* 20:6877-6888; Tuschlら、1999, *Genes Dev.* 13:3191-3197; Hutvagnerら、2002, *Scienceexpress* 297:2056-2060)。最近の1つの報告は、siRNA標的化転写体切断には完全な配列相補性が必要であるが、部分的な相補性はマイクロRNAの形で、転写体分解無しで翻訳抑制を引き起こすと予測している (Hutvagnerら、2002, *Scienceexpress* 297:2056-2060)。しかしsiRNA誘導性遺伝子サイレンシングの公表された分析のほとんどは、標的遺伝子以外に1つまたは数個の遺伝子を調べるのみで、「電柱の下の鍵を捜す」アプローチに似ていないこともない。

10

【0005】

siRNAとmiRNAを含む小制御性RNAの生物学的機能はあまりわかっていない。1つのよくある質問は、これらの2つのクラスの制御性RNAの明確なサイレンシング経路が決定される機構に関するものである。miRNAはゲノムから発現される制御性RNAであり、前駆体ステム-ループ構造から処理されて、標的mRNAの3'UTR中の配列に結合する1本鎖核酸を産生する (Leeら、1993, *Cell* 75:843-854; Reinhartら、2000, *Nature* 403:901-906; Leeら、2001, *Science* 294:862-864; Lauら、2001, *Science* 294:858-862; Hutvagnerら、2001, *Science* 293:834-838)。miRNAはほんのわずかの相補性で転写体配列に結合し (Zengら、2002, *Molec. Cell* 9:1327-1333)、定常状態RNAレベルに影響を与えることなく翻訳を抑制する (Leeら、1993, *Cell* 75:843-854; Wightmanら、1993, *Cell* 75:855-862)。miRNAとsiRNAの両方とも、ダイサー (Dicer) により処理され、RNA誘導性サイレンシング複合体の成分に結合する (Hutvagnerら、2001, *Science* 293:834-838; Grishokら、2001, *Cell* 106:23-34; Kettingら、2001, *Genes Dev.* 15:2654-2659; Williamsら、2002, *Proc. Natl. Acad. Sci. USA* 99:6889-6894; Hammondら、2001, *Science* 293:1146-1150; Moulato sら、*Genes Dev.* 16:720-728)。最近の報告 (Hutvagnerら、2002, *Scienceexpress* 297:2056-2060) は、miRNA経路対siRNA経路を介する遺伝子制御が、標的転写体に対する相補性の程度によってのみ決定されるという仮説を報告している。miRNA標的に対してほんのわずかの同一性を有するsiRNAは、RNA分解を開始させるのではなく、miRNAと同様に翻訳抑制で機能すると推測されている。

20

30

【0006】

また、siRNAとshRNAは *in vivo* で遺伝子をサイレンシングするのに使用できることも証明されている。*in vivo* の遺伝子サイレンシングのためにsiRNAとshRNAを利用する能力は、治療用途のsiRNAの選択と開発とを可能にする可能性がある。最近の報告は、siRNAの治療的応用の可能性を強調している。広範囲の肝臓疾患においてFas介在アポトーシスの関与が示唆されており、肝細胞のアポトーシス性死滅を阻害することにより生命を救うことができるかも知れない。Song (Songら、2003, *Nat. Medicine* 9:347-351) は、Fas受容体を標的とするsiRNAをマウスに静脈内注射した。Fas遺伝子は、mRNAとタンパク質レベルでマウス肝細胞にサイレンシングされ、アポトーシスを防ぎ、マウスを肝炎誘導肝臓傷害から防御した。すなわちFas発現をサイレンシングすることは、細胞毒性から肝細胞を防御して肝臓傷害を防ぐのに有望である。別の例として、マウスにTNF- α を標的とするsiRNAを腹腔内注入した。リポ多糖誘導性TNF- α 遺伝子発現が阻害され、これらのマウスは敗血症から防御された。まとめるとこれらの結果は、siRNAがインピボで機能し、治療薬として有望であることを示唆する (Sorensenら、2003, *J. Mol. Biol.* 327:761-766)。

40

【0007】

US Patent No. 6,506,559は、細胞中の標的遺伝子の発現を阻害するためのRNA干渉法を

50

開示する。この方法は、標的遺伝子中の配列と同一の配列を2本鎖領域中に有する部分的にまたは完全に2本鎖のRNAを、細胞または細胞外環境に導入することを含む。標的配列に対して挿入、欠失、および単一の点突然変異を有するRNA配列はまた、発現阻害に有効であることがわかっている。

【0008】

米国特許出願公報No.2002/0086356は、21~23ヌクレオチドの長さのRNAセグメントを使用してキイロショウジョウバエ (*Drosophila*) の *in vitro*系でのRNA干渉を開示する。本特許出願公報は、これらの21~23nt断片を精製し、キイロショウジョウバエ (*Drosophila*) 抽出物に戻し加えると、これらは長いdsRNAの非存在下で配列特異的RNA干渉を仲介することを教示する。本特許出願公報はまた、哺乳動物細胞中の分解について特異的mRNAを標

10

【0009】

PCT公報W002/44321は、19~23ntの長さの2本鎖RNA (dsRNA) が、*in vitro*でキイロショウジョウバエ (*Drosophila*) の配列特異的翻訳後遺伝子サイレンシングを誘導することを開示する。このPCT公報は、長いdsRNAまたはオーバーハング3'末端を有する化学合成したsiRNA2本鎖からRNaseIII様プロセッシング反応により作成した小干渉性RNA (siRNA) が、溶解物中の標的RNA切断を仲介し、切断部位がガイディングsiRNAがまたがる領域の中心近くに位置することを教示する。このPCT公報はまた、センスまたはアンチセンス標的RNAが、

20

産生されたsiRNA複合体により切断されることができるとい証拠を提供する。

【0010】

米国特許出願公報US2002/016216は、標的遺伝子のヌクレオチド配列にストリンジェントな条件下でハイブリダイズするヌクレオチド配列を含む2本鎖RNA (dsRNA) を、標的遺伝子の発現を弱めるのに十分な量で細胞中に導入することにより、培養細胞中の標的遺伝子の発現を弱めるための方法を開示する。

【0011】

PCT公報W003/006477は、細胞中で発現されると細胞により処理されて、細胞自体のRNA干渉 (RNAi) 経路を使用して標的化遺伝子を選択的にサイレンシングする (特異的mRNAを除去することにより) 標的化小干渉性RNA (siRNA) を産生する遺伝子操作したRNA前駆体

30

を開示する。このPCT公報は、これらの遺伝子操作したRNA前駆体をコードする核酸分子を、適切な制御配列を有する細胞に *in vivo*で導入することにより、遺伝子操作されたRNA前駆体の発現を、一時的にかつ空間的に、すなわち特定の時間におよび/または特定の組織、臓器、または細胞中で選択的に制御することができることを教示する。

【0012】

DNAアレイ技術は、多数の遺伝子転写体の発現レベルを一度に追跡することを可能にした (例えば、Schenら、1995, *Science* 270:467-470; Lochhartら、1996, *Nature Biotechnology* 14:1675-1680; Blanchardら、1996, *Nature Biotechnology* 14:1649; Ashbyら、US Patent No. 5,569,588、1996年10月29日発行、を参照)。DNAアレイの2つの主要なフォーマットのうち、スポットcDNAアレイは、サイズが約0.6~2.4kbのcDNA断片、完全長cD

40

NA、ESTなどのPCR産物を適当な表面上に沈積することにより調製される (例えば、DeRisiら、1996, *Nature Genetics* 14:457-460; Shalonら、1996, *Genome Res.* 6:689-645; Schenaら、1995, *Proc. Natl. Acad. Sci. USA* 93:10539-11286; および Dugganら、*Nature Genetics Supplement* 21:10-14)。あるいは、表面上の規定の位置に規定の配列に相補的な数千のオリゴヌクレオチドを含有する高密度オリゴヌクレオチドアレイは、例えば写真平版法により表面上に *in situ*で合成される (例えば、Fodorら、1991, *Science* 251:767-773; Peaseら、1994, *Proc. Natl. Acad. Sci. USA* 91:5022-5026; Lockhartら、1996, *Nature Biotechnology* 14:1675; McGallら、1996, *Proc. Natl. Acad. Sci. USA* 93:13555-13560; US Patent No. 5,578,832; 5,556,752; 5,510,270; および6,040,138参照)。 *in situ*オリゴヌクレオチド合成のためにインクジェット技術を使用してアレイを作成する

50

方法が、当該分野で公知である（Blanchard, 国際特許公報W098/41531、1998年9月24日公表；Blanchardら、1996, Biosensors and Bioelectronics 11:687-690；Blanchard, 1998, 「遺伝子工学における合成DNAアレイ（Synthetic DNA Arrays in Genetic Engineering）」、第20巻、J.K. Setlow編、プレナム・プレス（Plenum Press）、ニューヨーク、111～123頁参照）。DNAアレイの情報能力をさらに上昇させる試みは、DNAアレイ上の構造サイズをさらに縮小して、ある表面積中のプローブの数をさらに増加させることから、各標的核酸の検出に必要な重複プローブの数を低下させることを目的とする感度と特異性ベースのプローブ設計と選択をして、プローブ密度を上昇させることなく追跡される標的核酸の数を増加させることにわたる（例えば、Friendら、米国特許出願第09/364,751号（1999年6月30日出願）；およびFriendら、米国特許出願第09/561,487号（2000年4月28日出願）参照）。

【0013】

数万の遺伝子を同時に追跡することにより、DNAアレイ技術は特に、細胞または細胞タイプまたは任意の生物学的サンプル中のmRNA発現のゲノム全体にわたる分析を可能にする。高度なデータ管理と分析法に補助されて、細胞または細胞タイプの転写状態ならびに外部摂動（特に限定されないが薬剤の摂動）に応答した転写状態の変化を、mRNAレベルで解析することができる（例えば、Stoughtonら、国際公報W0003/39336（2000年6月6日公開）；Friendら、国際公報W000/24936（2000年5月4日公開）；およびShoemaker、国際公報W002/16550（2002年2月28日）。かかる技術の応用には、例えば種々の生理学的状態（特に疾患状態）でアップレギュレートまたはダウンレギュレートされている遺伝子の同定を含む。DNAアレイの追加の用途例には、シグナル伝達経路のメンバーの分析、および種々の薬物の標的の同定がある。例えばFriendとHartwell、国際公報W098/38329（1998年9月3日公開）；Stoughton、国際公報W099/66067（1999年12月23日公開）；StoughtonとFriend、国際公報W099/58708（1999年11月18日公開）；FriendとStoughton、国際公報W099/59037（1999年11月18日公開）；Friendら、米国特許出願第98/334,328（1999年6月16日出願）を参照されたい。

【0014】

本明細書に記載の文献の考察または引用は、これらが本発明の先行技術であることを認めるものではない。

【発明の開示】

【0015】

3. 発明の要約

本発明は、真核細胞を小干渉性RNA（siRNA）の分子に付すことを含むRNA干渉により該細胞中の標的遺伝子をサイレンシングさせる方法であって、前記siRNAは、前記標的遺伝子の転写体の配列と同一の、センス鎖またはアンチセンス鎖の少なくとも11ヌクレオチドの連続的ヌクレオチド配列を含むが、前記siRNAは、前記転写体中のどの配列とも完全長のセンス鎖またはアンチセンス鎖配列の同一性を持たず、前記連続的ヌクレオチド配列は前記siRNAの中央領域にあることを特徴とする方法を提供する。本発明のある実施形態においてsiRNAは、前記標的遺伝子の前記転写体と同一の16、15、14、13、12、または11ヌクレオチドより大きいセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まない。

【0016】

本発明はまた、真核細胞を小干渉性RNA（siRNA）の分子に付すことを含むRNA干渉により該細胞中の標的遺伝子をサイレンシングさせる方法であって、前記siRNAは、前記標的遺伝子の転写体の配列と同一の、センス鎖またはアンチセンス鎖の少なくとも9ヌクレオチドの連続的ヌクレオチド配列を含むが、前記siRNAは、前記転写体中のどの配列とも完全長のセンス鎖またはアンチセンス鎖配列の同一性を持たず、前記連続的ヌクレオチド配列は前記siRNAの3'末端にあることを特徴とする方法を提供する。本発明のある実施形態においてsiRNAは、前記標的遺伝子の前記転写体と同一の16、15、14、13、12、11、10、または9ヌクレオチドより大きいセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配

列を含まない。

【0017】

本発明はまた、複数の異なる遺伝子のそれぞれの転写体配列が該複数の異なる遺伝子に共通する9~18ヌクレオチドのヌクレオチド配列を含む、RNA干渉により真核細胞中の複数の異なる遺伝子をサイレンシングさせる方法であって、前記細胞を小干渉性RNA (siRNA) の分子に付すことを含み、前記 siRNA は、(i) 前記共通配列中の配列と同一の少なくとも11ヌクレオチドのセンス鎖またはアンチセンス鎖の中央の連続的ヌクレオチド配列および / または (ii) 該共通配列中の配列と同一の少なくとも9ヌクレオチドの3'センス鎖またはアンチセンス鎖の連続的ヌクレオチド配列とを含む方法を提供する。ある実施形態において中央の連続的ヌクレオチド配列は、11~15、14~15、13、12、または11ヌクレオチドの長さである。

10

【0018】

本発明はまた、(i) 第1遺伝子の転写体配列と同一の少なくとも11ヌクレオチドのセンス鎖またはアンチセンス鎖の中央の連続的ヌクレオチド配列、または (ii) 前記第1遺伝子の転写体配列と同一の少なくとも9ヌクレオチドの3'センス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含む siRNA 分子に前記真核細胞を付すことを含み、RNA干渉により真核細胞中の第1遺伝子をサイレンシングするが第2遺伝子はサイレンシングしない方法であって、前記 siRNA は、前記第2遺伝子の前記転写体の配列と同一の10ヌクレオチドを超えるセンス鎖またはアンチセンス鎖の中央の連続的ヌクレオチド配列を含まず、かつ前記 siRNA は、前記第2遺伝子の前記転写体の配列と同一の8ヌクレオチドを超える3'センス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まないことを特徴とする方法を提供する。ある実施形態において、siRNA分子は、前記第2遺伝子の転写体の任意の配列と同一の8ヌクレオチドを超える長さの連続的ヌクレオチド配列を含まない。ある実施形態において、(i) の中央の連続的ヌクレオチド配列は11~15、14~15、13、12、または11ヌクレオチドの長さである。別の実施形態において、(ii) の連続的ヌクレオチド配列は9~15、9~12、10、または9ヌクレオチドの長さである。

20

【0019】

本発明はまた、(i) 第1遺伝子の転写体配列と同一の少なくとも11ヌクレオチドのセンス鎖またはアンチセンス鎖の中央の連続的ヌクレオチド配列、または (ii) 第1遺伝子の転写体配列と同一の少なくとも9ヌクレオチドの3'センス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含む siRNA 分子を同定することを含み、RNA干渉により真核細胞中の第1遺伝子をサイレンシングするが第2遺伝子はサイレンシングしないための小干渉性RNAを設計する方法であって、該 siRNA は前記第2遺伝子の転写体の配列と同一の10ヌクレオチドを超える長さのセンス鎖またはアンチセンス鎖の中央の連続的ヌクレオチド配列を含まず、かつ前記 siRNA は前記第2遺伝子の転写体の任意の配列と同一の8ヌクレオチドを超える長さの3'センス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まないことを特徴とする方法を提供する。ある実施形態において、siRNA分子は、該第2遺伝子の転写体の任意の配列と同一の8ヌクレオチドを超える長さの連続的ヌクレオチド配列を含まない。別の実施形態において、(i) の中央の連続的ヌクレオチド配列は11~15、14~15、13、12、または11ヌクレオチドの長さである。別の実施形態において、(ii) の連続的ヌクレオチド配列は9~15、9~12、10、または9ヌクレオチドの長さである。

30

40

【0020】

本発明はまた、各異なる siRNA は遺伝子の転写体中の異なる配列を標的とするように設計されることを特徴とする、真核細胞中の遺伝子をサイレンシングするために複数の異なる siRNA から1つ以上の小干渉性RNA (siRNA) を選択する方法であって、(a) 各 siRNA について応答プロファイルを測定し、(該応答プロファイルは、複数の遺伝子の発現レベルの測定値を含む) ; そして (b) その応答プロファイルに基づいて1つ以上の siRNA を選択することを含む方法を提供する。ある実施形態において、複数の異なる siRNA は、その配列が前記標的遺伝子の一部のまたは全てのコード配列を分割する siRNA を含む。別の実施形態において、複数の異なる siRNA は、その配列が10塩基の間隔で前記標的遺伝子の一部の

50

または全てのコード配列をうめる(tile across)するsiRNAを含む。別の実施形態において、複数の異なるsiRNAは、その配列が5塩基の間隔で前記標的遺伝子の一部のまたは全てのコード配列をうめるsiRNAを含む。別の実施形態において、複数の異なるsiRNAは、その配列が1塩基の間隔で前記標的遺伝子の一部のまたは全てのコード配列をうめるsiRNAを含む。ある実施形態において1つ以上のsiRNAは、各siRNAの応答プロファイルを所望の応答プロファイルと比較し、その応答プロファイルが所望の応答プロファイルに一致する1つ以上のsiRNAを選択することにより選択される。

【0021】

本発明はまた、複数の異なる遺伝子は9~18ヌクレオチドの共通のヌクレオチド配列を共有することを特徴とする、真核細胞中の複数の異なる遺伝子を標的とする小干渉性RNA (siRNA) を設計する方法であって、(i)前記共通配列中の配列と同一の少なくとも11ヌクレオチドのセンス鎖またはアンチセンス鎖の中央の連続的ヌクレオチド配列および/または(ii)該共通配列中の配列と同一の少なくとも9ヌクレオチドの3'センス鎖またはアンチセンス鎖の連続的ヌクレオチド配列とを含むsiRNAを選択することを含む方法を提供する。ある実施形態において、中央の連続的ヌクレオチド配列は11~15、14~15、13、12、または11ヌクレオチドの長さである。

10

【0022】

本発明はまた、真核細胞をsiRNAに付した後の選択した時点で該真核細胞の発現プロファイルを測定することを含む、前記真核細胞に対するsiRNAの作用を測定する方法であって、各発現プロファイルは複数の異なる遺伝子の測定された転写体レベルを含む方法を提供する。本発明の方法において、複数の異なる遺伝子は5、10、100、1000、10,000、または25,000の異なる遺伝子を含むことができる。

20

【0023】

本発明はまた、真核細胞に対するsiRNAの作用を測定する方法であって、(a)該真核細胞を前記siRNAに付した後の複数の異なる各時点で前記真核細胞の発現プロファイルを測定し(ここで各発現プロファイルは複数の異なる遺伝子の測定された転写体レベルを含む);そして、(b)測定された転写体レベルの動力学的挙動に基づいて、前記複数の遺伝子を異なる動力学群に分類することを含む方法を提供する。本発明の方法において、複数の異なる遺伝子は5、10、100、1000、10,000、または25,000の異なる遺伝子を含むことができる。

30

【0024】

本発明はまた、siRNAが真核細胞中の標的遺伝子をサイレンシングするように設計される、該小干渉性RNA (siRNA) によりサイレンシングされる前記真核細胞中の1つ以上の遺伝子を同定する方法であって、(a)前記真核細胞を前記siRNAに付した後の複数の異なる各時点で前記真核細胞の発現プロファイルを測定し(ここで各発現プロファイルは複数の異なる遺伝子の測定された転写体レベルを含む);(b)その転写体レベルが、標的遺伝子にコードされるタンパク質レベルの低下より実質的に速く低下する、前記複数の異なる遺伝子の中の1つ以上の遺伝子を同定することを含む方法を提供する。ある実施形態において、標的遺伝子によりコードされる前記タンパク質のレベルがその撓動されないレベルの約50%まで低下する前に、前記1つ以上の遺伝子の転写体レベルは少なくとも50%低下する。別の実施形態において前記タンパク質レベルの低下は、前記複数の異なる各時点で前記タンパク質の量を測定することを含む方法により測定される。本発明の方法において、複数の異なる遺伝子は5、10、100、1000、10,000、または25,000の異なる遺伝子を含むことができる。

40

【0025】

本発明はまた、各siRNAは、真核細胞中の同じ標的遺伝子をサイレンシングするように設計される、複数の異なる小干渉性RNA分子 (siRNA) により直接サイレンシングされる真核細胞中の1つ以上の遺伝子を同定する方法であって、(a)複数の異なる各siRNAについて前記真核細胞の発現プロファイルを測定し(ここで各発現プロファイルは複数の異なる遺伝子の測定された転写体レベルを含む);そして、(b)その転写体レベルが前記複数のsiR

50

NAにより一般的に影響を受ける、前記複数の異なる遺伝子の中の1つ以上の遺伝子を同定することを含む方法を提供する。ある実施形態において、工程(b)は、異なるsiRNAの発現プロファイルの中に前記複数の遺伝子を集めることを含む方法により行われる。別の実施形態において各発現プロファイルは、各siRNAの導入後に同時に測定される。別の実施形態において、前記の時点は、標的遺伝子によりコードされるタンパク質レベルのその撓動されないレベルの約50%低下の時間スケールより実質的に短い。本発明の方法において、複数の異なる遺伝子は5、10、100、1000、10,000、または25,000の異なる遺伝子を含むことができる。

【0026】

本発明はさらに、1つ以上の候補遺伝子の発現レベルの変化が真核細胞中の表現型の特徴になる、該真核細胞中の1つ以上の候補遺伝子を同定する方法であって、(a)真核細胞への小干渉性RNA (siRNA) の導入に関連する真核細胞の表現型特徴を同定し (該siRNAは、真核細胞中の標的遺伝子をサイレンシングするように設計される) ; (b)siRNAの導入後の複数の異なる各時点で真核細胞の発現プロファイルを測定し (各発現プロファイルは、複数の異なる遺伝子の測定された転写体レベルを含む) ; (c)その転写体レベルが、前記標的遺伝子によりコードされるタンパク質のレベルの低下より実質的に速く低下する、前記複数の遺伝子から1つ以上の遺伝子を同定し ; そして、(d)前記1つ以上の遺伝子を1つ以上の候補遺伝子として同定する、ことを含む方法を提供する。ある実施形態において、標的遺伝子によりコードされる前記タンパク質のレベルがその撓動されないレベルの約50%まで低下する前に、前記1つ以上の遺伝子の転写体レベルは少なくとも50%低下する。別の実施形態において前記タンパク質レベルの低下は、前記複数の異なる各時点で前記タンパク質の量を測定することを含む方法により測定される。本発明の方法において、複数の異なる遺伝子は5、10、100、1000、10,000、または25,000の異なる遺伝子を含むことができる。

【0027】

本発明はさらに、1つ以上の候補遺伝子の発現レベルの変化が真核細胞中の表現型の特徴になる、該真核細胞中の1つ以上の候補遺伝子を同定する方法であって、(a)複数の異なる各小干渉性RNA分子 (siRNA) のそれぞれの導入に通常関連する真核細胞の表現型特徴を同定し (前記siRNAは、真核細胞中の同じ標的遺伝子をサイレンシングするように設計される) ; (b)複数の異なる各小干渉性RNA分子 (siRNA) のそれぞれの発現プロファイルを測定し (各siRNAは前記真核細胞中の同じ標的遺伝子をサイレンシングするように設計され、かつ各発現プロファイルは、複数の異なる遺伝子の測定された転写体レベルを含む) ; (c)その転写体レベルが、前記複数のsiRNAにより通常影響を受ける、前記複数の遺伝子から1つ以上の遺伝子を同定し ; そして(d)前記1つ以上の遺伝子を1つ以上の候補遺伝子として同定する、ことを含む方法を提供する。ある実施形態において、工程(c)は、異なるsiRNAの発現プロファイルの中に前記複数の遺伝子を集めることを含む方法により行われる。別の実施形態において、各発現プロファイルは、各siRNAの導入後に同じ時点で測定される。別の実施形態においてその時点は、標的遺伝子によりコードされるタンパク質レベルのその撓動されないレベルの約50%低下の時間スケールより実質的に短い。本発明の方法において、複数の異なる遺伝子は5、10、100、1000、10,000、または25,000の異なる遺伝子を含むことができる。

【0028】

本発明はさらに、真核細胞中で第1タンパク質により制御されるが第2タンパク質には制御されない1つ以上の遺伝子を同定する方法であって、(a)(i)該第1タンパク質をコードする第1遺伝子の転写体配列と同一の少なくとも11ヌクレオチドのセンス鎖またはアンチセンス鎖の中央の連続的ヌクレオチド配列、または(ii)前記第1遺伝子の転写体配列と同一の少なくとも9ヌクレオチドの3'センス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含む、siRNAの分子に前記真核細胞を付し (前記siRNAは、第2タンパク質をコードする第2遺伝子の転写体配列と同一の10ヌクレオチドを超えるセンス鎖またはアンチセンス鎖の中央の連続的ヌクレオチド配列を含まず、かつsiRNAは、第2遺伝子の転写体配列と

同一の8ヌクレオチドを超える3'センス鎖またはアンチセンス鎖の連続的ヌクレオチド配列とを含まない) ; (b)前記真核細胞を前記siRNAに付した後のある期間後に、真核細胞の発現プロファイルを測定し(各発現プロファイルは、複数の異なる遺伝子の測定された転写体レベルを含み、その期間は、siRNAに付されていない細胞中の第1タンパク質のレベルと比較して、第1タンパク質のレベルの50%低下が観察される時間より長い) ;そして、(c)第1タンパク質によりサイレンシングされるが第2タンパク質にはサイレンシングされない遺伝子として、siRNAに付されていない細胞中のレベルからはその転写体レベルが変化している1つ以上の遺伝子を同定する、ことを含む方法を提供する。ある実施形態において、siRNA分子は、第2遺伝子の転写体の任意の配列と同一の、8ヌクレオチドを超える長さの連続的ヌクレオチド配列を含まない。別の実施形態において(i)中の中央の連続的ヌクレオチド配列は11~15、14~15、13、12、または11ヌクレオチドの長さである。別の実施形態において、(ii)の連続的ヌクレオチド配列は9~15、9~12、10、または9ヌクレオチドの長さである。本発明の方法において、複数の異なる遺伝子は5、10、100、1000、10,000、または25,000の異なる遺伝子を含むことができる。

10

20

30

40

50

【0029】

本発明はさらに、小干渉性RNA(siRNA)を含む真核細胞であって、該siRNAは、該真核細胞中の遺伝子の転写体配列と同一の少なくとも11ヌクレオチドのセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含むが、前記真核細胞のゲノム中の他の遺伝子の転写体の任意の配列と完全長のセンス鎖またはアンチセンス鎖の配列同一性は無く、前記連続的ヌクレオチド配列は前記siRNA分子の中央領域にある真核細胞を提供する。本発明のある実施形態においてsiRNAは、前記標的遺伝子の転写体配列と同一の16、15、14、13、12、または11ヌクレオチドより大きいセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まない。本発明はまた、小干渉性RNA(siRNA)を含む真核細胞であって、該siRNAは、該真核細胞中の遺伝子の転写体配列と同一の少なくとも11ヌクレオチドのセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含むが、前記真核細胞のゲノム中の他の遺伝子の転写体の任意の配列と完全長のセンス鎖またはアンチセンス鎖の配列同一性は無く、前記連続的ヌクレオチド配列は前記siRNA分子の3'末端にある真核細胞を提供する。本発明のある実施形態においてsiRNAは、前記標的遺伝子の転写体配列と同一の16、15、14、13、12、11、10、または9ヌクレオチドより大きいセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まない。真核細胞はヒト細胞でもよい。

【0030】

本発明はまた、真核細胞中の標的遺伝子をサイレンシングするように設計された第1のsiRNAと第2のsiRNAにより鑑別的にサイレンシングされる遺伝子を同定する方法であって、(a)前記真核細胞の第1の発現プロファイルと該真核細胞の第2の発現プロファイルとを比較し(第1の発現プロファイルは、前記細胞への第1siRNAの導入後の選択された時点で測定され、第2の発現プロファイルは前記細胞への第2siRNAの導入後の前記の選択された時点で測定され、前記発現プロファイルは複数の異なる遺伝子の測定された転写体レベルを含む) ;そして(b)その転写体レベルが前記第1および第2siRNAにより鑑別的に影響を受ける、前記標的遺伝子以外の遺伝子を同定する、ことを含む方法を提供する。ある実施形態において、前記の選択された時点は、標的遺伝子によりコードされるタンパク質レベルのその撓動されないレベルの50%低下の時間スケールより実質的に短く、前記同定される遺伝子は、前記第1および第2siRNAにより直接鑑別的にサイレンシングされる。別の実施形態において、前記の選択された時点は、標的遺伝子によりコードされるタンパク質レベルのその撓動されないレベルから、前記標的遺伝子によりコードされるタンパク質のレベルの50%低下の時間スケールより実質的に長く、前記同定される遺伝子は、直接サイレンシングされた遺伝子の鑑別的サイレンシングの結果として、前記第1および第2siRNAにより鑑別的にサイレンシングされる。本発明の方法において、複数の異なる遺伝子は5、10、100、1000、10,000、または25,000の異なる遺伝子を含むことができる。

【0031】

本発明の任意の方法において、真核細胞はヒト細胞である。

【0032】

本発明はさらに、哺乳動物に治療的に十分な量のsiRNAを投与することを含む哺乳動物の疾患または好ましくない症状の治療方法であって、該siRNAはその発現が該疾患または好ましくない症状を引き起こす遺伝子を標的とし、前記siRNAは、(i)前記遺伝子の転写体中の配列と同一の少なくとも11ヌクレオチドであるが18ヌクレオチド以下のセンス鎖またはアンチセンス鎖の中央の連続的ヌクレオチド配列および/または(ii)前記転写体中の配列と同一の少なくとも9ヌクレオチドであるが18ヌクレオチド以下の3'センス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含む、治療方法を提供する。本発明のある実施形態においてsiRNAは、該遺伝子の転写体配列と同一の16、15、14、13、12、または11ヌクレオチドより大きいセンス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まない。ある実施形態において哺乳動物はヒトである。

【0033】

別の態様において本発明は、RNA干渉により真核細胞中の標的遺伝子をサイレンシングする方法を提供する。本方法は、標的遺伝子をサイレンシングするように設計された複数の異なるsiRNAを真核細胞中に挿入することを含む。好適な実施形態において複数の異なるsiRNAは、少なくとも3、5、9、12、15、20、50、または100個の異なるsiRNAからなる。好適な実施形態において複数のsiRNAの総濃度は、標的遺伝子をサイレンシングするための最適の濃度である。かかる最適濃度は、その濃度をさらに上昇させてもサイレンシングレベルが実質的に上昇しない濃度でもよい。例えば最適濃度は、その濃度をさらに上昇させても前記標的遺伝子のサイレンシングレベルが5%、10%または20%より大きく上昇しない濃度でもよい。好適な実施形態において複数の組成物(複数には、異なるsiRNAの数と各異なるsiRNAの濃度を含む)は、複数のsiRNAが標的ではない遺伝子の30%、20%、10%、または5%、1%、0.1%または0.01%未満のサイレンシングを引き起こすように選択される。別の好適な実施形態において複数の異なるsiRNAは、等しい比率で各siRNAを含む。さらに別の好適な実施形態において複数の異なるsiRNAは、5%、10%、20%または50%未満だけ互いに異なる比率で各siRNAを含む。さらに別の好適な実施形態において、複数の異なるsiRNAのいずれも、該複数のsiRNA中で総siRNA濃度の90%、80%、70%、50%、または20%より大きい濃度を構成しない。他のある実施形態において、複数のsiRNA中の各siRNAは、他のsiRNAの非存在下または標的遺伝子をサイレンシングするように設計された他のsiRNAの非存在下で使用される時、少なくとも30%、50%、75%、80%、85%、90%、または95%のサイレンシングを達成するのに有効なsiRNAの濃度より低い濃度を有する。ある実施形態において各siRNAは、他のsiRNAの非存在下または標的遺伝子をサイレンシングするように設計された他のsiRNAの非存在下で使用される時、30%、20%、10%、または5%未満のサイレンシングを引き起こす濃度を有する。

【0034】

別の態様において本発明は、1つ以上の遺伝子のそれぞれの転写体配列との第1鎖の配列類似性と第2鎖の配列類似性とを比較することを含む、siRNAの第1鎖と第2鎖の相対的遺伝子サイレンシング活性を評価する方法を提供する。好ましくは1つ以上の遺伝子は、2、5、10、100、または1,000個の異なる遺伝子を含む。ある実施形態において1つ以上の遺伝子は、siRNAにより直接サイレンシングされる遺伝子からなる。ある実施形態においてsiRNAにより直接サイレンシングされる遺伝子は、同じ動力学によりsiRNAによりダウンレギュレートされる遺伝子である。ある実施形態において1つ以上の遺伝子は、siRNAにより直接サイレンシングされる遺伝子とsiRNAにより間接的にサイレンシングされる遺伝子とを含む。別の実施形態において1つ以上の遺伝子は、細胞へのsiRNAの導入後のある時点で検出されるsiRNAによりその転写体レベルが低下させられる遺伝子からなる。ある時点は、細胞へのsiRNAの導入の約12時間後または24時間後でもよい。

【0035】

ある実施形態において比較は、(a)1つ以上の遺伝子のそれぞれについて、遺伝子の転写体配列とsiRNAの第1鎖とをアラインメントさせた場合の配列同一性の最も長い連続区間の長さを測定し；(b)1つ以上の遺伝子のそれぞれについて、遺伝子の転写体配列とsiRNA

の第2鎖とをアラインメントさせた場合の配列同一性の最も長い連続区間の長さを測定し；そして(c)トータルの第1鎖の同一性長さトータルの第2鎖の同一性長さとを比較する(ここで、トータルの第1鎖の同一性長さは、1つ以上の遺伝子について工程(a)で測定した長さを合計し、トータルの第2鎖の同一性長さは、1つ以上の遺伝子について工程(b)で測定した長さを合計することにより得られる)ことを含む方法により行われる。

【0036】

別の実施形態において、遺伝子の転写体配列への第1鎖の配列類似性は、(i)該遺伝子の転写体配列と第1鎖とをアラインメントさせた場合の配列同一性の最も長い連続区間の長さで示され、遺伝子の転写体配列への第2鎖の配列類似性は、(ii)前記遺伝子の転写体配列と第2鎖とをアラインメントさせた場合の配列同一性の最も長い連続区間の長さで示される。好ましくはその比較は、遺伝子の(i)と(ii)との差を測定することを含む。ある実施形態において差は、(i)と(ii)の短い方より過剰にある(i)と(ii)の長い方のヌクレオチドの数であり、配列類似性の比較は、(a)1つ以上の遺伝子中の第2鎖より過剰の第1鎖中の総ヌクレオチド数を決定し；(b)1つ以上の遺伝子中の第2鎖より過剰の第2鎖中の総ヌクレオチド数を決定し；そして(c)もし工程(a)で決定した総ヌクレオチド数が工程(b)で決定した総ヌクレオチド数より大きい場合は、siRNAを第1鎖の遺伝子サイレンシング活性より高い第2鎖のサイレンシング活性を有するとして確定し、またはもし工程(b)で決定した総ヌクレオチド数が工程(a)で決定した総ヌクレオチド数より大きい場合は、第2鎖の遺伝子サイレンシング活性より高い第1鎖のサイレンシング活性を有するとして確定することを含む方法により行われる。ある実施形態において工程(c)における比較は、以下の式に従ってSLRスコア算出することにより行われる

$$SLR = \log \left(\frac{\text{1つ以上の遺伝子中の第2鎖より過剰の第1鎖中の総ヌクレオチド数}}{\text{1つ以上の遺伝子中の第2鎖より過剰の第2鎖中の総ヌクレオチド数}} \right)$$

【0037】

別の実施形態において第1鎖および第2鎖中の総ヌクレオチド数は、バックグラウンドの過剰のヌクレオチド数により補正される。

【0038】

さらに別の実施形態において比較は、(a)siRNAの第1鎖の配列と同一の少なくとも7塩基の配列の連続区間を有し、かつ該siRNAの第1鎖の3'末端の3塩基内で終止する、1つ以上の遺伝子中の各遺伝子を同定し；(b)前記siRNAの第2鎖の配列と同一の少なくとも7塩基の配列の連続区間を有し、かつ前記siRNAの第2鎖の3'末端の3塩基内で終止する、1つ以上の遺伝子中の各遺伝子を同定し；そして(c)工程(a)で同定された遺伝子の数と工程(b)で同定された遺伝子の数とを比較することにより行われ、ここでもし工程(a)で同定された遺伝子の数が工程(b)で同定された遺伝子の数より大きい場合は、前記siRNAは第1鎖のサイレンシング活性より高い第2鎖のサイレンシング活性を有すると確定され、またはもし工程(b)で同定された遺伝子の数が工程(a)で同定された遺伝子の数より大きい場合は、前記siRNAは第2鎖のサイレンシング活性より高い第1鎖のサイレンシング活性を有すると確定される。ある実施形態において工程(c)の比較は、以下の式に従ってSLRスコア算出することにより行われる

$$SLR = \log \left(\frac{\text{第1鎖同一遺伝子の数}}{\text{第2鎖同一遺伝子の数}} \right)$$

ここで、第1鎖同一遺伝子の数は工程(a)で同定された遺伝子の数であり、第2鎖同一遺伝子の数は工程(b)で同定された遺伝子の数である。

【発明を実施するための最良の形態】

【0039】

4. 図面の簡単な説明

図1A~Bは、発現プロファイルにより証明されるsiRNA配列に特異的な遺伝子発現パターンを示す。図1A：MAPK14コード領域に対する8個の異なるsiRNA2本鎖標的を、HeLa細胞中で遺伝子サイレンシングのために使用した。Luc、ルシフェラーゼに対するsiRNA標的。図1B：IGF1Rコード領域に対する16個の異なるsiRNA2本鎖標的を遺伝子サイレンシングのために使用した。オリゴフェタミン(Invitrogen)とウェル当たり100nM siRNA2本鎖を使用

して、細胞を6ウェルプレート中でトランスフェクトした。siRNAはDharmacon(コロラド州ボルダー)から、アニーリングした精製2本鎖として得られた。siRNAトランスフェクトした細胞からのRNAを、モックトランスフェクト細胞(RNA2本鎖の非存在下でトランスフェクション試薬で処理した)からのRNAに対してハイブリダイズさせた。総RNAを、Qiagen RNeasyキットにより精製し、約21,000のヒト遺伝子に対応するオリゴヌクレオチドを含有するマイクロアレイへのハイブリダイゼーションのために処理した。マイクロアレイは、Agilent Technologiesから購入したかまたは合成した。各行は、個々のsiRNAのトランスフェクションにより生じる発現プロファイルである。記載のデータは、モックトランスフェクトした細胞に対して発現レベルの少なくとも2倍の変化(p 値 < 0.01 、および \log_{10} 強度 > 1)を示す遺伝子である。薄い灰色は発現の低下を示し、黒は発現の上昇を示す。データはRosetta Resolver(登録商標)ソフトウェアを使用して解析した。棒グラフは、siRNAトランスフェクション後に残存する標的タンパク質(灰色の棒)とRNA(黒い棒)の割合である。RNA定量は、IGF1RのためのAP Biosystems TaqManプレデベロップアッセイ試薬(#4319442)を使用してリアルタイムPCRにより行った。MAPK14のプライマープローブは、Primer Expressソフトウェアを使用してカスタム設計した。IGF1RとMAPK14のRNA値は、アクチン(#4326315)のRNAに対して標準化した。IGF1Rタンパク質は、IGF1R特異的モノクローナル抗体(BD Biosciences #555998)とフィコエリトリン結合2次抗体(BD Biosciences #550589)で染色後にフローサイトメトリーにより定量した。星印は、タンパク質レベルを少なくとも60%低下させたIGF1R siRNA2本鎖を示す。MAPK14タンパク質は、MAPK14特異的モノクローナル抗体(BD Biosciences)を用いて細胞溶解物をウェスタンブロットした後、化学発光免疫プロットのKodak画像分析により定量した。MAPK14タンパク質レベルは、アクチンレベルに対して標準化した。誤差棒は、少なくとも3つの独立した実験の標準偏差である。

10

20

【0040】

図2はオフターゲット遺伝子サイレンシングの用量効果を示す。オフターゲット遺伝子サイレンシングは、siRNA濃度の低下により排除されなかった。HeLa細胞を記載の濃度のMAPK14-1 siRNAでトランスフェクトした。トランスフェクションの48時間後にRNAを抽出し、図1に記載のように分析した。

【0041】

図3A~Cは、RNAiによるMAPK14タンパク質とRNAノックダウンの動力学的分析を示す。図3A、HeLa細胞をsiRNA MAPK14-1でトランスフェクション後の記載の時間にタンパク質抽出物を採取し、MAPK14タンパク質の免疫プロット分析に付した。図3B、HeLa細胞をトランスフェクション後の記載の時間にRNA抽出物を採取し、図1に記載のように発現プロファイルのために処理した。siRNAでトランスフェクトした細胞からのRNAを、同じ時点のモックトランスフェクトした細胞からのRNAに対してハイブリダイズした。MAPK14に対する8つのRNAのうちの5つにより共通に制御される遺伝子のセットで集めることにより測定した。各行は、単一の時点での遺伝子発現パターンを示す。データは、2つの独立した実験(1~24時間の短時間コースと24~96時間の長時間コース)をまとめたものである。すなわち、24時間の時点について2つの独立したデータセットがあり、実験の再現性を証明している。棒グラフは、リアルタイムPCRにより定量した、siRNAトランスフェクション後に残存しているMAPK14 RNAの割合である。図3C、(b)のマイクロアレイ分析からの転写体データをトレンドプロットにより分析した。制御した転写体の選択を示す(9つの時点のうちの3つで p 値 < 0.1 、9つの時点のうちの2つで \log_{10} 比 > 0.3 、および9つの時点のうちの2つで \log_{10} 強度 > -1)。異なる時点で一貫しない発現パターンを示したいいくつかの遺伝子を、分析から除去した(NMP200、HIRA、HKE2、EYA4、FLJ20281、LMNB1、EGR1)。残りの転写体を、半最大転写体分解のタイミングに基づき一時的に6つの群に分類した。データは、トランスフェクション後の時間の関数としてプロットした発現比の \log_{10} として示す。第1群: MAPK14; 第2群: KPNB3、RAP2A、FLJ20291、RRAD、RPA2、DKFZp564J157、AF093680、および2つの未解析ESTコンティグ(contig)(図4A参照)。第3群: MGC4809、NCF2、IFI44、コンティグ41538_RC、CBFA2T3、ISG15、LGP2、SCYA5。第4群: H2AFL。第5群: KIAA1460。第6群

30

40

50

: VCP、TIMP4、AL162069、BIRC3、CTGF。

【0042】

図4A~Cは、オフターゲット遺伝子制御への配列類似性の寄与を示す。図4A: MAPK14と類似の動力学で制御される遺伝子の配列。MAPK14配列と完全な同一性を有するヌクレオチドを太字で示し、ミスマッチヌクレオチドを小さい文字で示す。MAPK14転写体との配列同一性の程度は、連続同一ヌクレオチドの数/同一ヌクレオチドの総数として示される。図4B: 動力学的第2群の転写体の遺伝子サイレンシングのクラスター分析。HeLa細胞を、同族のMAPK14 siRNA、または配列類似性の程度を小さくするための単一ヌクレオチド置換を含有するsiRNAでトランスフェクトした。トランスフェクションの48時間後にRNA抽出物を採取した。各行は、迅速にサイレンシングされた転写体の遺伝子発現のパターンに対する単一ヌクレオチド置換の作用を示す。緑色は遺伝子発現の低下を示し、黒は鑑別的制御の欠如を示す (siRNA配列については表1を参照)。図4C: オフターゲット遺伝子KPNB3とFJLJ20291のサイレンシングにより生じる遺伝子発現。HeLa細胞を、オフターゲット遺伝子KPNB3 (上のパネル) およびFLJ20291 (下のパネル) に対応するsiRNA配列でトランスフェクトした。マイクロアレイデータを、発現比のlog10に対する蛍光強度のlog10としてプロットする。標的化した遺伝子ならびにMAPK14を示す。

10

【0043】

図5は、MAPK14とIGF1Rサイレンシングは、明確な遺伝子発現パターンを示す。8つのMAPK14 siRNAと16個のIGF1R siRNAの転写体発現データを、1つの実験群にまとめた。統計的な要件は、少なくとも7つのsiRNAについてp値 < 0.01とlog10発現比 = 0.3であった。点線は、IGF1R発現プロファイルをMAPK14発現プロファイルから分離する。

20

【0044】

図6A~Cは、IGF1R siRNAが受容体発現と機能を妨害することを示す。図6A、IGF1R siRNAでMCF7細胞をトランスフェクションした96時間後に、タンパク質抽出物を採取した。IGF1Rタンパク質を、図1に記載したようにフローサイトメトリーにより定量した。図6B、IGF1R siRNAでMCF7細胞をトランスフェクションした72時間後に、細胞を24時間血清枯渇させ、次に記載の用量のIGFで2時間刺激した。タンパク質抽出物を、このタンパク質のリン酸化型に特異的な抗体を使用してリン酸化AKTの免疫プロット分析に付した (細胞シグナル伝達)。図6C、血清枯渇させIGFで刺激したモックトランスフェクトした細胞またはsiRNAトランスフェクトした細胞からのRNAを、同様にトランスフェクトし枯渇させIGFで刺激しなかった細胞とハイブリダイズさせた。

30

【0045】

図7は、siRNAオリゴによるMAPK14 mRNAのノックダウンに対する単一ヌクレオチドミスマッチの作用を示す。このミスマッチヌクレオチドを下線で示す。

【0046】

図8は、本発明の方法を実施するのに有用なコンピュータシステムの例を示す。

【0047】

図9は、RNA干渉の略図を示す。

【0048】

図10は、3つのsiRNAのsiRNAプールを使用して得られた結果を示す。個々のsiRNAを、33 nM濃度でHeLa細胞中にトランスフェクトした。3つのsiRNAのプールを100nM (それぞれ33nM) でトランスフェクトした。トランスフェクションの24時間後RNAを抽出し、モックトランスフェクトした細胞からのRNAに対してプロファイリングした。右のパネルは、オフターゲット遺伝子サイレンシングがプール中で維持されることを示す。オフターゲットイベントの数は、個々のsiRNAからのオフターゲットイベントのほぼ合計であるが、オフターゲットサイレンシングの強度は低下する。

40

【0049】

図11は、siRNAプールがサイレンシング特異性を上昇させることを示す。単一のsiRNA濃度の低下は、特異性を改善しない (上のパネル)。これに対して、プールサイズを上昇させることはオフターゲットサイレンシングを維持するが、オフターゲット遺伝子サイレン

50

シングの数と強度を低下させる。これは、リスク (RISC) との結合についての siRNA 間の競合のためかも知れない。その結果、オンターゲット：オフターゲット遺伝子サイレンシングの比は上昇し、特異性が上昇する。これは、さらに大きなプールサイズまで上昇させると特異性がさらに上昇することを示唆し、標的評価の試みで非常に有用であろう。

【 0 0 5 0 】

図12は、9つの siRNA の siRNA プールを使用して得られた結果を示す。個々の siRNA の濃度は 11nM であり、プールはそれぞれ 11nM の 9 つのメンバーを含有した。siRNA プール中の siRNA の数が増加すると、おそらくリスク (RISC) との結合についての siRNA 間の競合のために、プール中のシグナチャ (signature) 遺伝子の数が低下した。その結果、オフターゲットサイレンシングの強度が低下し、オンターゲット：オフターゲット遺伝子サイレンシングの比率が上昇し、特異性が上昇した。

10

【 0 0 5 1 】

図13A~Bは、siRNA とその逆相補体の GC 含量と悪い siRNA (bad siRNA) の GC 含量との比較を示し、悪い siRNA が良い siRNA (good siRNA) と同様のセンス鎖を有し、良い siRNA が悪い siRNA と同様のセンス鎖を有することを示す。RC: siRNA 標的配列の逆相補体。

【 0 0 5 2 】

図14は、あまり有効でない siRNA が活性なセンス鎖を有することを示す。61個の siRNA の鎖の偏りは、3' 偏り法 (3'-biased method) による発現プロファイル、siRNA の GC PSSM スコアの比較、およびこれらの逆相補体から予測された。鎖の偏り予測は、siRNA サイレncing 効率により区分けした。

20

【 0 0 5 3 】

図15は、センス鎖修飾が、「センス活性」 siRNA の特異性を変化させることを示す。100 nM の濃度で HeLa 細胞中にトランスフェクトされた siRNA。トランスフェクションの 12 時間後に RNA を抽出し、モックトランスフェクトした細胞からの RNA に対してプロファイリングした。オンターゲット siRNA はセンス鎖中に 2 つの 2'-O-メチル置換を含有して、この鎖を不活性化し、これが遺伝子サイレンシングで機能することを妨害する。以前はアンチセンス活性であると測定された siRNA 71 のシグナチャは、センス鎖不活性化により変化しなかった。これに対して siRNA 73 のシグナチャ (両方の鎖が活性である) は、センス鎖不活性化により大きく変化した。この状況で、センス鎖は遺伝子サイレンシングで機能することが妨害され、リスク (RISC) と相互作用することが妨害され、従ってアンチセンス鎖が優勢になった。この結果は、センス鎖の化学不活性化が、センス活性 siRNA のオフターゲットシグナチャを変化させることができるが、それを排除しないことを示す。

30

【 0 0 5 4 】

図16はオフターゲット制御における鎖の偏りを示す。siRNA をシグナチャ遺伝子とアラインメントさせた。センス鎖またはアンチセンス鎖とのアラインメントに対する偏りを、連続的アラインメント長の比較により測定した。ランダムセットの予測されるスコアを、灰色で示す。siRNA 73 について、非修飾 siRNA のシグナチャはアンチセンス鎖との同一性を示し、センス鎖によるサイレンシングを示している。この siRNA のセンス不活性化型では、シグナチャ遺伝子は、センス鎖との同一性を示し、アンチセンス鎖の活性によるサイレンシングを示している。これは両方の鎖が siRNA 73 について機能性であり、センス鎖が優勢であることを証明する。この鎖の不活性化は、アンチセンス鎖が優勢になることを可能にした。siRNA 71 について、シグナチャ遺伝子はセンス鎖との同一性を示し、アンチセンス鎖がこの 2 本鎖中では優先的に活性であり、センス鎖の化学的不活性化がこのシグナチャを変化させないことを証明している。

40

【 0 0 5 5 】

図17は、センス鎖修飾が「センス活性」 siRNA の力価を上昇させることを示す。siRNA 力価曲線は、より低い濃度でサイレンシング効率を維持する能力として定義される力価を評価するのに使用することができる。siRNA を記載の濃度で HeLa 細胞中にトランスフェクトさせた。トランスフェクションの 24 時間後に RNA を抽出し、オンターゲットサイレンシングをリアルタイム PCR により測定した。センス活性 siRNA (黒の実線) は、アンチセンス活

50

性siRNA（灰色の実線）と比較して力価が低下していることを示す。このアッセイはオンターゲットサイレンシングを測定し、これはアンチセンス鎖の活性の関数である。2本鎖73中のセンス鎖の優勢な活性は、アンチセンス鎖の活性に干渉し、こうして力価と最大効率の両方を制限する。この2本鎖のセンス鎖の化学的不活性化は、おそらくアンチセンス鎖との結合のためにリスク（RISC）を解放することにより、力価と最大効率（黒の点線）の両方を上昇させる。これは、センス鎖の不活性化が、より有効な標的遺伝子サイレンシングについてより強力なsiRNAを達成するための1つの機構であることを示唆する。

【0056】

5. 発明の詳細な説明

本発明は、RNA干渉による遺伝子サイレンシングのための方法と組成物を提供する。好ましくはこの方法と組成物は、生物の内因性遺伝子をサイレンシングするために使用される。特に本発明は、標的遺伝子転写体とわずかに部分的な配列相同性を有する小さいかまたは短い干渉性RNA（siRNA）（例えば、転写体の配列と同一のヌクレオチド配列のセンス鎖またはアンチセンス鎖中央領域を含むsiRNA、および転写体の配列と同一のセンス鎖またはアンチセンス鎖3'ヌクレオチド配列を有するsiRNA）を使用する遺伝子サイレンシングの方法を提供する。本出願では、siRNA鎖の配列と目的の遺伝子の配列の比較がしばしば行われる。この比較は、siRNAの特定の鎖と遺伝子の転写体との配列比較を意味することを理解されたい。本出願では、siRNAはまた遺伝子を標的とすると言われる。この説明がされる場合、siRNAは遺伝子の転写体を標的としこれの分解を引き起こすことを意味することを理解されたい。本出願では、siRNA鎖中のヌクレオチドまたはヌクレオチド配列の位置は、しばしばsiRNAの3'末端を参照して説明される。この説明が使用される時、3'オーバーハングの2つのヌクレオチドはヌクレオチドの番号付けに含まれず、すなわち3'末端のヌクレオチドの番号付けは、siRNAの2本鎖部分の最初のヌクレオチドで始まることを理解されたい。本発明はまた、遺伝子を標的化するように設計された複数の異なる小干渉性RNAにより通常制御される遺伝子を同定するための方法を提供する。本方法において、その発現が複数の異なるsiRNAによりサイレンシングされる1つ以上の遺伝子は、これらのsiRNAに対する測定された応答プロファイルに基づいて同定される。本発明はまた、それぞれが標的遺伝子中に異なる配列を標的化するsiRNAのプールを使用する遺伝子サイレンシングの方法を提供する。本発明はまた、鎖の優先、すなわち遺伝子サイレンシングにおけるsiRNAの2つの鎖の相対活性を測定する方法を提供する。本発明はさらに、遺伝子サイレンシングのためにsiRNAを設計する方法を提供する。

【0057】

本発明は少なくとも部分的に、遺伝子の転写体と同一の、例えば約11~14ヌクレオチドのヌクレオチド配列の中央領域、または約9ヌクレオチドの3'末端ヌクレオチド配列は、siRNAが遺伝子の発現をサイレンシングするのに充分であるという本発明者らによる発見に基づく。従って特定の遺伝子転写体を標的とするように設計されたsiRNAは、siRNA中の適切なサブ配列が他の遺伝子の転写体中の配列と同一なら、この1つ以上の他の遺伝子（標的遺伝子以外）の発現を改変することがある。さらに、部分的または完全な配列同一性で標的遺伝子を標的としない鎖（例えば、センス同一標的を標的とするsiRNAのセンス鎖、またはアンチセンス同一標的を標的とするアンチセンス鎖）もまた、1つ以上の遺伝子の発現を改変することがある。すなわち各siRNAは、細胞中の明確な遺伝子発現パターンを与えることがある。siRNAの遺伝子発現パターンは、siRNAに付された細胞を発現プロファイリングすることにより決定することができる。また、その発現がsiRNAの導入により影響を受ける遺伝子は、その応答の動力学に基づいて異なる群に分類することができる。例えば、ある群の遺伝子について、siRNAの導入後まもなく、転写体レベルが低下することがわかっている。これらの群は、siRNAにより直接サイレンシングされる遺伝子を含む。ある群の遺伝子については、十分に長期間が経過するまで転写体レベルは低下しない。これらの群は、siRNAにより標的化される遺伝子によりコードされるタンパク質の喪失の結果としてその転写体レベルが改変される遺伝子を含む。別に、特定の遺伝子を標的化する複数の異なるsiRNAを使用して、複数の異なるsiRNAにより通常制御される1つ以

上の遺伝子のセットを、その応答プロファイルに基づいて同定することができる。

【0058】

本出願では便宜上、本発明の方法と組成物を例示するために、21ヌクレオチドからなるsiRNAがしばしば使用される。より長いsiRNA(例えば22または23ヌクレオチドのsiRNA)が同様に応用可能であり、本発明に包含されることは理解されるであろう。本出願では、siRNAは小さいかまたは短い干渉性RNAを意味する。適切な配列を有する小さいヘアピンRNA(shRNA)をsiRNAの代わりに使用できることは、当業者により理解されるであろう(セクション5.5を参照、下記)。このshRNAは、本発明に包含されると企図される。本出願では、細胞は培養細胞またはin vivoの細胞でもよい。

【0059】

本出願では、siRNAのいずれかの鎖が遺伝子サイレンシングのガイド鎖として作用する。本明細書において便宜上siRNAの2つの鎖を、サイレンシングされるすべての標的転写配列の方向(5'から3'または3'から5')を記載せずに、センス鎖またはアンチセンス鎖と呼ぶ。一般に本発明におけるセンスまたはアンチセンスのsiRNA鎖の記載は、siRNAが標的とする選択された転写配列(参照標的配列)の配列の方向に関してである。すなわち、siRNAのアンチセンス鎖は、この参照標的配列に対するガイド鎖を意味し、一方センス鎖は、この参照標的配列と「同一」であるかまたは非ガイド鎖を意味する。siRNAが標的とすることができる任意の転写体配列は、参照標的配列として使用することができる。好適な実施形態において参照標的配列は、siRNAがサイレンシングするはずの目的とする標的配列である。この鎖の記載の慣習ではsiRNAのセンス鎖は、同じかまたは異なる転写体中の他の配列の標的化においてガイド鎖として作用することがある。

【0060】

5.1. 生物学的状態と発現プロファイル

細胞または他の生物学的サンプルの状態は、セクション5.1.1(後述)で規定される細胞成分(任意の測定可能な生物学的変数)で示される。これらの細胞成分は、摂動(perturbation)に反応してまたは異なる条件下で変化する。測定されたシグナルは、この細胞成分の測定値または細胞成分の応答の測定値でもよい。

【0061】

5.1.1 生物学的状態

本明細書において「生物学的サンプル」は、任意の細胞、組織、臓器または多細胞生物を含むと、広く定義される。生物学的サンプルは例えば、in vitroの細胞または組織培養物由来でもよい。あるいは生物学的サンプルは、生きた生物または単一細胞生物の集団に由来してもよい。好適な実施形態において生物学的サンプルは、生きた細胞または生物を含む。

【0062】

生物学的サンプルの状態は、その細胞成分の含量、活性または構造により測定することができる。本明細書において生物学的サンプルの状態は、特に限定されないがsiRNAまたは他の変化の影響の解析を含む目的について、細胞または生物を解析するのに十分な細胞成分の集団の状態から得られる。用語「細胞成分」はまた本開示において、任意の種類の測定可能な生物学的変数を包含すると広く定義される。これらの成分の状態についての測定値および/または観察結果は、例えばmRNAもしくはタンパク質のこれらの豊富さ(すなわち、生物学的サンプル中の量または濃度)、またはこれらの活性、またはこれらの修飾状態(例えばリン酸化)、または生物学的サンプルの生物学に関連する他の測定値でもよい。種々の実施形態において本発明は、細胞成分の異なる集団についてこの測定値および/または観察結果を作成することを含む。これらの異なる細胞成分の集団はまた本明細書において、生物学的サンプルの生物学的状態の態様と呼ばれる。

【0063】

本発明において有用に測定される生物学的サンプル(例えば細胞または細胞培養物)の生物学的状態の1つの態様は、その転写状態である。実際転写状態は、本発明で測定される生物学的状態の好適な態様である。生物学的サンプルの転写状態には、ある条件下での

10

20

30

40

50

細胞中の構成RNA種（特にmRNA）の本体および量を含む。好ましくは生物学的サンプル中の構成RNA種のかなりの割合が測定されるが、siRNAの作用または目的の他の摂動を解析するのに少なくとも十分な割合が測定される。生物学的サンプルの転写状態は、例えばいくつかの既存の遺伝子発現法の任意の方法を使用してcDNAの量を測定することにより決定することが便利である。本発明の1つの特に好適な実施形態は、多数の遺伝子のmRNAまたは転写体レベルを測定するためにDNAアレイを使用する。本発明の他の好適な実施形態は、ある生物のゲノム中の多数の遺伝子またはエキソンの発現レベルを測定するためにDNAアレイを使用する。

【0064】

本発明で有用に測定される生物学的サンプルの生物学的状態の他の態様は、転写状態である。生物学的サンプルの転写状態には、ある条件下での生物学的サンプル中の構成タンパク質種の本体および量を含む。好ましくは生物学的サンプル中の構成タンパク質種のかなりの割合が測定されるが、siRNAの作用または目的の他の摂動を解析するのに少なくとも十分な割合が測定される。当業者には公知のように、転写状態はしばしば翻訳状態を示す。

【0065】

生物学的サンプルの生物学的状態のさらに別の態様は、小分子状態（例えば代謝状態）である。生物学的サンプルの小分子状態は、細胞中に存在する小分子の本体および量を含む。小分子は、分子量が約5000未満の分子を意味し、特に限定されないが、脂肪酸、アミノ酸、ヌクレオチド、細胞プロセスの中間体（例えば代謝経路およびシグナル伝達経路の中間体）がある。

【0066】

生物学的サンプルの生物学的状態の他の態様もまた、本発明において有用である。例えば本明細書で使用される用語の生物学的サンプルの活性状態は、ある条件下の生物学的サンプル中の成分タンパク質種（およびまた、随時触媒活性のある核酸種）の活性を含む。当業者には公知のように、転写状態はしばしば活性状態を示す。

【0067】

本発明はまた、適宜生物学的サンプルの生物学的状態の「混合」状態に適用され、生物学的サンプルの生物学的状態の異なる態様の測定値が組合わされる。例えばある混合態様において、あるRNA種およびあるタンパク質種の量は、いくつかの他のタンパク質種の活性の測定値と組合わされる。さらに本発明はまた、測定可能な生物学的サンプルの生物学的状態の他の態様に適用可能であることが、以下から理解されるであろう。

【0068】

生物学的サンプル（例えば細胞または細胞培養物）の生物学的状態は、いくつかの数の細胞成分のプロファイルで示される。細胞成分のこのプロファイルは、ベクトルSで示される

$$S = (S_1, \dots, S_i, \dots, S_k) \quad (1)$$

ここで、 S_i は、 i 番目の細胞成分のレベル（例えば遺伝子 i の転写体レベル）であるか、またはタンパク質 i の量もしくは活性レベルである。好適な実施形態において、 k は2より大きく、好ましくは10より大きく、さらに好ましくは100より大きく、さらに好ましくは1000より大きく、さらに好ましくは10,000より大きく、さらに好ましくは25,000より大きく、さらに好ましくは50,000より大きく、最も好ましくは100,000より大きい。

【0069】

ある実施形態において細胞成分は連続的変量として測定される。例えば転写速度は、代表的には単位時間に合成される分子の数として測定される。転写速度はまた、対照速度のパーセントとして測定してもよい。しかし他のある実施形態では、細胞成分は範疇的変量として測定される。例えば転写速度は、「オン」または「オフ」として測定され、ここで値「オン」は所定の閾値より大きい転写速度を示し、「オフ」は、閾値より小さい転写速度を示す。

【0070】

5.1.2 生物学的応答と発現プロファイル

摂動（すなわち、例えばsiRNAの適用のようなある条件）に対する生物学的サンプルの応答は、生物学的サンプルの生物学的状態の変化を観察することにより測定することができる。例えば生物学的サンプルの応答は、摂動（例えばsiRNAの適用、遺伝子突然変異、環境変化など）に対する生きた細胞または生物の応答でもよい。応答プロファイルは、細胞成分の変化の集まりである。本発明において、摂動 m に対する生物学的サンプル（例えば細胞または細胞培養物）の応答プロファイルは、ベクトル $^{(m)}$ で示される

$$^{(m)} = (\quad_1^{(m)}, \dots, \quad_i^{(m)}, \dots, \quad_k^{(m)}) \quad (2)$$

ここで、 $\quad_i^{(m)}$ は、摂動 m 下での細胞成分 i の応答の強度である。本発明のある特に好適な実施形態において、siRNA、薬物、薬物候補の適用、または任意の他の摂動に対する生物学的応答は、少なくとも2つの遺伝子および/またはタンパク質、好ましくは10個より多い遺伝子および/またはタンパク質、さらに好ましくは100個より多い遺伝子および/またはタンパク質、さらに好ましくは1000個より多い遺伝子および/またはタンパク質、さらに好ましくは10,000個より多い遺伝子および/またはタンパク質、さらに好ましくは50,000個より多い遺伝子および/またはタンパク質、さらに好ましくは100,000個より多い遺伝子および/またはタンパク質の転写体レベルの誘導された変化により測定される。

10

【0071】

本発明のある実施形態において応答は単に、摂動の前と後の生物学的変量の差である。ある好適な実施形態において応答は、摂動が適用される前と後の細胞成分の比として定義される。他の実施形態において応答は、摂動後の時間の関数、すなわち $^{(m)} = \quad^{(m)}(t)$ である。例えば $\quad^{(m)}(t)$ は、摂動前と摂動後の時間 t における細胞成分の差または比である。

20

【0072】

ある好適な実施形態において、もし遺伝子 i の応答がある閾値強度または測定誤差挙動の知識から決定される信頼レベルより低い場合、 $\quad_i^{(m)}$ はゼロに設定される。この実施形態において、その測定応答が閾値より小さい細胞成分は応答値ゼロが与えられ、一方その測定応答が閾値より大きい細胞成分はその測定応答値が保持される。この応答ベクトルの切り捨ては、小さい応答のほとんどが測定誤差であると予測される場合は、優れた方法である。切り捨て後、応答ベクトル $^{(m)}$ はまた、同様の摂動の存在について「一致検出物 (matched detector)」に近づく（例えば、Van Trees, 1968, 「検出、推定、および調節理論 (Detection, Estimation, and Modulation Theory)」、Vol. I, Wiley & Sons参照）。切り捨てレベルは検出目的と測定誤差に基づいて設定できることは、当業者に公知であろう。例えばある実施形態において、その転写体レベルの変化が2倍より小さいかまたはさらに好ましくは4倍より小さい遺伝子は、ゼロの値が与えられる。

30

【0073】

ある好適な実施形態において、摂動はいくつかの強度レベルで適用される。例えば、その応答を観察するためにsiRNAの異なる量が生物学的サンプルに適用される。この実施形態において摂動応答は、摂動強度 u の単一のパラメータ化「モデル」関数によりそれぞれを近似することにより内挿される。転写状態データを近似するのに適したモデル関数の例は、ヒル (Hill) 関数であり、これは調整可能なパラメータ a 、 u_0 、および n を有する。

40

【0074】

$$H(u) = \frac{a(u/u_0)^n}{1+(u/u_0)^n} \quad (3)$$

【0075】

調整可能なパラメータは、摂動応答の各細胞成分について独立に選択される。好ましくは調整可能なパラメータは、各細胞成分についてモデル関数（例えば、ヒル関数、式3）と各摂動強度の対応する実験データとの差の平方和が最小になるように選択される。この

50

好適パラメータ調整法は、最小二乗適合として当該分野で公知である。他の可能なモデル関数は、例えば種々の既知のクラスの多項式による多項適合に基づく。モデル適合と生物学的応答のさらに詳細な説明は、FriendとStoughtonのMethods of Determining Protein Activity Levels Using Gene Expression Profiles、PCT公報W099/59037（これは参照することによりその全体が本明細書に組み込まれる）に開示されている。

【0076】

5.2. 遺伝子を標的化する複数の小干渉性RNAに対する共通の応答を同定する方法

本発明は、標的遺伝子をサイレンシングするように設計された複数の異なる小干渉性RNAにより共通にサイレンシングされ、かつ標的遺伝子とは異なる1つ以上の遺伝子を同定する方法を提供する。この1つ以上の遺伝子はまた、「非標的」または「オフターゲット」遺伝子と呼ばれる。本発明において非標的遺伝子は、複数のsiRNAに対する測定された応答プロファイルに基づいて同定することができる。

10

【0077】

5.2.1. 小干渉性RNAの応答プロファイル

本発明は、siRNAの応答プロファイルを測定する方法を提供する。この方法では、siRNAによる摂動に付された細胞中の複数の遺伝子のmRNAレベルおよび/またはコードされるタンパク質のレベルが測定される。好ましくはsiRNAの応答プロファイルは、細胞中のsiRNA分子の導入後の選択された時点で測定される。本発明のある実施形態において、siRNAに対する応答速度が測定される時、siRNA導入後の複数の時点で複数の応答プロファイルを測定することができる。

20

【0078】

siRNAによる摂動に付された細胞中のmRNAレベルおよび/またはタンパク質のレベルは、当該分野で公知の任意の方法により測定することができる。好ましくは細胞中のすべての成分種のかなりの割合が測定される。ある好適な実施形態においてmRNAレベルはDNAマイクロアレイを使用して測定される。好ましくは、5より大きい、10より大きい、100より大きい、1000より大きい、10,000より大きい、25,000より大きい、50,000より大きい、または100,000より大きいmRNA種が、siRNAの応答プロファイルについて測定される。別の好適な実施形態においてタンパク質レベルは、2Dゲル電気泳動またはタンパク質アレイにより測定される。好ましくは5より大きい、10より大きい、100より大きい、1000より大きい、10,000より大きい、25,000より大きい、50,000より大きい、または100,000より大きいタンパク質種が、siRNAの応答プロファイルについて測定される。

30

【0079】

例として標準的選択規則（例えば、Elbashirら、2002、Methods 26:199-213参照）に従って、IGF1RまたはMAPK14のコード領域を標的とする21ヌクレオチドの2本鎖siRNAを設計した。IGF1Rを標的とする全部で16個のsiRNAとMAPK14を標的とする8個のsiRNA（表1）を設計した。これらのsiRNAは、個々にHeLa細胞中にトランスフェクトした。異なるsiRNAによる同じ標的遺伝子のサイレンシングにより生じる発現プロファイルが得られた。発現プロファイルは、21,000個のヒト遺伝子に対応するオリゴヌクレオチドを含有するマイクロアレイを使用して測定された、約21,000個のヒト遺伝子の測定された転写体レベルを含有した。これらのsiRNAの測定された応答プロファイルの一部を図1Aと1Bに示す。MAPK14を標的とする8個のsiRNA2本鎖のそれぞれは、明確な発現パターンを与えた（図1A）。同様にIGF1Rに対する16個のsiRNA2本鎖のそれぞれは、ユニークな発現パターンを与えた（図1B）。3つの異なる独立した実験で、実質的に同一の遺伝子発現パターンが得られ、特定のsiRNAから生じる遺伝子制御が再現性あることを証明している。これらの結果は、同じ標的転写体に対する異なるsiRNAが、少数の共通の遺伝子制御を誘発するが、転写体発現パターンの大部分は標的的特異的であるというよりsiRNA特異的であった。

40

【0080】

改変された転写体の数と本体は、標的遺伝子以外の遺伝子を含む。MAPK14 siRNA 2本鎖のすべてが標的を有効にサイレンシングし、RNAとタンパク質レベルの80%を超える低下を示しているが、各siRNAはまた、MAPK14以外の異なる数の転写体の発現を制御する（図1

50

A)。16個のIGF1R siRNAのうち11個はIGF1Rタンパク質レベルを約60~80%低下させた(図1B)。2つのIGF1R siRNAはIGF1Rタンパク質とRNAレベルを再現性よく上昇させた。IGF1R-4はIGF1Rの発現を80%低下させるが、発現が改変された遺伝子はIGF1R-5より少なく、標的のわずかに30%のサイレンシングを引き起こしたのみであった。さらにルシフェラーゼを標的とするsiRNAは、ヒトゲノム中に相同的標的が無いにもかかわらずいくつかの遺伝子の発現を再現性よく制御し、ルシフェラーゼsiRNAにより制御された転写体の数は、同じ対照についての数より大きかった。すなわち発現プロファイルから、遺伝子制御の観察されたパターンは目的の標的ではなく、サイレンシングに利用されたsiRNA配列に特異的であることが証明される。

【0081】

従って本発明はまた、真核細胞に対するsiRNAの作用を測定する方法を提供する。この方法は、細胞へのsiRNA導入後の1つ以上の選択された時点で真核細胞の発現プロファイルを測定することを含む。

【0082】

5.2.2. 1つ以上の小干渉性RNAの共通の応答パターンを同定する方法

1つ以上のsiRNAに対する細胞の共通の応答パターンは、1つ以上のsiRNAに対する細胞の応答プロファイルを使用して同定することができる。好適な実施形態において単一のsiRNAにより共通にサイレンシングされる遺伝子は、siRNAの応答プロファイル動力学を分析することにより同定される。ある好適な実施形態において、複数の異なるsiRNAにより共通にサイレンシングされる遺伝子は、複数の異なるsiRNAの複数の応答プロファイルを分析することにより同定される。

【0083】

5.2.2.1. 動力学データを使用して小干渉性RNAにより直接サイレンシングされる遺伝子を同定する方法

本発明は、siRNAにより直接サイレンシングされる遺伝子、すなわちsiRNAによりサイレンシングされる1つ以上の他の遺伝子によりコードされる1つ以上のタンパク質の喪失のためではなく、siRNAの直接の作用のためにそのmRNAレベルが変化する遺伝子を同定する方法を提供する。

【0084】

好適な実施形態において、遺伝子のサイレンシングのsiRNA濃度への依存が測定され分析される。好適な実施形態において、遺伝子(例えば標的遺伝子または非標的遺伝子)のmRNAレベルおよび/またはコードされるタンパク質レベルが、siRNAの複数の異なる濃度で測定される。次に遺伝子のサイレンシングの濃度依存性の特徴が、非標的遺伝子と標的遺伝子サイレンシングとの間で比較される。ある実施形態において、濃度依存性を解析するためにsiRNA濃度についての半最大応答(half-maximal response)が使用される。ある実施形態において、オフターゲット遺伝子制御と目的の標的のサイレンシングの間に差が無い場合、オフターゲット遺伝子制御が単に高siRNA濃度のアーチファクトではないことを示す。例としてMAPK14タンパク質とsiRNA MAPK14-1によるRNAノックダウンの濃度分析を行った。siRNA濃度を1000倍低下させた時、標的遺伝子サイレンシングが検出されたが、オフターゲット遺伝子制御もまた検出された(図2)。これらの遺伝子の多くは、MAPK14としてsiRNA濃度についてほとんど同一の半最大応答を示す(~1nM)。目的の標的のサイレンシングからオフターゲット遺伝子制御を力価測定することができなかつたという事実は、オフターゲット遺伝子制御が単に高siRNA濃度のアーチファクトではないことを示す。

【0085】

別の好適な実施形態において、遺伝子のサイレンシングの動力学が測定され分析される。好適な実施形態において、遺伝子(例えば標的遺伝子または非標的遺伝子)のmRNAレベルおよび/またはコードされるタンパク質レベルが、細胞へのsiRNAの導入後の複数の異なる時点で測定される。次に遺伝子の転写体レベルの一時的プロファイルの特徴を使用して、遺伝子がsiRNAにより直接に、または例えばsiRNAによりサイレンシングされる遺伝子

10

20

30

40

50

によりコードされるタンパク質の喪失の結果としての2次的作用を介してサイレンシングされるかどうかを調べる。ある実施形態において、時間依存性を解析するために、mRNAの半最大分解の時間が使用される。siRNAにサイレンシングされる遺伝子によりコードされるタンパク質の寿命より速い速度で、例えばタンパク質の半減期より短い半最大分解の時間で非標的遺伝子をサイレンシングすることは、遺伝子が標的タンパク質の喪失によるより、直接サイレンシングされることを示す。

【0086】

発現プロファイルの一時的プロファイルは、当該分野で公知の任意の標準的方法を使用して測定することができる。本発明のある実施形態において、細胞へのsiRNAの送達後の複数の異なる時間で複数の遺伝子の発現レベルが測定される。この実施形態において、発現レベルは最も好ましくは、0時間、すなわちRNA導入直後から、標的遺伝子によりコードされるタンパク質の喪失による下流の遺伝子の発現レベルへの影響が観察されるのに十分な長い時点までの範囲にわたって測定される。好ましくは複数の時点は、0時間から標的遺伝子にコードされるタンパク質のレベルがその摂動の無いレベルの約50%まで低下する時点までの範囲にわたる時点を含む。ある実施形態において複数の時間は、標的遺伝子によりコードされるタンパク質のレベルがその摂動の無いレベルの約10%、20%、30%、40%、50%まで低下する時点を含む。標的遺伝子によりコードされるタンパク質のレベルがその摂動の無いレベルの約50%まで低下する時点は、タンパク質の「半減期」と呼ばれる。さらに好ましくは発現レベルは、標的遺伝子によりコードされるタンパク質の半減期の約2倍長い時間で測定される。例えばMAPK14タンパク質は、約40時間の半減期を有する。すなわちMAPK14が標的遺伝子である実施形態において、代表的なハイブリダイゼーション時間は、約0~40時間、さらに好ましくは約0~80時間である。好適な実施形態において標的タンパク質の半減期は、発現レベルの測定と同時に測定される。例えば発現プロファイルが測定される各時点で、標的タンパク質の量もまた測定される。

【0087】

ある実施形態において異なる時点での複数の遺伝子の発現プロファイルが、異なる同一のマイクロアレイ上で別々に測定される。この各測定について発現レベルが測定される時点で、短時間、好ましくは室温で、高~中程度の塩濃度(例えば0.5~3M塩濃度)の水溶液中で、すべての結合またはハイブリダイズしたポリヌクレオチドを保持しすべての非結合ポリヌクレオチドを除去する条件下で、マイクロアレイが洗浄される。次に、各プローブ上の残存するハイブリダイズしたポリヌクレオチド分子上の検出可能な標識物が、使用される具体的な標識法に適した方法により測定される。次に、得られる発現レベルを組合せて発現動力学曲線を生成する。別の実施形態において発現レベルは、単一のマイクロアレイを使用してリアルタイムで測定される。この実施形態においてマイクロアレイは、妨害無しでサンプルにハイブリダイズさせられ、マイクロアレイは各時点で非侵襲的に信号を送る。さらに別の実施形態において、1つのアレイを使用し、短時間ハイブリダイズさせ、洗浄し、ハイブリダイゼーションレベルを測定し、同じサンプルに戻り、別の期間ハイブリダイズさせ、洗浄し、再度測定して発現動力学曲線を得る。

【0088】

例として、一時的遺伝子発現パターンを測定し解析した。MAPK14タンパク質は、siRNAトランスフェクション後に約40時間の半減期を示した(図3A)。これに対してMAPK14転写体は急速に分解され、トランスフェクションの11時間後に半最大分解を示した(図3B)。発現プロファイリングにより、MAPKタンパク質の観察可能な低下の充分前の早い時点(6~12時間)で複数の非標的遺伝子の遺伝子制御が観察された。さらに分析すると、発現シグナチャは、いくつかの一時的な明確な転写体群に分類できた(図3C)。第1群は、単一の転写体の目的標的MAPK14を含有した。第2群は、MAPK14に対して同様のサイレンシング動力学(マイクロアレイにより測定すると半最大分解は7~13時間)を示す9つの転写体を含有した。同じ群の転写体は、別の実験で迅速な動力学でダウンレギュレートされ、これらの遺伝子がsiRNAにより再現性よくサイレンシングされることを証明している。転写体制御の迅速な速度は、これらが直接の転写体分解イベントであることを示唆した。これは

、動力学第3群と第4群（半最大分解は約40時間で起き、従って2次遺伝子発現変化である可能性がある）とは反対である。

【0089】

従って本発明はまた、真核細胞に対するsiRNAの作用を測定する方法であって、細胞へのsiRNAの導入の複数の異なる時間に複数の異なる遺伝子の転写体レベルの測定値を含む真核細胞の発現プロファイルを測定し、複数の異なる遺伝子を該測定された転写体レベルの動学的挙動に基づいて異なる動力学群に分類することを含む方法を提供する。この方法は、siRNAにより直接サイレンシングされる遺伝子と、直接サイレンシングされる遺伝子によりコードされるタンパク質の喪失により影響を受ける遺伝子とを区別するのに使用することができる。

10

【0090】

本発明はまた、標的遺伝子をサイレンシングするように設計されたsiRNAの応答プロファイルの動学的挙動に基づいて、標的遺伝子の機能（標的遺伝子により制御される下流の遺伝子）を同定する方法を提供する。好ましくはsiRNAは、標的遺伝子をサイレンシングするが1つ以上の他の遺伝子はサイレンシングしないように設計される。ある実施形態において本発明は、真核細胞中の第1タンパク質により制御されるが第2タンパク質によっては制御されない1つ以上の遺伝子を同定する方法であって、(a)該真核細胞中に、(i)第1タンパク質をコードする第1遺伝子の転写体配列と同一の11~18ヌクレオチドのセンス鎖またはアンチセンス鎖の中央の連続的ヌクレオチド配列、または(ii)前記第1遺伝子の転写体配列と同一の9~18ヌクレオチドの3'センス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含み、かつ(iii)第2タンパク質をコードする第2遺伝子の転写体配列と同一の11~18ヌクレオチドのセンス鎖またはアンチセンス鎖の中央の連続的ヌクレオチド配列を含まず、かつ(iv)第2遺伝子の転写体配列と同一の9~18ヌクレオチドの3'センス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含まないsiRNA分子を導入し；(b)前記siRNAの導入のある期間後に該真核細胞の発現プロファイルを測定し（ここで、各発現プロファイルは複数の異なる遺伝子の測定された転写体レベルを含み、その期間は該第1タンパク質の寿命より長い）；そして(c)その転写体レベルが、前記siRNAに付されていない細胞中のレベルからは変化している1つ以上の遺伝子を同定することを含む、方法を提供する。

20

【0091】

5.2.2.2. 共通の応答パターンを同定する方法

本発明は、1つ以上の遺伝子を標的とする複数のsiRNAに対する共通の応答パターンを同定する方法を提供する。複数のsiRNAは、3、5、8、10、16、50、100、1,000、またはそれ以上のsiRNAを含むことができる。遺伝子を標的とする複数のsiRNAの応答プロファイルセットの測定された発現レベルは、これらのsiRNAによる摂動にตอบสนองして同時変化する傾向に従って同時変化セットに分類することができる。こうして同定される遺伝子のセットは、複数のsiRNAに対する細胞の共通応答である。遺伝子の発現レベルはまた、コードされるタンパク質の量を分析して測定することができる。

30

【0092】

ある好適な実施形態において共通の応答は、細胞にsiRNAまたはsiRNAの前駆体（例えば、siRNAをコードするプラスミド）を導入後の維持時間に測定された応答プロファイルに基づいて測定される。ある実施形態において共通の応答は、早い時点で測定された応答プロファイルに基づいて測定される。早い時点は、標的遺伝子によりコードされるタンパク質のレベルが、非摂動レベル（例えば、siRNAに付されていない細胞中のタンパク質のレベル）の約50%まで低下していない時点を意味する。ある実施形態において共通の応答は、標的遺伝子によりコードされるタンパク質のレベルが、非摂動レベルの約10%まで低下していない時点で測定される応答プロファイルに基づいて測定される。別の実施形態において共通の応答は、標的遺伝子によりコードされるタンパク質のレベルが、非摂動レベルの約20%まで低下していない時点で測定される応答プロファイルに基づいて測定される。さらに別の実施形態において共通の応答は、標的遺伝子によりコードされるタンパク質の

40

50

レベルが、非摂動レベルの約30%まで低下していない時点で測定される応答プロファイルに基づいて測定される。さらに別の実施形態において共通の応答は、標的遺伝子によりコードされるタンパク質のレベルが、非摂動レベルの約40%まで低下していない時点で測定される応答プロファイルに基づいて測定される。早い時点でsiRNAにより共通にサイレンシングされる遺伝子は、しばしばsiRNAにより直接サイレンシングされる遺伝子である。

【0093】

別の実施形態において共通の応答は、遅い時点で測定された応答プロファイルに基づいて測定される。遅い時点は、標的遺伝子によりコードされるタンパク質のレベルが、非摂動レベルの少なくとも約50%まで低下する時点の意味する。ある実施形態において共通の応答は、標的遺伝子によりコードされるタンパク質のレベルが、非摂動レベルの約50%まで低下する時点で測定される応答プロファイルに基づいて測定される。別の実施形態において共通の応答は、標的遺伝子によりコードされるタンパク質のレベルが、非摂動レベルの約60%まで低下する時点で測定される応答プロファイルに基づいて測定される。さらに別の実施形態において共通の応答は、標的遺伝子によりコードされるタンパク質のレベルが、非摂動レベルの約75%まで低下する時点で測定される応答プロファイルに基づいて測定される。さらに別の実施形態において共通の応答は、標的遺伝子によりコードされるタンパク質のレベルが、非摂動レベルの約90%まで低下する時点で測定される応答プロファイルに基づいて測定される。さらに別の実施形態において共通の応答は、標的遺伝子によりコードされるタンパク質のレベルが、非摂動レベルの約50%まで低下する時点の2倍長い時点で測定される応答プロファイルに基づいて測定される。

【0094】

同時変動セット中の測定された発現レベルを分類し共通の応答を同定する方法は、当該分野で公知である。例えばUS Patent No. 6,203,987および米国特許出願No. 09/220,142（いずれも参照することによりその全体が本明細書に組み込まれる）を参照されたい。好ましくは本発明の同時変動セットは、クラスタリングアルゴリズム（すなわち「クラスタリング分析」）により同定される。

【0095】

本発明で使用されるクラスタリング法およびアルゴリズムは、両方とも「階層制」または「固定数の群」のアルゴリズムを含む（例えば、S-Plus Guide to Statistical and Mathematical Analysis v.3.3, 195, MathSoft, Inc.: StatSci. Division, ワシントン州シアトル、を参照）。かかるアルゴリズムは当該分野で公知（例えば、Fukunaga, 1990, Statistical Pattern Recognition, 第2版、サンジエゴ; Academic Press; Everitt, 1974, Cluster Analysis, ロンドン; Heinemann Educ. Books; Hartigan, 1975, Clustering Algorithms, ニューヨーク; Wiley; Sneath and Sokal, 1973, Numerical Taxonomy, Freeman; Anderberg, 1973, Cluster Analysis for Applications, New York; Academic Press参照）であり、例えば階層制集塊性クラスタリングアルゴリズム、Hartiganの「k手段」アルゴリズム、およびモデルベースのクラスタリングアルゴリズム（例えばMathSoft, Inc.によるmclust）を含む。本発明の方法において、好ましくは階層制クラスタリング法および/またはアルゴリズムが使用される。特に好適な実施形態において、hclustルーチンまたはアルゴリズムを使用して本発明のクラスタリング分析が行われる（例えば、ソフトウェアパッケージS-Plusからの「hclust」ルーチン、MathSoft, Inc.、マサチューセッツ州ケンブリッジ、参照）。

【0096】

本発明で使用されるクラスタリングアルゴリズムは、遺伝子発現測定値を含むデータ表の上で機能する。特に本発明のクラスタリング法により分析されるデータ表は、 $m \times k$ アレイまたはマトリックスを含む（ここで、 m は状態または摂動の総数、すなわち異なるsiRNAの総数であり、 k は測定および/または分析された細胞成分（例えば遺伝子の転写体）の数である）。

【0097】

クラスタリングアルゴリズムはかかるアレイまたはマトリックスを分析して、細胞成分

10

20

30

40

50

間の相違点を調べる。数学的には、細胞成分*i*と*j*の差は「距離」 l_{ij} である。例えばある実施形態において、ユークリッド幾何学的距離は、以下の式により決定される

$$I_{i,j} = \left(\sum_n |v_i^{(n)} - v_j^{(n)}|^2 \right)^{1/2} \quad (4)$$

(式中、 $v_i^{(n)}$ と $v_j^{(n)}$ はそれぞれ細胞成分*i*と*j*の摂動*n*に対する応答である)。別の実施形態において、上記式4のユークリッド幾何学的距離は平方されて、細胞成分に漸進的に重みを加ると、これらはさらに離れていく。別の実施形態において距離測定値 l_{ij} は以下の式により与えられるマンハッタン距離 (Manhattan distance) である

$$I_{i,j} = \sum_n |v_i^{(n)} - v_j^{(n)}| \quad (5)$$

別の実施形態において距離は、 $l_{ij} = 1 - r_{ij}$ として定義され、ここで r_{ij} は、応答ベクトル v_i と v_j の「相関係数」または標準化「ドット積」である。例えば r_{ij} は以下の式により定義され

$$r_{i,j} = \frac{v_i \cdot v_j}{|v_i| |v_j|} \quad (6)$$

ここで、ドット積 $v_i \cdot v_j$ は以下の式で定義され

$$v_i \cdot v_j = \sum_n v_i^{(n)} \cdot v_j^{(n)} \quad (7)$$

$$\text{および } |v_i| = (v_i \cdot v_i)^{1/2}; \text{ および } |v_j| = (v_j \cdot v_j)^{1/2}$$

そして、

さらに別の実施形態において、距離尺度はチェビシェフ距離 (Chebychev distance)、パワー距離 (power distance)、およびパーセント不一致 (percent disagreement) でもよく、これらはすべて当該分野で公知である。別の実施形態において距離尺度は $l_{ij} = 1 - r_{ij}$ であり、相関係数は応答ベクトル v_i と v_j の加重ドット積を含む。特にこの実施形態において、 r_{ij} は好ましくは以下の式により定義される

$$r_{i,j} = \frac{\sum_n \frac{v_i^{(n)} \cdot v_j^{(n)}}{\sigma_i^{(n)} \cdot \sigma_j^{(n)}}}{\left[\sum_n \left(\frac{v_i^{(n)}}{\sigma_i^{(n)}} \right)^2 \cdot \sum_n \left(\frac{v_j^{(n)}}{\sigma_j^{(n)}} \right)^2 \right]^{1/2}} \quad (8)$$

ここで、 $v_i^{(n)}$ と $v_j^{(n)}$ は、実験*n*でそれぞれ*i*番目と*j*番目の細胞成分に関連する標準誤差である。

【0098】

式6と8の相関係数は、+1(これは2つの応答が完全に相関しており基本的に同一であることを示す)と-1(これは、2つの応答ベクトルが「反相関」しており「アンチセンス」(すなわち反対である)ことを示す)の値の間で結ばれている。これらの相関係数は、細胞成分セットまたはクラスターが同じ符号の応答を有する成分であると考えられる場合に、特に好ましい。

【0099】

10

20

30

40

50

別の実施形態において、同じ生物学的応答または経路で同時制御されるかまたは含まれるが同様のかつ反相関応答を含む細胞成分セットまたはクラスターを、同定することが好ましい。かかる実施形態において、相関係数として式6または8の絶対値すなわち $|r_{i,j}|$ を使用することが好ましい。

【0100】

さらに別の実施形態において、同時制御および/または同時摂動細胞成分の関係はさらに複雑であり、例えば複数の生物学的経路（例えばシグナル伝達経路）が同じ細胞成分上に収束して異なる結果を生み出す。かかる例では、符号に無関係に同時摂動および/または同時制御細胞成分を同定できる相関係数 $r_{ij} = r_{ij}^{(change)}$ を使用することが好ましい。式9で特定される相関係数はかかる例で特に有用である。

【0101】

$$r_{i,j}^{change} = \frac{\sum_n \left| \frac{v_i^{(n)}}{\sigma_i^{(n)}} \right| \left| \frac{v_j^{(n)}}{\sigma_j^{(n)}} \right|}{\left[\sum_n \left(\frac{v_i^{(n)}}{\sigma_i^{(n)}} \right)^2 \cdot \sum_n \left(\frac{v_j^{(n)}}{\sigma_j^{(n)}} \right)^2 \right]^{1/2}} \quad (9)$$

【0102】

一般に本発明の方法で使用されるクラスタリングアルゴリズムはまた、1つ以上の結合規則を使用して、細胞成分を1つ以上のセットまたは「クラスター」に分類する。例えば単一結合または最隣接法（nearest neighbor method）は、データ表中の2つの最も近い対象間（すなわち、2つの最も近い細胞成分間）の距離を決定する。これに対して、完全結合法は、異なるクラスターまたはセット中の任意の2つの対象（すなわち細胞成分）の間の最大の距離を決定する。あるいは、重み付けをしない対になった群の平均は、2つのクラスター中のすべての対の対象（すなわち細胞成分）の間の平均距離を測定することにより、2つのクラスターまたはセットの間の「距離」を評価する。あるいは重み付けをした対になった群の平均は、2つのクラスター中のすべての対の対象の間の加重平均距離を測定することにより、2つのクラスターまたはセットの間の「距離」を評価し、加重因子は各クラスターのサイズに比例する。他の結合規則（例えば重み付けをしないおよび重み付けをした対になった群の重心、およびウォード（Ward）法）もまた、本発明のいくつかの実施形態に有用である（例えば、Ward, 1963, J. Am. Stat. Assn 58:236; Hartigan, 1975, Clustering Algorithms、ニューヨーク; Wiley）。

【0103】

クラスタリングアルゴリズムが例えば上記のような結合規則を適用してデータ表からの細胞成分をセットまたはクラスターにいったん分類すると、クラスタリング「ツリー」が作成されて、こうして決定された細胞成分のクラスターを例示する。図1Aと1Bでは、図に示した応答プロファイルデータの分析についてhclustクラスタリングアルゴリズムにより作成したクラスタリングツリーを、応答プロファイルデータ表示の上に示す。測定された細胞成分データ $\{v_i^{(n)}\}$ は、各異なる実験 n を含む対の条件（すなわち、siRNA摂動と摂動無し）中の各転写体 i 量の間の比の底10の対数を含む。

【0104】

細胞成分は、上記方法により得られたかまたは提供された細胞成分セットまたはクラスターに従って再アラインメントされ、視覚的に表示される。分析的には、かかる再アラインメントは、特定の摂動応答プロファイルのような特定の元々の生物学的応答プロファイル、 $v_i^{(n)} = \{v_i^{(n)}\}$ を、再アラインメント精製 $\{A_{(i)}^{(n)}\}$ （ここで i は細胞成分指数である）に変換することに対応する。

【0105】

図1Aと1Bは、細胞成分のかかる再アラインメント測定値を示す。特に図1Aは、9つの異

10

20

30

40

50

なる実験で測定された複数の遺伝子転写体（すなわち、細胞成分；横軸）を示し、そのうち8つはMAPK14遺伝子を標的とする異なるsiRNAに暴露し、1つはルシフェラーゼを標的とするsiRNAに暴露した（すなわち、摂動応答プロファイル、縦軸）。図1Bは、17の異なる実験で測定された複数の遺伝子転写体を示し、そのうち16はIGF1R遺伝子を標的とする異なるsiRNAに暴露し、1つはルシフェラーゼを標的とするsiRNAに暴露した。

【0106】

好適な実施形態において、siRNA群のメンバーにより活性化（または脱活性化）された細胞成分は、定性的または定量的方法により同定される。

【0107】

ある実施形態において細胞成分は、複数の摂動について応答プロファイルデータを視覚的に検査することにより同定される。好ましくはかかるデータは、例えば上記方法に従って再アラインメントされ、その結果同時変動細胞成分および同様の応答プロファイルがより容易に同定される。例えば図1Aは、複数の実験で測定した複数の遺伝子転写体（横軸）（すなわち応答プロファイル）を示し、ここでHeLa細胞は、横軸に示すように、MAPK14遺伝子と標的とするsiRNAとルシフェラーゼに暴露した。同時変動細胞成分（すなわち遺伝子セット）が容易に見えるように、細胞成分を分類し再アラインメントした。図1Aの視覚的検査は、表示の左側の行に対応する遺伝子の発現が、すべてまたは多数の異なるsiRNA実験で低下していることを証明する。これに対して、表示の右側の行に対応する遺伝子の発現は、上昇を含む多様な応答を示す。

10

【0108】

より定量的な実施形態において、好ましくは例えば応答プロファイル中の個々の応答強度を閾値で区切ることにより、遺伝子の交点が同定される。ある実施形態において応答強度は、摂動サンプルと非摂動サンプルの比であり、閾値は、発現レベルの2、4または10倍の低下として設定される。

20

【0109】

好適な実施形態において、1つ以上のプロファイル中の遺伝子の応7B54の統計的有意性も測定される。ある実施形態において遺伝子の測定された応答は、Weng、米国特許出願No. 10/349,364（2003年1月22日出願）およびWeng、米国特許出願No. 10/354,664（2003年1月30日出願）（それぞれが参照することによりその全体が本明細書に組み込まれる）に記載のような変換により変換される。次に、変換された応答に基づいて、応答の統計的有意性が決定される。ある実施形態において統計的有意性は、p値により特徴付けられ、変換された応答の摂動がランダムな誤差に由来する確率を示す。好適な実施形態において、その応答がある閾値レベルより小さいp値である閾値レベルより上の倍数変化を有する遺伝子は、有意に制御された遺伝子として選択される。

30

【0110】

別の実施形態において、遺伝子の応答の統計的有意性は、パーセントイルランクにより特徴付けられる（例えば、US Patent No. 6,351,712参照、これは参照することにより本明細書に組み込まれる）。ある実施形態において、目的の遺伝子がプロファイルのアップレギュレーションまたはダウンレギュレーションの上位1%中に存在するなら、プロファイル中のその遺伝子のパーセントイルランクは、p値=0.01として表される。kプロファイル中の遺伝子のパーセントイルランクは以下の式により与えられる

40

$$p = \prod_i p_i \quad (10)$$

ここで、 p_i は i 番目のプロファイル中の遺伝子のp値である。ある実施形態において、1つ以上のプロファイル中のそのp値が閾値より小さい遺伝子が同定される。好適な実施形態において、そのp値が応答プロファイルの少なくとも50%、70%、または90%で0.01未満である遺伝子は、共通に制御された遺伝子として同定される。

【0111】

50

5.2.3. 表現型の原因である候補遺伝子の測定法

本発明は、真核細胞の表現型特徴の原因である1つ以上の候補遺伝子の測定法を提供する。この方法は、細胞中の標的遺伝子をサイレンシングするように設計された1つ以上のsiRNAにより共通にサイレンシングされる1つ以上の遺伝子を同定すること、および1つ以上のsiRNAの摂動に関連する表現型特徴に遺伝子を関連させることを含む。

【0112】

ある実施形態において、標的遺伝子をサイレンシングするように設計されたsiRNAが細胞中に導入される。siRNAの導入に関連する表現型特徴が同定される。siRNAの導入後の複数の時点での細胞中のsiRNAの応答プロファイルが測定される。siRNAにより直接サイレンシングされる1つ以上の非標的遺伝子が、セクション5.2.2.1に記載の方法を使用して動力学に基づいて同定される。1つ以上の非標的遺伝子は標的遺伝子とともに、表現型特徴の原因である候補遺伝子として同定される。

10

【0113】

別の実施形態において標的遺伝子をサイレンシングするように設計された複数の異なるsiRNAが、ある細胞タイプの細胞に別々に導入される。siRNAの導入に共通に関連する表現型特徴が同定される。siRNAのそれぞれの各プロファイルを含む応答プロファイルが測定される。複数のsiRNAにより共通にサイレンシングされる1つ以上の遺伝子(標的遺伝子と非標的遺伝子の両方)が、セクション5.2.2.2に記載の方法を使用して同定される。表現型特徴の原因である候補遺伝子として、1つ以上の遺伝子が同定される。

【0114】

好適な実施形態において標的遺伝子をサイレンシングするように設計された複数の異なるsiRNAが、ある細胞タイプの細胞と一緒に導入される。siRNAの導入に共通に関連する表現型特徴が同定される。複数のsiRNAの応答プロファイルが測定される。標的遺伝子は、表現型特徴の原因である候補遺伝子として同定される。好適な実施形態において複数の異なるsiRNAは、3、5、9、12、15、20、50または100個の異なるsiRNAを含む。好ましくは複数のsiRNAの総濃度は、目的の標的遺伝子をサイレンシングするのに最適な濃度である。最適濃度は、それ以上濃度を上げてサイレンシングのレベルが実質的に上昇しない濃度である。ある実施形態において最適濃度は、それ以上濃度を上げてサイレンシングのレベルが5%、10%、または20%より大きく上昇しない濃度である。好適な実施形態において、複数のsiRNAが30%、20%、10%、または5%、1%、0.1%、または0.01%未満のオフターゲット遺伝子のサイレンシングしか引き起こさないように、複数の組成(複数中の異なるsiRNAの数と各異なるsiRNAの濃度を含む)が選択される。別の好適な実施形態において複数の異なるsiRNAは各siRNAを等しい比率で含む。さらに別の好適な実施形態において複数の異なるsiRNAは、各siRNAを互いに5%、10%、20%または50%未満しか異ならないように含む。さらに別の好適な実施形態において、複数の異なるsiRNA中のどのsiRNAも、その複数中の総siRNA濃度の90%、80%、70%、50%、または20%以上を構成しない。別の実施形態において複数中の各siRNAは、個々に使用される時の濃度より低い濃度を有する。好適な実施形態において複数の異なる各siRNAは、他のsiRNAの非存在下または遺伝子をサイレンシングするように設計された他のsiRNAの非存在下で使用される時、少なくとも30%、50%、75%、80%、85%、90%、または95%のサイレンシングを達成するのに有効なsiRNAの濃度より低い濃度を有する。別の実施形態において、複数の異なる各siRNAは、他のsiRNAの非存在下または遺伝子をサイレンシングするように設計された他のsiRNAの非存在下で使用される時、遺伝子の30%、20%、10%、または5%未満のサイレンシングしか引き起こさない濃度を有する。好適な実施形態において、複数のsiRNAが標的遺伝子の少なくとも80%または90%のサイレンシングを引き起こしても、各siRNAは単独で使用されると標的遺伝子の30%、20%、10%、または5%未満のサイレンシングしか引き起こさない濃度を有する。

20

30

40

【0115】

5.2.4. 遺伝子サイレンシングにおいて鎖の優先を測定する方法。

【0116】

50

本発明は、遺伝子サイレンシングにおける遺伝子の優先、すなわちどの遺伝子がより活性であるかを測定する方法を提供する。本発明は、siRNAによりダウンレギュレートされる複数の遺伝子（例えばプロファイリングにより同定されるダウンレギュレートされたシグナチャ遺伝子）の転写体配列との、siRNAのセンス鎖対アンチセンス鎖のアラインメントを比較することを含む（セクション5.2を参照）。ある実施形態において複数の遺伝子は、siRNAにより直接サイレンシングされるオフターゲット遺伝子である。かかるオフターゲット遺伝子は、例えばサイレンシング速度を調べることにより同定することができる（セクション5.2）。別の実施形態において複数の遺伝子は、ある時点で測定すると転写体レベルがsiRNAによりダウンレギュレートされる遺伝子を含む。すなわちこの実施形態における、複数の遺伝子は直接サイレンシングされたオフターゲット遺伝子と、およびsiRNAによりサイレンシングされる1つ以上の他の遺伝子によりコードされる1つ以上のタンパク質の喪失の結果としてその転写体レベルが低下する遺伝子とを含んでよい。

【0117】

ある実施形態において、センス鎖またはアンチセンス鎖と同一のアラインメントが含まれる。本発明者は、そのダウンレギュレーション速度を調べたsiRNAについて、センス鎖とアンチセンス鎖の両方とのアラインメントが観察されることを発見した（例えば、実施例6.3参照、後述）。これらの協調してダウンレギュレートされるオフターゲットシグナチャ遺伝子は、2つの方法でこれらを制御するsiRNAをアラインメントすることがわかった：(1)同一の中央連続区間：すべてオフターゲット遺伝子と同一のsiRNA2本鎖の中央部分の11以上の塩基；または(2)同一の3'ベースの連続区間：すべてオフターゲット遺伝子と同一のsiRNA2本鎖の3'末端の3塩基内で終止する8以上の塩基。

【0118】

アラインメントがセンス鎖と同じsiRNAについては、アンチセンス鎖がオフターゲット制御とオフターゲット制御との両方を引き起こしており；アラインメントがアンチセンス鎖と同じsiRNAについては、センス鎖がアンチセンス鎖の活性以外にオフターゲット制御を引き起こしていると推測される。

【0119】

ある実施形態において、単一時点のシグナチャは、1つのsiRNA鎖対他のsiRNA鎖とのアラインメントのより大きな程度について、すべてのシグナチャ遺伝子のすべてまたは一部（すなわち、ダウンレギュレートされた遺伝子）を投票（polling）することにより評価される（「投票法（polling method）」）。好適な実施形態においてシグナチャ遺伝子配列は、siRNAの両方の鎖とアラインメントされる。各鎖とのアラインメントは、最も長い連続的同一区間について調べられる。これらの同一区間の長さが比較される。遺伝子は、より長い連続的同一区間を有する鎖に投票したと見なされる。ある実施形態において投票は、2つの同一性の区間の短い方より過剰の2つの同一性の区間の長い方の塩基の数により重み付けされる。すなわち、配列同一性の2つの区間の長さのより大きな差を有するsiRNAが、より多くの票（vote）に寄与する。例えば、1つの鎖に15塩基の同一性区間を有し別の鎖に5塩基の同一性区間を有するsiRNAは、1つの鎖に8塩基の同一性区間を有し別の鎖に7塩基の同一性区間を有するsiRNAより、大きく重み付けをされる。従って票は、ある鎖の過剰の連続的同一性区間である。総票数は、それが優勢であったすべての遺伝子中の勝者の鎖の連続的同一性の総過剰長さである。

【0120】

ある実施形態においてsiRNAの鎖とのアラインメントの程度のバックグラウンド（バックグラウンド鎖優先）は、siRNAの両方の鎖と遺伝子（すなわち、siRNAについてその転写体レベルが調べられたすべての遺伝子）のより大きいセットとのアラインメントを調べることにより測定される。バックグラウンド鎖の優先は、上記したように鎖に投票させて、siRNAの各鎖について遺伝子当たりの平均過剰重みを計算することにより決定される。好適な実施形態において転写体レベルは、DNAマイクロアレイを使用して測定される。かかる実施形態において、測定したすべての遺伝子（例えば、siRNAプロファイルに含まれるすべての遺伝子）を使用することができる。好ましくは、siRNAの応答プロファイルについて5よ

り大きい、10より大きい、100より大きい、1000より大きい、10,000より大きい、25,000より大きい、50,000より大きい、または100,000より大きい転写体レベルが測定され、使用される。

【0121】

別の実施形態において、目的の標的遺伝子と同じ動力学でsiRNAによりダウンレギュレートされる複数の遺伝子を調べることにより、siRNAの鎖の優先を評価し測定するために投票法が使用される。かかる遺伝子は、セクション5.2に記載のように同定される。鎖優先は、1つのsiRNA鎖対他のsiRNA鎖とのアラインメントのより大きな程度について、かかるシグナチャ遺伝子を投票することにより評価される。

【0122】

別の実施形態において、目的の標的遺伝子と同じ動力学でダウンレギュレートした遺伝子とのsiRNAのアラインメントから開発された3'偏りモデル(3'偏りモデル)とのシグナチャ遺伝子のアラインメントの比較により、単一時点のシグナチャが評価される(セクション5.2参照)。ある実施形態において、シグナチャ遺伝子は分析されて、オフターゲット遺伝子アラインメントの3'偏りモデル(すなわち、オフターゲット遺伝子とすべて同一のsiRNA2本鎖の3'末端の3塩基内で終止する8つ以上の塩基の3'ベースの連続的同一区間)に一致する割合を調べる。シグナチャ遺伝子配列は、siRNAの両方の鎖とアラインメントされる。各鎖とのアラインメントが、連続的同一区間について調べられる。アラインメントは、もしa)少なくとも7つの連続的同一区間を有するなら;およびb)特定の鎖の3'末端の3塩基以内で終止するなら、印が付けられる。

【0123】

ある実施形態において、センス鎖との同一性およびアンチセンス鎖との同一性を有する上記基準a)とb)を満足するアラインメントの印は、以下の式のSLRスコアを計算することにより比較される

$$SLR = \log(\text{センス同一印} / \text{アンチセンス同一印}) \quad (11)$$

SLRの有意性は、その転写体レベルがsiRNAについて測定されるすべての遺伝子についてセンス同一印とアンチセンス同一印を有するシグナチャについて、センス同一印とアンチセンス同一印を比較することにより評価される。好適な実施形態において、転写体レベルはDNAマイクロアレイを使用して測定される。かかる実施形態においては、測定されるすべての遺伝子(例えばsiRNAプロファイル中に含まれるすべての遺伝子)を使用することができる。好ましくは、siRNAの応答プロファイルについて5より大きい、10より大きい、100より大きい、1000より大きい、10,000より大きい、25,000より大きい、50,000より大きい、または100,000より大きい転写体レベルが測定され、使用される。ある実施形態において有意性は以下の式に従って超幾何分布から計算される

$$h(x, n, M, N) = \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}} \quad (12)$$

ここで

- x、サンプル - 選択 = 1つの鎖のシグナチャ印
- n、サンプル - 合計 = 両方の鎖のシグナチャ印
- M、親 - 選択 = 同じ鎖のチップ印
- N、親 - 合計 = 両方の鎖のチップ印

超幾何分布は対称ではないため、センス鎖とアンチセンス鎖選択超幾何分布の両方のテイルが使用される。センス同一過剰に対応する2つのテイルを平均して過剰センス同一性のp値が得られ、アンチセンス同一過剰に対応する両方のテイルを平均して過剰アンチセンス同一性のp値が得られる。

【0124】

別の実施形態において、目的の標的遺伝子と同じ動力学でダウンレギュレートされた複数の遺伝子を調べることにより、siRNAの鎖の優先を評価し測定するために3'偏りモデルが使用される。かかる遺伝子は、セクション5.2に記載のように同定される。鎖優先は、1つのsiRNA鎖対他のsiRNA鎖とのアラインメントについてかかるシグナチャ遺伝子を印付けすることにより評価される。

【0125】

特に投票法は、12時間、24時間、および48時間シグナチャ（実施例3参照、後述）の分析に有効であることがわかった。24時間シグナチャが一般に最も有意な結果を与えた。実際、siRNAの鎖優先性の測定についてのシグナチャのかなりの割合またはシグナチャ全体の有用性は、細胞中の遺伝子発現に対するsiRNAの直接作用がシグナチャ遺伝子のかなりの割合により示されることを意味する。例えば両方法は、約35～40%の12時間IGF1R-73シグナチャが鎖の偏りの推定に関連と推定する（すなわち約12時間）。35～40%の24時間IGF1R-73シグナチャもまた関連する（すなわち約80遺伝子）。

10

【0126】

さらに別の実施形態において本発明は、siRNAの位置特異的塩基組成に基づいてsiRNAの鎖優先性および/または有効性および特異性を予測する方法を提供する。本発明者は、その塩基組成PSSMスコア（2003年10月27日出願の米国仮特許出願第60/515,223号、Jacksonら、これは参照することによりその全体が本明細書に組み込まれる）がその逆相補体の塩基組成PSSM（G/C PSSM）スコアより大きいsiRNAが、そのセンス鎖より活性の高いアンチセンス鎖を有すると予測されることを発見した。これに対して、その塩基組成PSSMスコアがその逆相補体の塩基組成PSSMスコアより小さいsiRNAは、そのアンチセンス鎖より活性の高いセンス鎖を有すると予測される。

20

【0127】

有効性の上昇は、アンチセンス鎖のより高い活性とセンス鎖のより低い活性に対応することが証明されている。本発明者は、悪いsiRNAとして強いセンス鎖を有するsiRNAを、良いsiRNAとして弱いセンス鎖を有するsiRNAから区別するのに使用できることを発見した。悪いsiRNAの逆相補体は、良いsiRNAよりも、悪いsiRNA自体とより大きく異なることがわかった。平均すると、悪いsiRNAの逆相補体は、良いsiRNAよりも5'末端でのG/C含量はるかに大きく、3'末端では良いsiRNAと同様のB/C含量であった。これに対して、良いsiRNAの逆相補体は、良いsiRNAよりも悪いsiRNAに実質的に類似していた。平均すると、良いsiRNAの逆相補体は、5'末端のG/C含量が悪いsiRNAとあまり異ならず、3'末端で悪いsiRNAよりG/C含量がわずかに少なかった。これらの結果は、G/C PSSMが、悪いsiRNAとして強いセンス鎖を有するsiRNAを、良いsiRNAとして弱いセンス鎖を有するsiRNAから区別していることを示す。

30

【0128】

図13Aは、19量体2本鎖領域内の悪いsiRNAの逆相補体の平均G/C含量と、悪いsiRNA自体の平均G/C含量の差を示す。良いsiRNAと悪いsiRNAの平均G/C含量の差を、比較のために示す。曲線は5のウィンドウにわたって平滑化した（配列の端で、5のウィンドウの部分）。

【0129】

図13Bは、19量体2本鎖領域内の良いsiRNAの逆相補体の平均G/C含量と、悪いsiRNAの平均G/C含量の差を示す。良いsiRNAと悪いsiRNAの平均G/C含量の差を、比較のために示す。曲線は5のウィンドウにわたって平滑化した（配列の端で、5のウィンドウの部分）。

40

【0130】

図14では、測定されたサイレンシング効率により区分けし、3'偏り法とG/C PSSM法によるセンス活性コールの頻度を比較した。これらの方法は異なる分析に基づくが、一致は極めて良好である。両方とも、高サイレンシングsiRNAに対して低サイレンシングsiRNAの高い比率がセンス活性であることを示す。（siRNA G/C PSSMスコア - 逆相補体G/C PSSMスコア）対 \log_{10} （センス同スコア / アンチセンス同スコア）の相関係数は、図14に区分けした61のsiRNAのセットについて0.59であった。

【0131】

50

すなわちある実施形態において本発明は、siRNAの位置特異的塩基組成に基づくsiRNAの鎖優先を予測する方法を提供する。ある実施形態においてこの方法は、siRNAのセンスおよびアンチセンス鎖の塩基組成を比較することにより、遺伝子サイレンシングにおけるsiRNAの鎖優先を評価する。別の実施形態においてこの方法は、siRNAのセンス配列および標的配列の逆相補体の塩基組成を比較することにより、遺伝子サイレンシングにおけるsiRNAの鎖優先を評価する。

【0132】

ある実施形態において、siRNAのアンチセンス鎖または転写体中のsiRNAの標的配列の逆相補体が、PSSMアプローチを使用してセンス鎖と比較される（2003年10月27日出願の米国仮特許出願第60/515,180号参照、Jacksonら、これは参照することによりその全体が本明細書に組み込まれる）。siRNAとその逆相補体は、加重マトリックスとして、2本鎖領域内の良いsiRNAと悪いsiRNAとの平滑化したG/C含量の差に基づいたPSSMを使用してスコア化される。ある実施形態において、図13Aに記載の塩基組成加重マトリックスが加重マトリックスとして使用される。好適例では各鎖のPSSMスコアは、siRNA鎖G/C含量とG/C含量差マトリックスとのドット積として計算することができる（米国仮特許出願第60/515,180号に記載の曲線モデルPSSMのスコア計算法として）。ある実施形態において、その逆相補体PSSMスコアがそれ自体のPSSMスコアより大きい場合、siRNAはセンス活性として同定される。

10

【0133】

別の実施形態において、siRNAの鎖優先を測定するために3'偏り法がPSSMスコアと一緒に使用される。かかる実施形態において、アンチセンス同一スコアがセンス同一スコアより大きい場合、鎖優先測定の3'偏り法によりsiRNAはセンス活性であると同定される。

20

【0134】

鎖の偏りの予測のためにsiRNAとその逆相補体のG/C PSSMの比較に基づく方法を、3'偏り法によるsiRNA発現プロファイルからの鎖偏りの推定との比較により試験した。

【0135】

本発明はまた、siRNAのサイレンシング特異性を改良する方法を提供する。この方法は、オフターゲットサイレンシングにおけるsiRNAの鎖優先を測定する。siRNAがセンス鎖優先を有するとして測定される場合、センス鎖活性が低下するか排除されるように、センス鎖が修飾される。ある実施形態において修飾は、センス鎖中に2つの2'-O-メチル置換を導入することにより行われる。

30

【0136】

5.3. 遺伝子サイレンシングの方法と組成物

本発明は、標的遺伝子に対してほんの部分的な配列相同性を有するsiRNAを使用する、遺伝子サイレンシングの方法と組成物を提供する。好適な実施形態において本発明は、標的遺伝子の転写体配列と同一であるが転写体中のどの配列とも完全長相同性を持たない11~18ヌクレオチドのセンス鎖またはアンチセンス鎖連続的ヌクレオチド配列を含むsiRNAを使用して、標的遺伝子をサイレンシングするための方法と組成物を提供する。好ましくは連続的ヌクレオチド配列は、siRNA分子の中央領域にある。siRNAの中央領域にある連続的ヌクレオチド配列は、3'末端で開始しないsiRNA中のヌクレオチド配列の任意の連続区間である。例えば11ヌクレオチドの連続的ヌクレオチド配列は、ヌクレオチド配列2~12、3~13、4~14、5~15、6~16、7~17、8~18、または9~19でもよい。好適な実施形態において連続的ヌクレオチド配列は、11~16、11~15、14~15、11、12、または13ヌクレオチドの長さである。

40

【0137】

別の好適な実施形態において本発明は、標的遺伝子の転写体配列と同一の9~18ヌクレオチドの3'センス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含むが、転写体の連続配列と完全長配列同一性を持たないsiRNAを使用して、標的遺伝子をサイレンシングするための方法と組成物を提供する。本出願において、3'の9~18ヌクレオチド配列は、第1の対になった塩基で始まるヌクレオチドの連続区間であり、すなわちこれは2塩基3'オ

50

ーバーハングを含まない。すなわち、特定のヌクレオチド配列がsiRNAの3'末端にあると言う場合、2塩基オーバーハングは考慮されない。好適な実施形態において連続的ヌクレオチド配列は、9~16、9~15、9~12、11、10、または9ヌクレオチドの長さである。

【0138】

別の好適な実施形態において、標的遺伝子をサイレンシングするように設計された複数の異なるsiRNAは、一緒にある細胞タイプの細胞中に導入される。異なる各siRNAは異なる配列を含む。好適な実施形態において複数の異なるsiRNAは、3、5、9、12、15、20、50、または100の異なるsiRNAを含む。好ましくは複数のsiRNAの総濃度は、目的の標的遺伝子をサイレンシングするのに最適な濃度である。最適濃度は、それ以上濃度を上げてサイレンシングのレベルが実質的に上昇しない濃度である。ある実施形態において最適濃度は、それ以上濃度を上げてサイレンシングのレベルが5%、10%、または20%より大きく上昇しない濃度である。好適な実施形態において、複数のsiRNAが30%、20%、10%、または5%、1%、0.1%、または0.01%未満のオフターゲット遺伝子のサイレンシングしか引き起こさないように、複数の組成（複数中の異なるsiRNAの数と各異なるsiRNAの濃度を含む）が選択される。別の好適な実施形態において、複数の異なるsiRNAは各siRNAを等しい比率で含む。さらに別の好適な実施形態において、複数の異なるsiRNAは、各siRNAを互いに5%、10%、20%または50%未満しか異ならないように含む。さらに別の好適な実施形態において、複数の異なるsiRNA中のどのsiRNAも、その複数中の総siRNA濃度の90%、80%、70%、50%、または20%以上を構成しない。別の実施形態において、複数中の各siRNAは、個々に使用される時の濃度より低い濃度を有する。好適な実施形態において、複数の異なる各siRNAは、他のsiRNAの非存在下または遺伝子をサイレンシングするように設計された他のsiRNAの非存在下で使用される時、少なくとも30%、50%、75%、80%、85%、90%、または95%のサイレンシングを達成するのに有効なsiRNAの濃度より低い濃度を有する。別の実施形態において、複数の異なる各siRNAは、他のsiRNAの非存在下または遺伝子をサイレンシングするように設計された他のsiRNAの非存在下で使用される時、遺伝子の30%、20%、10%、または5%未満のサイレンシングしか引き起こさない濃度を有する。好適な実施形態において、複数のsiRNAが標的遺伝子の少なくとも80%または90%のサイレンシングを引き起こしても、各siRNAは単独で使用されると標的遺伝子の30%、20%、10%、または5%未満のサイレンシングしか引き起こさない濃度を有する。

【0139】

具体的な実施形態において本発明は、KPNB3、RAP2A、FLJ20291、RRAD、RPA2、DKFZp564J157、AF093680、および2つの未解析ESTコンティグ（contig）に対応する遺伝子、コンティグ53709_RCとコンティグ56528_RCの転写体を、これらの遺伝子中の配列とほんの部分的な配列相同性しか持たないsiRNAを使用して、サイレンシングするための方法と組成物を提供する。これらの遺伝子をサイレンシングするために使用されるsiRNAは、表1に示すようにMAPK14-1~MAPK14-8である。配列アラインメントは、これらの遺伝子が2つの亜群に分けられることを示した。3つの遺伝子からなる1つの亜群は、siRNA配列の中央領域を含む類似性の14~15ヌクレオチドのコアを含有した。第2の亜群は、siRNAセンス鎖配列の3'末端に9つのヌクレオチドを含む類似性のより小さいコアを含有した。これは、siRNA配列中にランダムに分布した類似性の短い区間（<6~8ヌクレオチド）を示した動力学群3~5中の転写体とは反対である。すなわちsiRNAの3'末端を包含する配列類似性のコアについての偏りは、急速にサイレンシングされる転写体にユニークである。この可能性を試験するために、siRNA配列の各位置のヌクレオチドに系統的な置換を行い、改変配列の発現シグナチャへの作用を調べた（図7）。代表的結果を図5Bに示す。4位の単一ヌクレオチド置換はMAPK14のサイレンシングを劇的に低下させ、この位置でMAPK14との類似性を有する亜群1中の3つのオフターゲット遺伝子のサイレンシングを排除した。しかし、この領域にMAPK14との類似性を持たない亜群2の6つのオフターゲット遺伝子では、サイレンシングは排除されなかった。5位の単一ヌクレオチド置換は、MAPK14サイレンシングを低下させたが排除はせず、この領域にMAPK14との類似性を有する亜群1中の3つの遺伝子のサイレンシ

ングを排除した。この領域に類似性を持たない6つのオフターゲット遺伝子の発現レベルは、このミスマッチにより影響を受けなかった。これらの結果は、オフターゲット遺伝子のサイレンシングが、MAPK14発現の喪失には依存しないことを確認する。15位の単一ヌクレオチド置換もまた、MAPK14のサイレンシングを低下させ、9つすべてのオフターゲット遺伝子のサイレンシングを排除した。15位のミスマッチの影響はより劇的にであり、これはおそらく、9つすべての転写体がこの領域でMAPK14との類似性を有するためであろう。まとめるとこれらの結果は、これらの転写体のsiRNA MAPK14-1~8との配列類似性が、MAPK14 siRNAによるこれらのサイレンシングの原因であることを確認している。

【0140】

観察されたオフターゲット遺伝子サイレンシングが、低下したMAPK14発現の結果ではなく配列類似性に基づくことのさらなる証拠として、異なるオリゴヌクレオチド配列(図5A)を有する異なるsiRNA2本鎖を試験した。この第2のsiRNA(MAPK14-2)はMAPK14をサイレンシングさせたが、第2群転写体のいずれもサイレンシングさせず、これらは新しいsiRNA配列に似ていない(図5B)。

【0141】

別の具体例において、MAPK14遺伝子中の配列とわずかな配列相同性しか持たないsiRNAを使用してMAPK14をサイレンシングする方法と組成物が提供される。2つの転写体KPNB3とFLJ20291(図5C)のsiRNAは、その目的の標的以外にMAPK14の発現をサイレンシングさせた(図5C)。KPNB3 siRNAはMAPK14と、14の連続ヌクレオチドおよび全部で15のヌクレオチドを共有する。FLJ20291 siRNAはMAPK14と、わずかに11の連続ヌクレオチドおよび全部で15のヌクレオチドを共有する。すなわち非標的転写体を直接サイレンシングするのに、わずかに11個の配列同一性の連続ヌクレオチドで充分である。

【0142】

本発明はまた、疾患または好ましくない症状に罹った臓器または組織の細胞中の1つ以上の遺伝子の発現を低下させることにより、哺乳動物のような動物の疾患または他の好ましくない症状を治療する方法を提供する。この方法は、動物に、本発明のsiRNA、例えば11~15、14~15、11、12、または13ヌクレオチドのセンスまたはアンチセンス鎖の中央の連続的ヌクレオチド配列を含むsiRNA、または9~15、9~12、11、10、または9ヌクレオチドの3'センス鎖またはアンチセンス鎖の連続的ヌクレオチド配列を含むsiRNA、を投与することを含む。siRNAは当該分野で公知の任意の方法を使用して投与される。siRNAのin vivo送達のために、好ましくはセクション5.5(後述)に記載の方法が使用される。ある実施形態においてこの方法は、ヒトの疾患または好ましくない症状を治療するのに使用される。

【0143】

好適な実施形態において本発明は、そのダウンレギュレーションが疾患の治療に有効な標的遺伝子をサイレンシングするように設計された複数の異なるsiRNAを動物に投与することにより、哺乳動物のような動物の疾患または好ましくない症状を治療する方法を提供する。好適な実施形態において複数の異なるsiRNAは、3、5、9、12、15、20、50、または100の異なるsiRNAを含む。好ましくは複数のsiRNAの総濃度は、目的の標的遺伝子をサイレンシングするのに最適な濃度である。最適濃度は、それ以上濃度を上げてサイレンシングのレベルが実質的に上昇しない濃度である。ある実施形態において最適濃度は、それ以上濃度を上げてサイレンシングのレベルが5%、10%、または20%より大きく上昇しない濃度である。好適な実施形態において、複数のsiRNAが30%、20%、10%、または5%、1%、0.1%、または0.01%未満のオフターゲット遺伝子のサイレンシングしか引き起こさないように、複数の組成(複数中の異なるsiRNAの数と各異なるsiRNAの濃度を含む)が選択される。別の好適な実施形態において、複数の異なるsiRNAは各siRNAを等しい比率で含む。さらに別の好適な実施形態において、複数の異なるsiRNAは、各siRNAを互いに5%、10%、20%または50%未満しか異ならないように含む。さらに別の好適な実施形態において、複数の異なるsiRNA中のどのsiRNAも、その複数中の総siRNA濃度の90%、80%、70

10

20

30

40

50

%、50%、または20%以上を構成しない。別の実施形態において、複数中の各siRNAは、個々に使用される時の濃度より低い濃度を有する。好適な実施形態において、複数の異なる各siRNAは、他のsiRNAの非存在下または遺伝子をサイレンシングするように設計された他のsiRNAの非存在下で使用される時、少なくとも30%、50%、75%、80%、85%、90%、または95%のサイレンシングを達成するのに有効なsiRNAの濃度より低い濃度を有する。別の実施形態において、複数の異なる各siRNAは、他のsiRNAの非存在下または遺伝子をサイレンシングするように設計された他のsiRNAの非存在下で使用される時、遺伝子の30%、20%、10%、または5%未満のサイレンシングしか引き起こさない濃度を有する。好適な実施形態において、複数のsiRNAが標的遺伝子の少なくとも80%または90%のサイレンシングを引き起こしても、各siRNAは単独で使用されると標的遺伝子の30%、20%、10%、または5%未満のサイレンシングしか引き起こさない濃度を有する。

10

【0144】

5.4. 遺伝子サイレンシングのための小干渉性RNAの設計法

本発明は、1つ以上の遺伝子をサイレンシングするためのsiRNAを設計する方法を提供する。

【0145】

ある実施形態において本発明は、遺伝子を標的とするが1つ以上の選択された非標的遺伝子を標的としないsiRNAの配列を設計する方法を提供する。好適な実施形態において、siRNAの配列は、中央領域のセンス鎖またはアンチセンス鎖が1つ以上の選択された非標的遺伝子の転写体中の任意の配列と同一ではないように選択される。好適な実施形態において中央領域は約10~14ヌクレオチドを含む。別の好適な実施形態においてsiRNAの中央領域の配列は、1つ以上の非標的遺伝子中の配列と比較して少なくとも1、2、3または5つのミスマッチを有する。

20

【0146】

別の好適な実施形態においてsiRNAの配列は、3'センス鎖またはアンチセンス鎖が1つ以上の選択された非標的遺伝子の転写体中の任意の配列と同一ではないように選択される。好適な実施形態において3'配列は約9ヌクレオチドを含む。別の好適な実施形態においてsiRNAの3'配列は、1つ以上の非標的遺伝子中の配列と比較して少なくとも1、2、3または5つのミスマッチを有する。

【0147】

別の好適な実施形態において本発明は、9~18ヌクレオチドの共通のヌクレオチド配列を共有する複数の異なる遺伝子を標的とするsiRNAの配列を設計する方法を提供する。この方法は、(i)該共通配列を含む11~18ヌクレオチドのセンス鎖またはアンチセンス鎖の中央の連続的ヌクレオチド配列と、および/または(ii)該共通配列中の配列と同一の少なくとも9ヌクレオチドの3'センス鎖またはアンチセンス鎖の連続的ヌクレオチド配列とを含む、小干渉性RNA(siRNA)を選択することを含む。好適な実施形態において中央の連続的ヌクレオチド配列は、11~15、14~15、11、12、または13ヌクレオチドの長さである。別の好適な実施形態において3'連続的ヌクレオチド配列は、9~15、9~12、11、10、または9ヌクレオチドの長さである。

30

【0148】

核酸配列を設計するための任意の公知の方法を、この目的に使用することができる。

40

【0149】

さらに別の好適な実施形態において本発明は、標的遺伝子をサイレンシングするためのsiRNAを設計する方法を提供する。この方法は、それぞれが標的遺伝子の転写体中の異なる配列を標的とするように設計された、複数の異なるsiRNAを作成することを含む。次に複数のsiRNAのそれぞれについて、応答プロファイルが測定される。次に所望の応答プロファイルを有する1つ以上のsiRNAが同定される。ある実施形態において、最小数のオフターゲット遺伝子を有する1つ以上のsiRNAが同定され選択される。別の実施形態において、1つ以上の選択された遺伝子に影響を与えない(例えば、1つ以上の選択された遺伝子の転写体レベルの50%、25%、10%、または5%未満の低下を引き起こす)1つ以上のs

50

siRNAが同定され選択される。好適な実施形態において複数の異なるsiRNAは、標的遺伝子の一部のまたはすべてのコード配列に重ねた配列を有する（すなわち、標的遺伝子の一部のまたはすべてのコード配列に重ねた重複配列を有する）siRNAを含む。ある実施形態において複数の異なるsiRNAは、20、15、10、5、2または1塩基の間隔で標的遺伝子の全コード配列に重ねた配列を有するsiRNAを含む。別の実施形態において複数の異なるsiRNAは、20、15、10、5、2または1塩基の間隔で標的遺伝子のすべての公知のまたは予測されるエキソンおよびエキソン変種の配列に重ねた配列を有するsiRNAを含む。

【0150】

選択されたsiRNAはさらに、標的遺伝子をサイレンシングする効率および/または特異性について評価することができる。ある実施形態において、標的転写体レベルの最大の低下を示すsiRNAが選択される。別の実施形態において、標的遺伝子の転写体レベルの特定の低下レベル（例えば、標的遺伝子の転写体レベルの低下の少なくとも5%、10%、25%、50%、75%、90%、または95%）が所望の時、転写体レベルの所望の低下を示すsiRNAが選択される。好適な実施形態において、2003年10月27日出願の米国仮特許出願第60/515,180号参照、Jacksonら（これは参照することによりその全体が本明細書に組み込まれる）に記載の方法が、遺伝子サイレンシングにおけるsiRNAの効率および/または特異性を評価するのに使用される。

【0151】

5.5 RNA干渉を実施するための方法

RNA干渉を実施するのに当該分野で公知の任意の方法が使用できる。ある実施形態において、細胞にsiRNAを提示することにより遺伝子サイレンシングが誘導され、これはダイサー（Dicer）切断の生成物を模倣する（例えば、Elbashirら、2001, Nature 411:494-498; Elbashirら、2001, Genes Dev. 15:188-200参照、これらのすべては参照することによりその全体が本明細書に組み込まれる）。合成siRNA2本鎖は、リスク（RISC）に結合し、mRNA転写体のサイレンシングを指令する能力を保持し、従って哺乳動物細胞中で遺伝子サイレンシングを行うための強力な手段を科学者に提供する。siRNAは化学的に合成できるか、または組換えダイサーによる2本鎖RNAの切断から得られる。

【0152】

遺伝子サイレンシングのための2本鎖DNA（dsDNA）を導入する別の方法は、shRNA（短いヘアピンRNA）である（例えば、Paddisonら、2002, Genes Dev. 16:948-958; Brummelkampら、2002, Science 296:550-553; Sui, G.ら、2002, Proc. Natl. Acad. Sci. USA 99:5515-5520参照、これらのすべては参照することによりその全体が本明細書に組み込まれる）。この方法では所望のsiRNA配列が、ヘアピン構造を形成するための介在ループ配列を有する反転繰り返し配列としてプラスミド（またはウイルス）から発現される。ヘアピンを含む生じるRNA転写体は次に、ダイサーにより処理されてサイレンシングのためのsiRNAを生成する。プラスミドベースのsiRNAは細胞中で安定に発現され、in vitroとin vivo（例えば動物中）の両方で細胞中の長期遺伝子サイレンシングを可能にする（McCaffreyら、2002, Nature 418:38-39; Xiaら、2002, Nature Biotech. 20:1006-1010; Lewisら、2002, Nat. Genetics 32:107-108; Rubbinsonら、2003, Nat. Genetics 33:401-406; Tiscorniaら、2003, Proc. Natl. Acad. Sci. USA 100:1844-1848参照、これらのすべては参照することによりその全体が本明細書に組み込まれる）。

【0153】

さらに別の方法ではsiRNAは、動物（例えばヒト）でin vivoで臓器または組織に送達することができる（例えば、Songら、2003, Nat. Medicine 9:347-351; Sorensenら、2003, J. Mol. Biol. 327:761-766; Lewisら、2002, Nat. Genetics 32:107-108参照、これらのすべては参照することによりその全体が本明細書に組み込まれる）。この方法では、siRNAの溶液が動物に静脈内注射される。次にsiRNAは目的の臓器または組織に達し、動物の臓器または組織中で標的遺伝子の発現を有効に低下させる。

【0154】

5.6. 生物学的状態と生物学的応答を測定する方法

10

20

30

40

50

本発明は、siRNAの応答プロファイルを測定する方法を提供する。測定された応答は、細胞または生物中の細胞成分の測定値、またはsiRNAによる摂動に対する細胞または生物の応答でもよい。細胞サンプルは、RNA干渉が起きる任意の生物、例えば真核生物、哺乳動物、霊長類、ヒト、非ヒト動物（例えばイヌ、ネコ、ウマ、ウシ、マウス、ラット、キイロショウジョウバエ (*Drosophila*)、シー・エレガンス (*C. elegans*) など)、植物（例えばイネ、コムギ、マメ、タバコなど）、および真菌のサンプルでもよい。細胞サンプルは、死滅したかまたは健康な生物、または疾患に罹りやすい生物由来でもよい。細胞サンプルは、特定の組織タイプでも成長段階でも、特定のsiRNA摂動に付したサンプルでもよい。このセクションと以後のサブセクションは、細胞サンプルの応答プロファイルを得るための方法のいくつかの例を示す。当業者は、本発明が生物学的システムの発現プロファイルおよび応答を測定するための以下の具体的な方法に限定されないことを理解するであろう。

10

【0155】

5.6.1. マイクロアレイを使用する転写体アッセイ

本発明は、発現プロファイルを追跡することにより、細胞または細胞タイプまたは任意の他の細胞サンプルの発現状態または転写状態を測定するのに特に有用である。本発明の1つの態様は、複数の遺伝子の発現レベルの同時測定のためのポリヌクレオチドプローブアレイ、およびかかるポリヌクレオチドプローブアレイを設計し作成するための方法を提供する。

【0156】

20

遺伝子中のヌクレオチド配列の発現レベルは、任意の大量処理法により測定することができる。いかに測定しても、結果は転写体の絶対量または相対的量、または応答データ（特に限定されないが、量比を示す値を含む）である。

【0157】

好ましくは発現プロファイルの測定は、転写体アレイへのハイブリダイゼーションにより行われる（これはこのサブセクションで説明される）。

【0158】

好適な実施形態において本発明は、「転写体アレイ」または「プロファイリングアレイ」を利用する。転写体アレイは、細胞サンプル中の発現プロファイルを分析するために、および特に特定の組織型の、または成長段階の、または目的のsiRNAに暴露されたか、もしくは目的の生物学的経路への摂動に暴露された細胞サンプルの発現プロファイルを測定するために使用される。別の実施形態において細胞サンプルは、患者、例えば疾患細胞サンプル由来でもよく、好ましくは健康細胞サンプルと比較することができる。

30

【0159】

ある実施形態において発現プロファイルは、細胞中に存在するmRNA転写体中のヌクレオチド配列である検出可能に標識されたポリヌクレオチド（例えば、総細胞mRNAから合成された蛍光標識されたcDNA）をマイクロアレイにハイブリダイズさせることにより得られる。マイクロアレイは、多くの方法で作成することができ、そのいくつかは本明細書に記載される。いかに産生されても、マイクロアレイはいくつかの特徴を共有する。アレイは再現性があり、あるアレイの複数のコピーを産生することが可能であり、容易に互いに比較することができる。好ましくはマイクロアレイは、結合（例えば核酸ハイブリダイゼーション）条件下で安定な材料から作成される。マイクロアレイは好ましくは小さく、例えば約 $1\text{cm}^2 \sim 25\text{cm}^2$ 、好ましくは約 $1 \sim 3\text{cm}^2$ である。しかし大きいアレイも小さいアレイも企図され、例えば大量の異なるプローブを同時評価するためには好ましい。

40

【0160】

好ましくはマイクロアレイ中のある結合部位または結合部位のユニークなセットは、細胞または生物からの単一の遺伝子のヌクレオチド配列（例えば、特異的mRNAまたはそこから得られる特異的cDNA）に特異的に結合（例えばハイブリダイズ）する。

【0161】

本発明の方法と組成物で使用されるマイクロアレイは、1つ以上の試験プローブを含み

50

、その各々が、検出すべきRNAまたはDNAのサブ配列に相補的なポリヌクレオチド配列を有する。各プローブは好ましくは、異なる核酸配列を有し、アレイの固相表面上の各プローブの位置は好ましくは公知である。実際マイクロアレイは好ましくは、アドレス可能なアレイであり、さらに好ましくは位置的にアドレス可能なアレイである。さらに詳しくはアレイの各プローブは好ましくは、アレイ上の位置（すなわち、支持体または表面上）から各プローブの本体（すなわち配列）が決定できるように、固体支持体上の既知の所定の位置にある。本発明のある実施形態において、アレイは規則的なアレイである。

【0162】

好ましくはマイクロアレイまたはマイクロアレイのセット上のプローブの密度は、 1cm^2 当たり約100個の異なる（すなわち同一ではない）プローブまたはそれ以上である。さらに好ましくは本発明の方法で使用されるマイクロアレイは、 1cm^2 当たり少なくとも550個のプローブ、 1cm^2 当たり少なくとも1,000個のプローブ、 1cm^2 当たり少なくとも1,500個のプローブ、または 1cm^2 当たり少なくとも2,000個のプローブを有する。特に好適な例では、マイクロアレイは、好ましくは 1cm^2 当たり少なくとも2,500個のプローブの密度を有する高密度アレイである。従って本発明で使用されるマイクロアレイは、少なくとも2,500、少なくとも5,000、少なくとも10,000、少なくとも15,000、少なくとも20,000、少なくとも25,000、少なくとも50,000、または少なくとも55,000の異なる（すなわち非同一の）プローブを含有する。

10

【0163】

ある実施形態においてマイクロアレイは、各位置が遺伝子によりコードされる転写体のヌクレオチド配列（例えば、mRNAのエキソンまたはそこから得られるcDNA）の別個の結合部位であるアレイ（すなわちマトリックス）である。マイクロアレイ上の結合部位の集合は、複数の遺伝子の結合部位のセットを含有する。例えば種々の実施形態において、本発明のマイクロアレイは、生物のゲノム中の50%より少ない遺伝子によりコードされる生成物の結合部位を含むことができる。あるいは本発明のマイクロアレイは、生物のゲノム中の少なくとも50%、少なくとも75%、少なくとも85%、少なくとも90%、少なくとも95%、少なくとも99%、または100%によりコードされる生成物の結合部位を有することができる。別の実施形態において本発明のマイクロアレイは、生物の細胞により発現される50%より少ない、少なくとも75%、少なくとも85%、少なくとも90%、少なくとも95%、少なくとも99%、または100%の遺伝子によりコードされる生成物の結合部位を有すること

20

30

【0164】

本発明のある実施形態において遺伝子または遺伝子中のエキソンは、遺伝子または遺伝子のエキソンの異なるコード配列セグメントに相補的な異なるポリヌクレオチドとのプローブを含む結合部位のセットによりプロファイリングアレイ中で示される。かかるポリヌクレオチドは好ましくは、15~200塩基の長さ、さらに好ましくは20~100塩基の長さ、最も好ましくは40~60塩基の長さである。各プローブ配列はまた、標的配列に相補的な配列以外にリンカー配列を含有してもよいことは理解されるであろう。本明細書においてリンカー配列は、その標的配列に相補的な配列と支持体の表面との間の配列を意味する。例えば好適な実施形態において、本発明のプロファイリングアレイは各標的遺伝子またはエキソンに特異的な1つのプローブを含む。しかし所望であれば、プロファイリングアレイは、いくつかの標的遺伝子またはエキソンに特異的な少なくとも2、5、10、100、1000プローブを含有してもよい。例えばアレイは、単一塩基工程で遺伝子の最も長いmRNAアイソフォームの配列にわたって重なったプローブを含有してもよい。

40

【0165】

好適な実施形態において、異なる条件からの細胞サンプルのcDNAは、2色プロトコールを使用してマイクロアレイの結合部位にハイブリダイズされる。siRNA応答の場合、1つの細胞サンプルはsiRNAに暴露され、同じタイプの他の細胞サンプルはsiRNAに暴露されな

50

い。経路応答の場合、1つの細胞は経路摂動に暴露され、同じタイプの他の細胞は経路摂動に暴露されない。2つの細胞タイプのそれぞれから得られるcDNAは、区別できるように別々に標識される(例えばCy3とCy5)。例えばある実施形態において、例えばsiRNAで処理した(または経路摂動に暴露した)細胞からのcDNAは、フルオレセイン標識dNTPを使用して合成され、siRNAに暴露されていない第2の細胞からのcDNAは、ローダミン標識dNTPを使用して合成される。2つのcDNAが混合されマイクロアレイにハイブリダイズされると、アレイ上の各部位について各cDNAセットからのシグナルの相対強度が測定され、特定の遺伝子の量の相対的差が検出される。

【0166】

上記の例では、siRNA処理(または経路摂動された)細胞からのcDNAは、蛍光物質が刺激されると緑色の蛍光を出し、未処理細胞からのcDNAは赤色の蛍光を出す。その結果、siRNA処理が細胞中の特定の遺伝子の転写および/または翻訳後スプライシングに対して直接でも間接でも何の影響も無い時、遺伝子および/またはエキソン発現パターンは両方の細胞で区別できず、逆転写により、赤標識および緑標識cDNAは同程度に存在するであろう。マイクロアレイにハイブリダイズさせると、そのRNA種の結合部位は両方の蛍光物質に特徴的な波長の光を放出するであろう。これに対してsiRNAに暴露された細胞が、細胞中の特定の遺伝子の転写および/または翻訳後スプライシングを直接または間接に変化させるsiRNAで処理されると、各遺伝子またはエキソン結合部位の緑と赤の蛍光の比で示される遺伝子および/またはエキソン発現パターンは変化するであろう。siRNAがmRNAの量を上昇させると、mRNA中で発現される各遺伝子またはエキソンの比は上昇するが、siRNAがmRNAの量を低下させると、mRNA中で発現される各遺伝子またはエキソンの比は低下するであろう。

【0167】

遺伝子発現の変化を規定するための2色蛍光標識と検出スキームの使用は、mRNAの検出に関連して記載されており、例えばShenら、1995、相補的DNAマイクロアレイを用いる遺伝子発現パターンの定量的モニタリング、*Science* 270:467-470がある(これは参照することによりその全体が本明細書に組み込まれる)。このスキームは、遺伝子またはエキソンの標識と検出に等しく適用可能である。2つの異なる蛍光物質で標識したcDNAを使用する利点は、2つの細胞状態で各アラインメントした遺伝子に対応するmRNAまたはエキソン発現レベルの直接および内部制御比較ができることであり、実験条件(例えばハイブリダイゼーション条件)の小さな差による変動が以後の分析に影響を与えないことである。しかし、単一の細胞のcDNAを使用して、例えばsiRNA処理または経路摂動細胞および未処理細胞中の特定の遺伝子もしくはエキソンの絶対量を比較することも可能であることは認識されるであろう。さらに3つ以上の色で標識することも本発明で企図される。本発明のある実施形態では、少なくとも5、10、20、または100の異なる色の色素を標識に使用することができる。かかる標識は、同じアレイへの区別できるように標識されたcDNA集団の同時ハイブリダイゼーションを可能にし、こうして3つ以上のサンプルから得られるmRNA分子を測定し、随時その発現レベルを比較することを可能にする。使用される色素には、特に限定されないが、フルオレセインとその誘導体、ローダミンとその誘導体、テキサスレッド、5'-カルボキシ-フルオレセイン(「FMA」)、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(「JOE」)、N,N,N',N'-テトラメチル-6-カルボキシ-ローダミン(「TAMRA」)、6'-カルボキシ-X-ローダミン(「ROX」)、HEX、TET、IRD40、およびIRD41、シアミン色素(特に限定されないが、Cy3、Cy3.5およびCy5を含む)、BODIPY色素(特に限定されないが、BODIPY-FL、BODIPY-TR、BODIPY-TMR、BODIPY-630-650、およびBODIPY-650-670を含む)、およびALEXA色素(特に限定されないが、ALEXA-488、ALEXA-532、ALEXA-546、ALEXA-568、およびALEXA-594を含む)、ならびに当業者に公知の他の蛍光色素がある。

【0168】

本発明のある実施形態において、ハイブリダイゼーションレベルの平衡への進展が測定できるように、複数の異なるハイブリダイゼーション時間にハイブリダイゼーションデー

タが測定される。かかる実施形態においてハイブリダイゼーションレベルは、混合物が実質的に到達した平衡に近く、2本鎖が拡散より親和性と量に依存する濃度であるように、最も好ましくは0から、標識ポリヌクレオチドにより結合したポリヌクレオチド（すなわちプローブ）のサンプリングに必要な時間より多い時間にわたるハイブリダイゼーション時間で測定される。しかしハイブリダイゼーション時間は、好ましくは標識ポリヌクレオチドとプローブおよび/または表面との不可逆的結合相互作用が起きないかまたは少なくとも限定されるように十分に小さいものである。例えば、断片化ポリヌクレオチドの複雑な混合物をプローブ結合するのにポリヌクレオチドアレイが使用される実施形態において、代表的ハイブリダイゼーション時間は約0~72時間である。他の実施態様の適切なハイブリダイゼーション時間は、使用される特定のポリヌクレオチドおよびプローブに依存し、当業者により決定される（例えば、Sambrookら編、1989、「分子クローニング：実験室マニュアル（Molecular Cloning: A Laboratory Manual）」第2版、第1~3巻、コールドスプリングハーバーラボラトリー（Cold Spring Harbor Laboratory）、コールドスプリングハーバー（Cold Spring Harbor）、ニューヨーク参照）。

10

【0169】

ある実施形態において、異なるハイブリダイゼーション時間のハイブリダイゼーションレベルは、異なる、同一のマイクロアレイで別々に測定される。各かかる測定について、ハイブリダイゼーションレベルが測定されるハイブリダイゼーション時間に、マイクロアレイは短時間、好ましくは室温で高~中塩濃度（例えば、0.5~3M塩濃度）の水溶液中で、すべての結合またはハイブリダイズしたポリヌクレオチドを保持するがすべての非結合ポリヌクレオチドを除去する条件下で洗浄される。次に、各プローブ上の残存するハイブリダイズしたポリヌクレオチド分子の検出可能な標識は、使用される特定の標識法に適した方法で測定される。別の実施形態においてハイブリダイゼーションレベルは、単一のマイクロアレイを使用してリアルタイムで測定される。この実施形態では、マイクロアレイは中断することなくサンプルにハイブリダイズされ、マイクロアレイは各ハイブリダイゼーション時間に非侵襲的に信号を送る。さらに別の実施形態において、1つのアレイを使用し、短時間ハイブリダイズし、洗浄し、ハイブリダイゼーションレベルを測定し、同じサンプルに戻り、さらなる時間ハイブリダイズし、洗浄し、再度測定してハイブリダイゼーション時間曲線を得ることができる。

20

【0170】

好ましくは2つの異なるハイブリダイゼーション時間に少なくとも2つのハイブリダイゼーションレベルが測定され、最初のレベルは交差ハイブリダイゼーション平衡の時間スケールに近いハイブリダイゼーション時間で、第2のレベルは第1より長いハイブリダイゼーション時間で測定される。交差ハイブリダイゼーション平衡の時間スケールは、特にサンプル組成とプローブ配列に依存し、当業者により決定される。好適な実施形態において最初のハイブリダイゼーションレベルは1~10時間に測定され、第2のハイブリダイゼーション時間は第1のハイブリダイゼーション時間の約2、4、6、10、12、16、18、48または72倍長い時間に測定される。

30

【0171】

5.6.2. マイクロアレイのプローブの調製

上記したように、本発明の特定のポリヌクレオチド分子（例えば遺伝子またはエキソン）が特異的にハイブリダイズする「プローブ」は、相補的ポリヌクレオチド配列である。好ましくは各標的遺伝子またはエキソンのために1つ以上のプローブが選択される。例えば遺伝子またはエキソンの検出のために最小数のプローブが使用される時は、プローブは通常約40塩基を超える長さのヌクレオチド配列を含む。あるいは、遺伝子またはエキソンのために重複プローブの大きなセットが使用される時、プローブは通常約40~60塩基のヌクレオチド配列を含む。プローブはまた、完全長エキソンに相補的な配列を含むことができる。エキソンの長さは、50塩基未満~200塩基を超える範囲でもよい。すなわち、エキソンより長いプローブ長が使用される時、プローブ長が標的エキソンを含有する連続的mRNA断片に相補的になるように、隣接する構成性にスプライスしたエキソン配列でエキソン

40

50

配列を増強することが好ましい。これは、エキソプロファイリングアレイのプローブ中で類似のハイブリダイゼーションストリンジェンシーを可能にするであろう。各プローブ配列はまた、標的配列に相補的な配列以外に、リンカー配列を含むことは理解されるであろう。

【0172】

プローブは、生物のゲノム中の遺伝子または遺伝子のエキソンの一部に対応するDNAまたはDNA「模倣物」（例えば誘導体と類似体）を含んでよい。ある実施形態においてマイクロアレイプローブは、相補的RNAまたはRNA模倣物である。DNA模倣物は、DNAと特異的ワトソン-クリック様ハイブリダイゼーションまたはRNAとの特異的ハイブリダイゼーションが可能なサブユニットからなるポリマーである。核酸は、塩基成分、糖成分、またはリン酸塩骨格で修飾することができる。DNA模倣物の例には、例えばホスホロチオエートがある。DNAは例えばゲノムDNA、cDNA（例えばRT-PCRにより）、またはクローン化配列からの遺伝子またはエキソンセグメントのポリメラーゼ連鎖反応（PCR）増幅により得ることができる。PCRプライマーは、好ましくはユニークな断片（すなわち、マイクロアレイ上の他の断片と連続的同一配列の10塩基を超えて共有することは無い断片）の増幅を与える遺伝子またはエキソンまたはcDNAの既知の配列に基づいて選択される。必要な特異性と最適増幅性を有するプライマーの設計には、当該分野で公知のコンピュータプログラムが有用であり、例えばOligoバージョン5.0（National Biosciences）がある。代表的にはマイクロアレイ上の各プローブは、20塩基～600塩基であり、通常30～200塩基の長さである。PCR法は当該分野で公知であり、例えばInnisら編、1990、「PCRプロトコル：方法と応用のガイド（PCR Protocols: A Guide to Methods and Applications）」、アカデミックプレスインク（Academic Press inc.）（サンジエゴ、カリフォルニア州）に記載されている。核酸を単離し増幅するのに制御されたロボットシステムが有用であることは、当業者に明らかであろう。

【0173】

マイクロアレイのポリヌクレオチドプローブを作成するための別の好適な手段は、N-ホスホネートまたはホスホラミダイト化学を使用する合成ポリヌクレオチドまたはオリゴヌクレオチドの合成である（Froehlerら、1986、Nucleic Acids Res. 14:5399-5407；McBrideら、1983、Tetrahedron Lett. 24:246-248）。合成配列は代表的には、約15～約600塩基の長さ、さらに代表的には約20～約100塩基、最も好ましくは約40～約70塩基の長さである。ある実施形態において合成核酸は、非天然の塩基（例えば特に限定されないが、イノシン）を含む。上記したように核酸類似体は、ハイブリダイゼーションの結合部位として使用される。適当な核酸類似体の例はペプチド核酸である（例えば、Egholmら、1993、Nature 363:566-568；US Patent No. 5,539,083参照）。

【0174】

別の実施態様においてハイブリダイゼーション部位（すなわちプローブ）は、遺伝子のプラスミドまたはファージクローン、cDNA（例えば発現された配列タグ）、またはその挿入体から作成される（Nguyenら、1995、Genomics 29:207-209）。

【0175】

5.6.3. 固体表面へのプローブの結合

あらかじめ生成されたポリヌクレオチドプローブは、支持体上に沈積させてアレイを形成することができる。あるいはポリヌクレオチドプローブは直接支持体上で合成してアレイを形成することができる。プローブは、ガラス、プラスチック（例えば、ポリプロピレン、ナイロン）、ポリアクリルアミド、ニトロセルロース、ゲル、または他の多孔性もしくは非多孔性材料から作成された固体支持体または表面上に結合される。

【0176】

核酸を表面に結合させるための好適な方法は、Schenら、1995、Science 270:467-470に一般的に記載されたようにガラスプレート上へのプリントである。この方法は、cDNAのマイクロアレイを調製するのに特に有用である（また、DeRisiら、1996、Nature Genetics 14:457-460；Shalonら、1996、Genome Res. 6:639-645；およびSchenら、1995、Proc

10

20

30

40

50

. Natl. Acad. Sci. USA 93:10539-11286参照)。

【0177】

マイクロアレイを作成するための第2の好適な方法は、高密度ポリヌクレオチドアレイを作成することによる。in situ合成のための写真平版法(Fodorら、1991, Science 251:767-773; Peaseら、1994, Proc. Natl. Acad. Sci. USA 91:5022-5026; Lockhartら、1996, Nature Biotechnology 14:1675; US Patent No. 5,578,832; 5,556,752; および5,510,270参照)、または規定のオリゴヌクレオチドの迅速な合成と沈積のための他の方法(Blanchardら、Biosensors & Bioelectronics 11:687-690)を使用して、規定の位置に規定の配列に相補的な数千のオリゴヌクレオチドを含有するアレイを製造するための方法は公知である。これらの方法が使用される時、既知の配列のオリゴヌクレオチド(例えば60量体)は、誘導体化ガラススライドのような表面で直接合成される。産生されたアレイは重複してもよく、1つの遺伝子もしくはエキソン当たりいくつかのポリヌクレオチド分子がある。

10

【0178】

例えばマスクング(MaskosとSouthern、1992, Nucleic Acids Res. 20:1679-1684)によりマイクロアレイを作成するための他の方法も使用される。原理的にかつ上記したように、任意のタイプのアレイ、例えばドットプロットまたはナイロンハイブリダイゼーション膜(Sambrookら参照、前述)を使用してもよい。しかし当業者には明らかなように、ハイブリダイゼーション容量が小さくなるために、非常に小さなアレイがしばしば好適である。

20

【0179】

特に好適な実施形態において本発明のマイクロアレイは、オリゴヌクレオチド合成のためのインクジェット印刷装置を使用して、例えばBlanchardの国際特許公報W098/41531(1998年9月24日公開); Blanchardら、1996, Biosensors & Bioelectronics 11:687-690; Blanchard, 1998, Synthetic DNA Arrays in Genetic Engineering, Vol.20, J.K. Setlow編、プレナム・プレス(Plenum Press)、ニューヨーク、111~123頁; およびBlanchardのUS Patent No. 6,028,1889に記載の方法とシステムを使用して製造される。具体的にはかかるマイクロアレイ中のポリヌクレオチドプローブは好ましくはアレイ(例えばガラススライド上)で、炭酸プロピレンのような高表面張力溶媒の「微小液滴」中に個々のヌクレオチド塩基を連続して沈積することにより合成される。微小液滴は容量が小さく(例えば、100pLまたはそれ以下、さらに好ましくは50pLまたはそれ以下)、マイクロアレイ上で互いに分離されて、アレイ要素(すなわち異なるプローブ)の位置を画定する円形表面張力ウェルを形成する。ポリヌクレオチドプローブは通常、ポリヌクレオチドの3'末端で表面に共有結合される。あるいは、ポリヌクレオチドプローブは、ポリヌクレオチドの5'末端で表面に共有結合される(例えば、Blanchard, 1998, Synthetic DNA Arrays in Genetic Engineering, Vol.20, J.K. Setlow編、プレナム・プレス(Plenum Press)、ニューヨーク、111~123頁参照)。

30

【0180】

5.6.4. ポリヌクレオチド分子を標的とする

本発明の方法と組成物により分析される標的ポリヌクレオチドには、RNA分子、例えば特に限定されないが、メッセンジャーRNA(mRNA)分子、リボゾームRNA(rRNA)分子、cRNA分子(すなわち、in vivoで転写されるcDNA分子から調製されるRNA分子)、およびこれらの断片がある。本発明の方法と組成物により分析される標的ポリヌクレオチドにはまた、特に限定されないがDNA分子、例えばゲノムDNA分子、cDNA分子、およびこれらの断片(オリゴヌクレオチド、EST、STSなどを含む)がある。

40

【0181】

標的ポリヌクレオチドは任意の供給源から得られる。例えば標的ポリヌクレオチド分子は、天然に存在する核酸分子、例えば生物から単離されるDNA分子もしくはゲノム外DNA分子、または生物から単離されるmRNA分子のようなRNA分子でもよい。あるいはポリヌクレオチド分子は合成され、例えばin vivoもしくはin vitroで酵素的に合成される核酸分子

50

、例えばcDNA分子、またはPCRにより合成されるポリヌクレオチド分子、*in vitro*転写により合成されるRNA分子などがある。標的ポリヌクレオチドのサンプルは、例えばDNA、RNA、またはDNAとRNAのコポリマーの分子を含んでよい。好適な実施形態において本発明の標的ポリヌクレオチドは、特定の遺伝子または特定の遺伝子転写体（例えば、細胞中で発現されるmRNA配列、またはかかるmRNA配列から得られる特定のcDNA配列）に対応する。しかし多くの実施形態、特にポリヌクレオチド分子が哺乳動物細胞から得られる実施形態において、標的ポリヌクレオチドは遺伝子転写体の特定の断片に対応してもよい。例えば標的ポリヌクレオチドは、その遺伝子の異なるスプライス変種が検出されおよび/または分析されるように、同じ遺伝子の異なるエキソンに対応してもよい。

【0182】

好適な実施形態において分析される標的ポリヌクレオチドは、細胞から抽出された核酸から*in vitro*で調製される。例えばある実施形態においてRNAは、細胞から抽出され（例えば、総細胞RNA、ポリ(A)⁺メッセンジャーRNA、これらの断片）、メッセンジャーRNAは総抽出RNAから精製される。総およびポリ(A)⁺RNAを調製する方法は当該分野で公知であり、例えばSambrookら（前述）に一般的に記載されている。ある実施形態においてRNAは、本発明の目的の種々のタイプの細胞から、チオシアン酸グアニジン溶解し、次にCsCl遠心分離しオリゴdT精製することにより抽出される（Chirgwinら、1979, *Biochemistry* 18: 5294-5299）。別の実施形態においてRNAは、チオシアン酸グアニジン溶解し、次にRNeasyカラム（キアゲン（Qiagen））で精製して抽出される。次に精製されたmRNAから、例えばオリゴdTまたはランダムプライマーを使用して精製mRNAからcDNAが合成される。好適な実施形態において標的ポリヌクレオチドは、細胞から抽出された精製されたメッセンジャーRNAから調製されたcRNAである。本明細書においてcRNAは、供給源RNAに相補的なRNAとして定義される。抽出されたRNAは、アンチセンスRNAの転写を指令することができる方向でRNAポリメラーゼプロモーターに結合したプライマーを使用してRNAから2本鎖cDNAが合成される方法を使用して、増幅される。次にアンチセンスRNAまたはcRNAは、RNAポリメラーゼを使用して2本鎖cDNAの第2鎖から転写される（例えば、US Patent No. 5,891,636、5,716,785；5,545,522および6,132,997；また米国特許出願第09/411,074（1999年10月4日出願、LinsleyとSchelter）、およびPCT公報W002/44399参照）。RNAポリメラーゼプロモーターまたはその相補体を含むオリゴdTプライマー（US Patent No. 5,545,522と6,132,997）およびランダムプライマー（PCT公報W002/44399）の両方が使用できる。好ましくは標的ポリヌクレオチドは、細胞の元々の核酸集団を代表する短いおよび/または断片化したポリヌクレオチド分子である。

【0183】

本発明の方法と組成物により分析される標的ポリヌクレオチドは、好ましくは検出できるように標識される。例えばcDNAは、直接、例えばヌクレオチド類似体により、または間接に、例えば第1鎖を鋳型として使用して第2の標識cDNA鎖を作成することにより標識することができる。あるいは2本鎖cDNAをcRNAに転写し標識することができる。

【0184】

好ましくは検出可能な標識物は、例えばヌクレオチド類似体の取り込みによる蛍光標識物である。本発明での使用に適した他の標識物には、特に限定されないが、ビオチン、イミノビオチン、抗原、補助因子、ジニトロフェノール、リポ酸、オレフィン化合物、検出可能なポリペプチド、電子リッチな分子、基質への作用により検出可能なシグナルを生成することができる酵素、および放射活性アイソトープがある。好適な放射活性アイソトープには、³²P、³⁵S、¹⁴C、¹⁵N、および¹²⁵Iがある。本発明に適した蛍光分子には、特に限定されないが、フルオレセインおよびその誘導体、ローダミンおよびその誘導体、テキサスレッド、5'-カルボキシ-フルオレセイン（「FMA」）、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（「JOE」）、N,N,N',N'-テトラメチル-6-カルボキシ-ローダミン（「TAMRA」）、6'-カルボキシ-X-ローダミン（「ROX」）、HEX、TET、IRD40、およびIRD41がある。本発明の適した蛍光分子にはさらに、シアミン色素（特に限定されないが、Cy3、Cy3.5およびCy5を含む）、BODIPY色素（特に限定されないが、BODIPY-FL

10

20

30

40

50

、BODIPY-TR、BODIPY-TMR、BODIPY-630/650、およびBODIPY-650/670を含む)、およびALEXA色素(特に限定されないが、ALEXA-488、ALEXA-532、ALEXA-546、ALEXA-568、およびALEXA-594を含む)、ならびに当業者に公知の他の蛍光色素がある。本発明に適した電子リッチな指標分子には、特に限定されないが、フェリチン、ヘモシアニン、およびコロイド金がある。あるいはやや好適な実施形態では、標的ポリヌクレオチドは第1の基をポリヌクレオチドに特異的に複合体化することにより標識してもよい。指標分子に共有結合し第1の基に対して親和性を有する第2の基は、標的ポリヌクレオチドを間接的に検出するのに使用することができる。かかる実施形態において第1基として使用するのに適した化合物には、特に限定されないがピオチンとイミノピオチンがある。第2基としての使用に適した化合物には、特に限定されないがアビジンおよびストレプトアビジンがある。

10

【0185】

5.6.5. マイクロアレイへのハイブリダイゼーション

上記したように核酸ハイブリダイゼーションと洗浄条件は、本発明により分析されるポリヌクレオチド分子(本明細書において「標的ポリヌクレオチド分子」と呼ぶ)が、その相補的DNAが位置するアレイの相補的ポリヌクレオチド配列(好ましくは特異的アレイ部位)に、特異的に結合または特異的にハイブリダイズするように選択される。

【0186】

その上に2本鎖プローブDNAを含有するアレイは好ましくは、標的ポリヌクレオチド分子との接触前にDNAを1本鎖にするために変性条件に付される。1本鎖プローブDNA(例えば、合成オリゴデオキシリボ核酸)を含有するアレイは、例えば自己相補配列のために形成するヘアピンまたはダイマーを除去するために、標的ポリヌクレオチド分子に接触する前に、変性されることが必要である。

20

【0187】

最適ハイブリダイゼーション条件は、プローブと標的核酸の長さ(例えば、200塩基より大きいオリゴマー対ポリヌクレオチド)とタイプ(例えばRNAまたはDNA)に依存するであろう。核酸の特異的(すなわち厳密)ハイブリダイゼーション条件の一般的パラメータは、Sambrookら(前述)とAusubelら、1987、「分子生物学の現代のプロトコール(Current Protocols in Molecular Biology)」、Greene Publishing and Wiley-Interscience、ニューヨークに記載されている。SchenららのcDNAマイクロアレイが使用される時、代表的なハイブリダイゼーション条件は、5×SSC+0.2% SDS中65ECで4時間、次に25ECで低ストリンジェンシーバッファ(1×SSC+0.2% SDS)で洗浄、次に25ECで高ストリンジェンシーバッファ(0.1×SSC+0.2% SDS)で10分洗浄である(Shenら、1996, Proc. Natl. Acad. Sci. USA 93:10614)。有用なハイブリダイゼーション条件はまた、例えばTijssen, 1993, 「核酸プローブとのハイブリダイゼーション(Hybridization With Nucleic Acid probes)」、Elsevier Science Publishers B.V. and Kricka, 1992, 「非アイソトープDNAプローブ技術(Nonisotopic DNA Probe Techniques)」、アカデミックプレス(Academic Press)、サンジエゴ、カリフォルニア州に記載されている。

30

【0188】

本発明のスクリーニングおよび/またはシグナル伝達チップで使用するのに特に好適なハイブリダイゼーション条件は、プローブの平均融点またはその近くの温度(例えば、5

40

【0189】

5.6.6. シグナル検出とデータ解析

細胞のRNAに相補的な標的配列(例えばcDNAまたはcRNA)が作成され、適当なハイブリダイゼーション条件下でマイクロアレイにハイブリダイズされる時、遺伝子または特定の遺伝子のエキソンに対応するアレイ中の部位へのハイブリダイゼーションのレベルが、その遺伝子から転写されたエキソンを含有するmRNAの細胞中での存在量を反映することは理解されるであろう。例えば、総細胞mRNAに相補的な検出できるように標識された(例えば蛍光物質で)cDNAがマイクロアレイにハイブリダイズする時、細胞中のRNAスプライシン

50

グ中に転写されないかまたは除去される遺伝子のエキソン（すなわち、遺伝子発現の生成物に特異的に結合することができる）に対応するアレイ上の部位は、ほとんどまたは全くシグナル（例えば蛍光シグナル）を持たず、エキソンを発現するコードされたmRNAが多い遺伝子のエキソンが比較的強いシグナルを有するであろう。次に、代替スプライシングにより同じ遺伝子から産生される異なるmRNAの相対量は、遺伝子について追跡されるエキソンの全セットにわたってシグナル強度パターンにより決定される。

【0190】

好適な実施形態において、2つの異なる細胞からの標的配列（例えば、cDNAまたはcRNA）は、マイクロアレイの結合部位にハイブリダイズされる。siRNA応答の場合、1つの細胞サンプルはsiRNAに暴露され、同じタイプの他の細胞サンプルはsiRNAに暴露されない。経路応答の場合、1つの細胞サンプルは経路摂動に暴露され、同じタイプの他の細胞サンプルは経路摂動に暴露されない。2つの細胞タイプのそれぞれから得られるcDNAまたはcRNAは、区別できるようにするために鑑別的に標識される。ある実施形態において、例えばsiRNAで処理された（または経路摂動に暴露された）細胞からのcDNAは、フルオレセイン標識dNTPを使用して合成され、第2の細胞からのcDNA（siRNAに暴露されていない）はローダミン標識dNTPを使用して合成される。2つのcDNAが混合されマイクロアレイにハイブリダイズされると、各cDNAからのシグナルの相対強度がアレイ上の各部位について測定され、特定のエキソンの量の相対的差が検出される。

【0191】

上記した例では、siRNA処理（または経路摂動された）細胞からのcDNAは、蛍光物質が刺激されると緑色の蛍光を出し、未処理細胞からのcDNAは赤色の蛍光を出す。その結果、siRNA処理が細胞中の特定の遺伝子の転写および/または翻訳後スプライシングに対して直接でも間接でも何の影響も無い時、遺伝子および/またはエキソン発現パターンは両方の細胞で区別できず、逆転写により、赤標識および緑標識cDNAは同程度に存在するであろう。マイクロアレイにハイブリダイズさせると、そのRNA種の結合部位は両方の蛍光物質に特徴的な波長の光を放出するであろう。これに対してsiRNAに暴露された細胞が、細胞中の特定の遺伝子の転写および/または翻訳後スプライシングを直接または間接に変化させるsiRNAで処理されると、各遺伝子またはエキソン結合部位の緑と赤の蛍光の比で示される遺伝子および/またはエキソン発現パターンは変化するであろう。siRNAがmRNAの量を上昇させると、mRNA中で発現される各遺伝子またはエキソンの比は上昇するが、siRNAがmRNAの量を低下させると、mRNA中で発現される各遺伝子またはエキソンの比は低下するであろう。

【0192】

遺伝子発現の変化を規定するための2色蛍光標識と検出スキームの使用は、mRNAの検出に関連して記載されており、例えばShenら、1995、相補的DNAマイクロアレイを用いる遺伝子発現パターンの定量的モニタリング、Science 270:467-470がある（これは参照することによりその全体が本明細書に組み込まれる）。このスキームは、エキソンの標識と検出に等しく適用可能である。2つの異なる蛍光物質で標識した標的配列（例えばcDNAまたはcRNA）を使用する利点は、2つの細胞状態で各アラインメントした遺伝子に対応するmRNAまたはエキソン発現レベルの直接および内部制御比較ができることであり、実験条件（例えばハイブリダイゼーション条件）の小さな差による変動が以後の分析に影響を与えないことである。しかし、単一の細胞のcDNAを使用して、例えばsiRNA処理または経路摂動細胞および未処理細胞中の特定の遺伝子もしくはエキソンの絶対量を比較することも可能であることは認識されるであろう。

【0193】

別の好適な実施形態において、例えば1色蛍光標識を使用するシングルチャンネル検出法が使用される（米国仮特許出願第60/227,966号、2000年8月25日出願）。この実施形態において逆相補体（RC）プローブを含むアレイが設計され作成される。DNA配列の逆相補体は、種々の尺度（例えば、GC含量およびGCトレンドのような尺度は逆相補体下では不変である）について標的配列に相補的な対応する前進鎖（FS）プローブに等しい配列複雑さを

10

20

30

40

50

有するため、対応するFSプローブへの非特異的クロスハイブリダイゼーションのレベルの測定のための対照プローブとしてRCプローブが使用される。標的配列のFSプローブ強度の有意性は、各測定誤差とともにFSプローブの生強度測定値とRCプローブの対応する生強度測定値を比較することにより決定される。好適な実施形態において、FSプローブと対応するRCプローブとの強度差が有意なら、遺伝子またはエキソンが存在すると言われる。さらに好ましくは、FSプローブ強度がまたバックグラウンドレベルより有意に高い場合、遺伝子またはエキソンが存在すると言われる。シングルチャンネル検出法は多色標識とともに使用することができる。ある実施形態において複数の異なるサンプル（それぞれ異なる色で標識されている）がアレイにハイブリダイズされる。各色についてのFSとRCプローブとの差は、対応するサンプルのハイブリダイゼーションレベルを測定するのに使用される。

10

【0194】

蛍光標識されたプローブが使用される時、転写体アレイの各部位での蛍光放出は、共焦点レーザー走査顕微鏡により検出することができる。ある実施形態において、使用される2つの蛍光物質のそれぞれについて、適切な励起線を使用して別のスキャンが行われる。あるいは、2つの蛍光物質に特異的な波長でサンプルを同時照射することを可能にするレーザーを使用することができ、2つの蛍光物質からの発光が同時に分析される（Shalonら、1996, *Genome Res.* 6:639-645）。好適な実施形態においてアレイは、コンピュータ制御されたX-Yステージと顕微鏡対物レンズを有するレーザー蛍光スキャナーを用いてスキャンされる。多重線混合ガスレーザーを用いる2つの蛍光物質の連続励起が行われ、2つの光電子増倍管により検出される。かかる蛍光レーザースキャニング装置は、例えばSchenら、1996, *Genome Res.* 6:639-645に記載されている。あるいは、Fergusonら、1996, *Nature Biotech.* 14:1681-1684に記載の光ファイバー束を使用して、多数の部位のmRNAレベルを同時に測定してもよい。

20

【0195】

シグナルは記録され、好適な実施形態では、例えば12ビットアナログデジタル基板を使用してコンピュータで解析される。ある実施形態においてスキャンされた画像は、グラフィクスプログラム（例えばHijaak Graphics Suite）を使用して斑点の除去が行われ、次に画像グリッドプログラムを使用して解析されて、各波長で各部位で平均ハイブリダイゼーションのスプレッドシートが作成される。必要であれば、2つの蛍光物質のチャンネル間の「クロストーク」（または重複）について実験的に決定された補正が行われる。転写体アレイ上の任意の具体的なハイブリダイゼーション部位について、2つの蛍光物質の発光の比が計算される。この比は、同系遺伝子の絶対発現レベルには依存しないが、その発現がsiRNAトランスフェクション、遺伝子欠失、または他の試験されたイベントで有意に調節される遺伝子に有用である。

30

【0196】

本発明の方法に従って、2つの細胞または細胞株中のmRNAおよび/またはmRNA中で発現されたエキソンの相対的量は、摂動した（すなわち、試験したmRNAの2つの供給源で量は異なる）か、または摂動しない（すなわち、相対量は同じである）としてスコアが付けられる。本明細書において、RNAの2つの供給源間の差が少なくとも約25%（すなわち、ある供給源中のRNAが他の供給源中より25%多い）、さらには約50%、さらには約2倍（すなわち2倍多い）、3倍（3倍多い）、または5倍（5倍多い）多いものは、摂動としてスコアが付けられる。本検出法は、約3倍～約5倍のオーダーの差の信頼できる検出を可能にするが、より高感度の方法が開発されると期待される。

40

【0197】

しかし、2つの細胞または2つの細胞株中のmRNAおよび/またはmRNA中で発現されるエキソンの量の相対差の強度を測定することも有用である。これは上記したように、鑑別標識に使用された2つの蛍光物質の発光の比を計算するか、または当業者に容易に明らか類似の方法により行うことができる。

【0198】

5.6.7. 転写状態測定の方法

50

当該分野で公知の他の遺伝子発現法により、細胞の転写状態を測定してもよい。いくつかのかかる技術は、電気泳動分析のための限定的な複雑さの制限断片のプールを作成し、例えば2重制限酵素消化を段階的プライマーと組合せた方法（例えば、ヨーロッパ特許0534858A1、1992年9月24日出願、Zabeauら）、または規定のmRNA末端に最も近い部位を有する制限断片を選択する方法（例えば、Prasharら、1996、Proc. Natl. Acad. Sci. USA 93:659-663）。他の方法は、例えば複数のcDNAのそれぞれの十分な塩基（例えば20～50塩基）を配列決定して各cDNAを同定する、または規定のmRNA末端に対して既知の位置で生成される短いタグ（例えば9～10塩基）を配列決定することにより、cDNAプールを統計的にサンプリングする（例えば、Velculescu、1995、Science 270:484-487）。

【0199】

10

5.7. 生物学的状態の他の態様の測定

本発明の種々の実施形態において、転写状態以外の生物学的状態の態様（翻訳状態、活性状態、または混合態様）を測定して、本発明で分析され測定されたシグナルを生成することができる。すなわちかかる実施形態において、遺伝子発現データは、翻訳状態測定値またはタンパク質発現測定値を含んでよい。実際、ある実施形態において、遺伝子発現に基づく遺伝子発現相互作用地図を使用するより、タンパク質発現地図に基づくタンパク質相互作用地図が使用される。転写状態以外の生物学的状態の態様の詳細が本セクションで説明される。

【0200】

5.7.1. 翻訳状態測定値に基づく実施形態

20

翻訳状態の測定はいくつかの方法により行われる。例えばマイクロアレイを構築することにより、タンパク質の全ゲノムモニタリング（すなわち「プロテオーム」、Goffeauら、1996、Science 274:546-567；Aebersoldら、1999、Nature Biotechnology 10:994-999）を行うことができ、ここで結合部位は、細胞ゲノムにコードされる複数のタンパク質種に特異的な固定化された（好ましくはモノクローナル）抗体を含む（例えば、Zhuら、2001、Science 293:2101-2105；MacBeathら、2000、Science 289:1760-63；de Wildtら、2000、Nature Biotechnology 18:989-994参照）。好ましくは、コードされたタンパク質、または少なくとも目的のsiRNAの作用に関連するモノクローナル抗体の実質的な割合について抗体が存在する。モノクローナル抗体を作成する方法は公知である（例えば、ハーロー（Harlow）とレーン（Lane）、1988、「抗体：実験室マニュアル（Antibodies: A Laboratory Manual）」、コールドスプリングハーバー（Cold Spring Harbor）、ニューヨーク、これは参照することによりその全体が本明細書に組み込まれる）。好適な実施形態において、細胞のゲノム配列に基づいて設計された合成ペプチド断片に対してモノクローナル抗体が作成される。かかる抗体アレイを使用して、細胞からのタンパク質はアレイに接触され、その結合が当該分野で公知のアッセイにより測定される。

30

【0201】

あるいはタンパク質は、2次元ゲル電気泳動システムにより分離し測定することができる。2次元ゲル電気泳動は当該分野で公知であり、代表的には1次元に沿って等電点電気泳動を行い、次に2次元に沿ってSDS-PAGE電気泳動を行う。例えば、Hamesら、1990、「タンパク質のゲル電気泳動：実際的アプローチ（Gel Electrophoresis of Proteins: A Practical Approach）」、アイアールエルプレス（IRL Press）、ニューヨーク；Shevchenkoら、1996、Proc. Natl. Acad. Sci. USA 93:1440-1445；Saglioccoら、1996、Yeast 12:1519-1533；Lander、1996、Science 274:536-539；およびBeaumontら、Life Science News 7、2001、アマシャム・ファルマシア・バイオテック（Amersham Pharmacia Biotech）。生じる電気泳動図は、多くの方法（ポリクローナル抗体およびモノクローナル抗体を使用する質量スペクトル法、ウェスタンブロッティング法、および免疫プロット分析、および内部とN末端マイクロ配列決定を含む）により分析することができる。これらの方法を使用して、siRNAに暴露された細胞（例えば酵母）中、または特異的遺伝子の欠失または過剰発現により修飾された細胞中の、ある生理学的条件下で産生されるすべてのタンパク質の実質的な割合を同定することができる。

40

50

る。コンピュータシステムはまた代表的にはネットワークリンク807に連結され、これは例えば、ローカルエリアネットワーク（「LAN」）と他のローカルコンピュータシステム、および/またはワイドエリアネットワーク（「WAN」）、例えば他の遠くのコンピュータシステムに連結されたインターネットでもよい。例えば好適な実施形態において、上記したように、コンピュータシステムが複数のノードを含む場合、各ノードは好ましくはネットワーク（好ましくはNFSネットワーク）に連結し、その結果コンピュータシステムのノードは互いに通信し、随時他のコンピュータシステムとネットワークにより通信し、従ってデータを共有し、互いにタスクを処理する。

【0208】

かかるコンピュータシステムの操作中にメモリーにいくつかのソフトウェア成分がロードされ、これはまた図8に示す。ソフトウェア成分は、当該分野で標準的なソフトウェア成分と本発明に特殊な成分とを含む。これらのソフトウェア成分は代表的には、大容量記憶装置例えばハードドライブ804に保存されるが、他のコンピュータで読める媒体、例えば1つ以上のフロッピーディスク、1つ以上のCD-ROMドライブ、1つ以上のDVDドライブ、または1つ以上のDATドライブを含む）に保存することができる。ソフトウェア成分810は、コンピュータシステムとそのネットワーク接続の管理を行うオペレーティングシステムである。オペレーティングシステムは、例えばマイクロソフトウィンドウズ（登録商標）ファミリー、例えばウィンドウズ95、ウィンドウズ98、ウィンドウズNT、ウィンドウズ2000またはウィンドウズXPでもよい。あるいはオペレーティングソフトウェアはマッキントッシュオペレーティングシステム、UNIXオペレーティングシステム、またはLINUXオペレーティングシステムでもよい。ソフトウェア成分811は、共通の言語とシステム中に存在することが好ましい機能を含んで、本発明に特異的な方法を実施するプログラムを支援する。本発明の分析方法をプログラムするのに使用できる言語には、例えばCとC++、フォートラン、PERL、HTML、JAVA、および任意のUNIXまたはLINUXシェル命令言語（例えばCシェルスクリプト言語）がある。本発明の方法はまた、式の記号エントリおよびプロセッシングの高レベル指定を可能にする数学的ソフトウェアパッケージ（使用される特異的アルゴリズムを含む）でプログラム化またはモデル化され、こうしてユーザーが個々の式およびアルゴリズムをプログラムする手続きを不要にする。かかるパッケージには、例えばMathworks（Natick、マサチューセッツ州）からのMatlab、Wolfram Research（Champaign、イリノイ州）からのMathematica、またはMathSoft（シアトル、ワシントン州）からのS-Plusがある。

10

20

30

【0209】

ソフトウェア成分812は、好ましくは操作言語または記号パッケージでプログラムされた前記した本発明の分析方法を含む。例えば、ソフトウェア成分812は好ましくは、プロセッサに複数の測定したシグナルを受け、測定したシグナルをメモリーに保存する工程を実施させるプログラムを含む。例えばコンピュータシステムは、ユーザーが手動的に入力した（例えばユーザーインターフェースにより）測定されたシグナルを受けることができる。しかしさらに好ましくは、プログラムはコンピュータシステムにデータベースから測定されたシグナルを検索させる。かかるデータベースは、大容量記憶装置（例えばハードドライブ）または他のコンピュータで読める媒体に保存することができ、コンピュータのメモリーにロードすることができ、またはネットワーク807を使用してコンピュータシステムによりデータベースにアクセスすることができる。

40

【0210】

上記のプログラム例とコンピュータシステム以外に、他の代替プログラム構造とコンピュータシステムは当業者に容易に明らかであろう。上記コンピュータシステムおよびプログラムの精神または範囲から逸脱することなくかかる代替システムは、添付の特許請求の範囲に包含されると企図される。

【実施例】

【0211】

6. 実施例

50

以下の例は本発明を例示するために提供され、決して本発明を限定するものではない。

【0212】

実施例1

6.1. 実施例1：発現プロファイリングはsiRNAの遺伝子制御を明らかにする

RNA干渉 (RNAi) またはRNAサイレンシングは、遺伝子発現を抑制するのに広く使用されている。RNAサイレンシングは、分解のために同族のmRNA種を標的とする21ntの短い干渉性RNA (siRNA) により開始される。サイレンシングは一般に極めて特異的であり、siRNAとその同種のmRNAとの間でほぼ完璧な相同性を要求すると考えられている。この例では、遺伝子発現プロファイリングを使用して、培養ヒト細胞中のsiRNAによる遺伝子サイレンシングの特異性を解析する。この解析は、異なるsiRNAがユニークなセットの遺伝子を抑制し、その中には目的の標的または標的サイレンシングの程度に無関係の遺伝子があることを明らかにした。これらの転写プロファイルは、siRNAに対してわずかに11の連続ヌクレオチドの相同性を含有する非標的遺伝子の直接サイレンシングを明らかにした。すなわち、この結果は、内因性遺伝子のサイレンシングが、予想されたよりsiRNAとのはるかに少ない相同性が必要なだけであり、siRNAは完全長ではない配列類似性を有する非標的遺伝子のmRNAと交差反応することを証明する。これらの結果は、RNAiによる遺伝子サイレンシングから生じる表現型へのオフターゲット遺伝子制御の寄与を考慮する必要があることを確認する。

10

【0213】

発現プロファイリングは、その発現が、生存シグナル伝達経路に關与する2つのタンパク質、MAPK14 (p38) とインスリン様増殖因子受容体 (IGF1R) のHeLa細胞中のRNAi誘導サイレンシングの結果として変化した遺伝子を同定するのに使用される。21ヌクレオチドの2本鎖siRNAを、標準的選択規則に従ってこれらの遺伝子のそれぞれのコード領域を標的とするように設計 (Elbashirら、2002, Methods 26:199-213) し、配列をBLAST解析に付して、ヒトゲノム中の他の遺伝子との大きな相同性を除外した。以前の報告は、siRNAと転写体との単一ヌクレオチドミスマッチがサイレンシングを排除することを示した (Elbashirら、2001, EMBO. J. 20:6877-6888)。この例では、標的遺伝子以外の遺伝子とは18未満のヌクレオチドの相同性を示した配列のみを選択した。MAPK14を標的とするように設計した8つのsiRNAとIGF1Rを標的とするように設計した16個のsiRNAを、個々にHeLa細胞中にトランスフェクトした。siRNAトランスフェクション後に、RNAを単離し、マイクロアレイ分析に付した (Hughesら、2001, Nature Biotech. 19:342-347)。遺伝子発現パターンの共通の変化を同定するために、異なるsiRNAにより同じ標的遺伝子のサイレンシングから生じる発現プロファイルを比較した。

20

30

【表 1】

表 I. MAPK14 および IGF1R siRNA 配列

| siRNA | 配列(センス鎖) | |
|--------------|--------------------------------------|----|
| MAPK14-1 | CCUACAGAGAACUGCGGUU-dTdT (配列番号 1) | |
| Pos.4 ミスマッチ | CCUGCAGAGAACUGCGGUU-dTdT (配列番号 2) | |
| Pos.5 ミスマッチ | CCUAAAGAGAACUGCGGUU-dTdT (配列番号 3) | |
| Pos.15 ミスマッチ | CCUACAGAGAACUGAGGUU-dTdT (配列番号 4) | 10 |
| MAPK14-2 | AUGUGAUUGGUCUGUUGGA-dTdT (配列番号 5) | |
| MAPK14-3 | UUCUCCGAGGUCUAAAGUA-dTdT (配列番号 6) | |
| MAPK14-4 | UAAUUCACAGGGACCUAAA-dTdT (配列番号 7) | |
| MAPK14-5 | CCAGUGGCCGAUCCUUAUG-dTdT (配列番号 8) | |
| MAPK14-6 | UGCCUACUUUGCUCAGUAC-dTdT (配列番号 9) | |
| MAPK14-7 | GUCAUCAGCUUUGUGCCAC-dTdT (配列番号 10) | |
| MAPK14-8 | GGCCUUUUCACGGGAACUC-dTdT (配列番号 11) | |
| IGF1R-1 | GCUCACGGUCAUUACCGAG-dTdT (配列番号 12) | 20 |
| IGF1R-2 | CCUGAGGAACAUUACUCGG-dTdT (配列番号 13) | |
| IGF1R-3 | UGCUGACCUCUGUUACCUC-dTdT (配列番号 14) | |
| IGF1R-4 | CGACACGGCCUGUGUAGCU-dTdT (配列番号 15) | |
| IGF1R-5 | GAUGAUUCAGAUGGCCGGA-dTdT (配列番号 16) | |
| IGF1R-6 | CUUGCAGCAACUGUGGGAC-dTdT (配列番号 17) | |
| IGF1R-7 | CCUCACGGUCAUCCGCGGC-dTdT (配列番号 18) | |
| IGF1R-8 | CUACGCCUGGUCAUCUUC-dTdT (配列番号 19) | |
| IGF1R-9 | UCUCAAGGAUUAUUGGCCUU-dTdT (配列番号 20) | 30 |
| IGF1R-10 | GGAUUAUUGGGCUUUAACAAC-dTdT (配列番号 21) | |
| IGF1R-11 | CAUUACUCGGGGGCCATC-dTdT (配列番号 22) | |
| IGF1R-12 | AAUGCUGACCUCUGUUACC-dTdT (配列番号 23) | |
| IGF1R-13 | CAUUACCGAGUACUUGCUGCU (配列番号 24) | |
| IGF1R-14 | CUUGCUGCUGUCCGAGUGGC (配列番号 25) | |
| IGF1R-15 | UCCGAGUGGCUGGCCUCGAGA (配列番号 26) | |
| IGF1R-16 | GGCCUCGAGAGCCUCGGAGAC (配列番号 27) | |
| luc | CGUACGCGAAUACUUCGA-dTdT (配列番号 28) | 40 |

【 0 2 1 4 】

発現プロファイルの初期分析は、同じ標的遺伝子に対する異なる siRNA に応答した、明らかな共通の遺伝子発現を示さなかった。その代わりに、MAPK14 を標的とする 8 つの siRNA 2 本鎖のそれぞれは、明確な発現パターンを与えた (図 1A)。同様に IGF1R に対する 16 個の 2 本鎖のそれぞれは、ユニークな発現パターンを与えた (図 1B)。3 つの独立した実験で実質的に同一の遺伝子発現パターンが観察され、特定の siRNA から生じる遺伝子制御が再現性あることを証明した。すなわち転写体発現パターンは siRNA 特異的であった。改変転写体の数と本体は、標的遺伝子をサイレンシングする siRNA の能力に対応しなかった。MAPK14 siRNA 2 本鎖のすべては標的を有効にサイレンシングし、RNA とタンパク質レベルの 90%

を超える低下を示した(図1A)。16個のIGF1R siRNAのうちの14個は、IGF1Rタンパク質とRNAレベルを60%より大きく低下させた(図1B)。IGF1R-4はIGF1Rの発現を80%低下させたが、発現が改変された遺伝子の数はIGF1R-5より少なく、標的の30%のみがサイレンシングされた。さらにルシフェラーゼを標的とするsiRNAは、ヒトゲノム中に同族の標的が無いにもかかわらずいくつかの遺伝子の発現を再現性よく制御した。すなわち遺伝子制御のパターンは、目的の標的ではなく、サイレンシングに使用したsiRNA配列に特異的である。

【0215】

siRNA誘導性サイレンシングによる遺伝子制御を理解するために、siRNA MAPK14-1によるMAPK14タンパク質とRNAノックダウンの濃度と動力学的詳細な分析を行った。siRNA濃度を1000倍低下させた時、標的遺伝子サイレンシングが検出できたが、オフターゲット遺伝子制御も検出できた(図2)。これらの遺伝子の多くは、siRNA濃度についてMAPK14(～1nM)とほとんど同じ半最大応答を示した。オフターゲット遺伝子制御は目的の標的のサイレンシングからは測定できず、これはオフターゲット遺伝子制御が単に高siRNA濃度のアーチファクトではないことを示している。

【0216】

MAPK14タンパク質とRNAノックダウンの詳細な動力学分析を行い、発現プロファイリングを使用して一時的な遺伝子発現パターンを分析した。MAPK14タンパク質は、siRNAトランスフェクション後に約40時間の半減期を示した(図3A)。これに対してMAPK14転写体は急速に分解され、半最大分解はトランスフェクション後11時間であった(図3B)。発現プロファイリングにより、MAPK14タンパク質の低下が観察されるよりかなり早い時点(6～12時間)での遺伝子制御が観察された。従って遺伝子発現の変化は、MAPK14機能の喪失により生じる2次的イベントである可能性は低かった。さらなる分析により、発現シグナチャがいくつかの一時的に明確な転写体群に分類できることが明らかになった(図3C)。第1群は、急速なサイレンシングを示す単一の転写体(標的MAPK14)を含有する。第2群は、半最大分解がマイクロアレイで測定すると7～13時間の、MAPK14へのサイレンシングの同様の動力学を示す9つの転写体を含有する。この同じ群の転写体は、独立した実験で急速な速度でダウンレギュレートされ、これらの遺伝子がこのsiRNAにより再現性よくサイレンシングされることを証明する。転写体制御の急速な速度は、これらが直接の遺伝子制御イベントであることを示唆する。これは第3群および第4群と反対であり、これらについては半最大分解は約40時間で起き、従って間接的な遺伝子発現の変化である可能性が高い。

【0217】

MAPK14との同時制御の基礎を理解するために、急速にダウンレギュレートされる転写体をさらに調べた。これらの転写体は、KPNB3、RAP2A、FLJ20291、RRAD、RPA2、DKFZp564J157、AF093680、および2つの未解析ESTコンティグを含む(図4A)。これらの遺伝子のいずれも、MAPK14経路で機能することはわかっていない。これらの転写体のすべては、siRNA2本鎖と部分的配列相同性の領域を含有することがわかった(図4A)。配列のアラインメントは、これらの遺伝子が2つの亜群に分類できることを示した。3つのオフターゲット遺伝子からなる1つの亜群は、siRNA配列の中央領域を含む相同性の14～15ヌクレオチドのコアを含有した。第2の亜群は、siRNA配列の3'末端に9つのヌクレオチドを含む相同性の小さいコアを含有した。これは、動力学群3～5と反対であり、これらは、siRNA配列中にランダムに分布した相同性のほんの短い区間(6～8ヌクレオチド)を示した。すなわちsiRNAの3'末端を含む配列相同性のコアの偏りは、急速にサイレンシングされる転写体にユニークである。siRNA配列中のヌクレオチド変化により排除される遺伝子サイレンシングの公開された報告(Elbashirら、2001、EMBO. J. 20:6877-6888; Holenら、2002 Nucleic Acids Res. 30:1757-1766)に基づくと、この限定された程度の配列相同性が転写体サイレンシングに充分であることは予測できなかったであろう。しかしこの可能性を検討するために、siRNA配列の各位置のヌクレオチドを系統的に置換し、発現シグナチャに対する改変配列の作用を調べた。代表的結果を図4Bに示す。第4位での単一ヌクレオチド置換はMAPK14のサイレンシングを排除し、この位置でMAPK14と相同性を有する亜群1の3つのオフタ

10

20

30

40

50

ターゲット遺伝子のサイレンシングを排除した。しかしこの領域にMAPK14との相同性を持たない亜群2の6つのオフターゲット遺伝子では、サイレンシングは排除されなかった。しかし第5位での単一ヌクレオチド置換は、MAPK14サイレンシングを排除はしないが低下させ、この領域でMAPK14と相同性を有する亜群1の3つのオフターゲット遺伝子のサイレンシングを排除した。

【0218】

この領域に相同性を共有しない6つのオフターゲット遺伝子の発現レベルは、このミスマッチにより影響を受けなかった。これらの結果は、オフターゲット遺伝子のサイレンシングがMAPK14発現の喪失に依存しないことを確認している。第15位での単一ヌクレオチド置換もまたMAPK14サイレンシングを低下させ、すべての9つのオフターゲット遺伝子のサイレンシングを排除した。第15位でのミスマッチの影響はより劇的であり、これはおそらくすべての9つの転写体がこの領域でMAPK14との相同性を有するためであろう。RPA2はすでにこの位置にCではなくTを含有し、これはMAPK14オリゴのアンチセンス鎖とG:U塩基対を生成するであろう。すなわちこの置換は、塩基対合とサイレンシングについて許容されるものである。改変されたsiRNA配列U:U塩基対合を生成し、これはあまり許容されず、従ってこの転写体のサイレンシングを排除するであろう。すなわち、一部のヌクレオチド置換はサイレンシングにとってより有害である。まとめるとこれらの結果は、これらの転写体のMAPK14との配列相同性は、MAPK14 siRNAによるそのサイレンシングの原因であることを確認する。

【0219】

観察されたオフターゲット遺伝子サイレンシングが配列相同性に基づき、MAPK14発現の低下の結果ではないことのさらなる証拠として、MAPK14を標的とするように設計したが異なるオリゴ配列を有する異なるsiRNA2本鎖を試験した。この2番目のsiRNA (MAPK14-2) はMAPK14のサイレンシングを保持したが、新しいsiRNA配列との相同性が無い第2群の転写体のいずれもサイレンシングしなかった (図4B)。最後に2つのオフターゲット転写体 (KPNB3とFLJ20291) について図4Aに例示した配列に対応するsiRNAは、その目的の標的以外にMAPK14の発現をサイレンシングした (図4C)。KPNB3 siRNAは、MAPK14と14個の連続ヌクレオチドを共有し、全部で15ヌクレオチドの相同性を共有する。FLJ20291 siRNAはMAPK14と、11個のみの連続ヌクレオチドを共有し、全部で15ヌクレオチドの相同性を共有する。すなわち、15ヌクレオチド、そしておそらくは11と少ない連続ヌクレオチドの配列相同性が、非標的転写体のサイレンシングを指令するのに充分であろう。従ってRNA干渉は所望の標的の確実なサイレンシングを引き起こすが、siRNA配列と部分的相同性を有するmRNA転写体の分解によりオフターゲット遺伝子制御が起きることがある。

【0220】

MAPK14に対するすべての8つのsiRNA2本鎖とIGF1Rに対する14個のsiRNA2本鎖が、有効に標的遺伝子をサイレンシングした。従って、標的遺伝子機能の喪失を示す各標的に対する異なるsiRNA中で共通の発現シグナチャを同定することが可能であろう。データに厳しい要件をつけて、複数のsiRNAによる標的遺伝子サイレンシングにより生じる発現パターンのみを含むようにすると、共通の発現パターンが観察された。両方の実験のデータを併せて1つの実験群にし、少なくとも7つのsiRNAによるサイレンシングに反応して発現が少なくとも2倍変化した遺伝子に限定したデータ解析を行った。この解析により、MAPK14 siRNAに反応して8つの遺伝子について発現低下の共通のパターンが得られた (図5)。これらの中に標的MAPK14があり、これでは8つすべてのsiRNAに反応して発現が10倍抑制された。MAPK14サイレンシングの共通の発現シグナチャはまた、通常インターフェロンに反応して誘導される多くの遺伝子を含んだ。この解析は、IGF1R siRNAに反応してその発現が低下した16個の遺伝子を証明した (図5)。IGF1R自体はこの細胞株中で十分に低レベルで発現されて統計的採用基準より低かったため、このクラスター解析により発現低下するとして同定されなかった。リアルタイムPCRによるIGF1R mRNAの定量は、IGF1Rの発現がsiRNA2本鎖のそれぞれに反応して低下していることを証明した (図1B)。IGF1Rサイレンシングの共通の発現シグナチャは、EGR1 (増殖因子シグナル伝達に反応して誘導される早期増殖応

10

20

30

40

50

答遺伝子)、いくつかの転写レギュレーター(HMCS、CAMTA2、CHOP、DDIT3)、および栄養物質検知とグルコース制御に関与する遺伝子(PCK2、ASNS)を含んだ。いくつかの転写体は、siRNAにより両方の標的遺伝子にアップレギュレートされることがわかった。これらは主に、非解析ESTコンテイングであり、およびおそらく長いdsRNAに反応したインターフェロン遺伝子の誘導に似た方法で、siRNAに対して反応して刺激された遺伝子を反映するであろう。1つの興味深い例外は、複数のsiRNAによるダイサーのIGF1Rへの誘導であるが、この遺伝子制御の意義は不明である。IGF1Rサイレンシングに反応してダウンレギュレートされる遺伝子は、MAPK14サイレンシングに反応して制御される遺伝子とは明確に異なる(図5)。すなわち、観察された発現プロファイルはオフターゲット遺伝子制御を含むが、オフターゲット遺伝子制御を示す発現パターンは区別できるであろう。

10

【0221】

IGF1Rに対してsiRNAにより誘発されるオフターゲット遺伝子制御をさらに解明するために、発現プロファイリングを使用して、IGF(インスリン様増殖因子)刺激に反応するMCF7乳癌細胞中のIGF1R遺伝子機能を解析した。siRNAトランスフェクション後に、細胞を血清枯渇させ、次にIGFで処理してIGF1R経路を刺激した。IGF1R siRNAは標的遺伝子発現を有効にサイレンシングし、下流のタンパク質Aktのリン酸化の低下により測定される受容体機能を妨害した(図6B)。siRNAの非存在下では、Aktリン酸化は非常に低濃度のIGFでも明らかであった。これに対してAktリン酸化は、飽和レベルのIGFでもIGF1Rサイレンシングに反応して大きく低下した。モックトランスフェクトした細胞のIGF刺激は、IGFシグナル伝達に反応してその発現が上昇した転写体を同定した(図6C)。この群の遺伝子は、同じ遺伝子がいくつかの他の細胞株で誘導されたため、IGFシグナル伝達経路を介して遺伝子制御を示す可能性がある。siRNAによるIGF1RサイレンシングのバックグラウンドでIGF刺激を行うと、これらのIGFシグナチャ遺伝子の発現は、変化しなかったかまたは低下した。すなわち標的遺伝子機能の干渉を反映する発現シグナチャは、発現プロファイリングにより同定することができる。まとめるとデータは、RNAi誘導遺伝子サイレンシングにより生じる遺伝子発現プロファイルは、オン標的遺伝子制御とオフターゲット遺伝子制御の合わさったものである。

20

【0222】

RNA干渉は、培養哺乳動物細胞中で遺伝子発現を操作するための強力手段である。しかしRNAiによる遺伝子サイレンシングから生じる表現型の解釈には、オフターゲット遺伝子制御の影響の評価を含めなければならない。この試験で同定されるオフターゲット遺伝子産物は、目的の標的の生物学的経路内で機能することが知られており、従って標的化シグナル伝達経路に関連する遺伝子のタンパク質またはRNA分析により、検出されていないであろう。これらのオフターゲットイベントが同定されたのは、偏りの無いマイクロアレイ分析によってのみである。オフターゲット遺伝子の詳細な配列解析は、siRNA2本鎖との相同性の15個のヌクレオチドのコアおよび11個の少ない連続ヌクレオチドでも、目的の標的以外に転写体の分解を指令することができることを証明した。siRNAの3'末端との相同性の9つのヌクレオチドの小さいコアは、転写体サイレンシングに充分であろう。

30

【0223】

siRNAとmiRNAを含む小制御性RNAの生物学的機能はあまりわかっていない。1つのよくある質問は、これらの2つのクラスの制御性RNAの明確なサイレンシング経路が決定される機構に関するものである。miRNAはゲノムから発現される制御性RNAであり、前駆体ステム-ループ構造から処理されて、標的mRNAの3'UTR中の配列に結合する1本鎖核酸を産生する(Leeら、1993, Cell 75:843-854; Reinhartら、2000, Nature 403:901-906; Leeら、2001, Science 294:862-864; Lauら、2001, Science 294:858-862; Hutvagnerら、2001, Science 293:834-838)。miRNAはほんのわずかの相補性で転写体配列に結合し(Zengら、2002, Molec. Cell 9:1327-1333)、定常状態RNAレベルに影響を与えることなく翻訳を抑制する(Leeら、1993, Cell 75:843-854; Wightmanら、1993, Cell 75:855-862)。miRNAとsiRNAの両方とも、ダイサー(Dicer)により処理され、RNA誘導性サイレンシング複合体

40

50

の成分に結合する (Hutvagnerら、2001, *Science* 293:834-838; Grishokら、2001, *Cell* 106:23-34; Kettingら、2001, *Genes Dev.* 15:2654-2659; Williamsら、2002, *Proc. Natl. Acad. Sci. USA* 99:6889-6894; Hammondら、2001, *Science* 293:1146-1150; Murlato sら、*Genes Dev.* 16:720-728)。最近の報告 (Hutvagnerら、2002, *Scienceexpress* 297:2056-2060) は、miRNA経路対siRNA経路を介する遺伝子制御が、標的転写体に対する相補性の程度によってのみ決定されるという仮説を報告している。著者らは、miRNA標的に対してほんのわずかの同一性を有するsiRNAは、RNA分解を開始させるのではなく、miRNAと同様に翻訳抑制で機能すると推測している。これに対して本例では、mRNA転写体に対してほんの部分的な相補性を有する合成siRNA2本鎖が、培養ヒト細胞中でこの転写体の分解を引き起こすことを明確に証明した。これらの結果は、標的転写体との相同性の程度が、siRNAの機能をmRNAの機能から区別する唯一の決定要因ではないことを示す。 10

【0224】

オフターゲット遺伝子制御に関係する相同性の程度は小さいため、目的の標的に絶対的に特異的なsiRNA配列を選択することは困難であろう。さらに、siRNAとの相同性を有する遺伝子のゲノムを探索してオフターゲットイベントを予測することは困難であろう。配列相同性のみでは、転写体配列がsiRNAにアクセスできることを保証できない。さらに、必ずしもすべてのオフターゲット転写体候補が、目的の細胞株中で発現されるとは限らない。すなわち配列相同性を有するすべての転写体がサイレンシングされるわけではない。さらに、サイレンシングされる転写体のすべてが配列相同性を有するわけではない。直接的な相同性ベースの遺伝子サイレンシングは早い時点で同定できるが、遅い時点で観察される遺伝子発現パターンは、標的遺伝子ならびにオフターゲット遺伝子のサイレンシングからの下流の遺伝子制御を含むであろう。さらなる複雑さは、siRNA2本鎖がクロマチンコンフォメーションの改変を介して遺伝子発現を制御する能力を有することから生じる。動原体反復配列に対するsiRNA2本鎖が最近、*S. pombe* (エス・ポンベ) 中のヘテロクロマチンのサイレンシングに影響を与えることが報告された (Volpeら、*Science* 297:1833-1837; Reinhartら、2000, *Science* 297:1831)。クロマチンコンフォメーションを維持するのに関与する転写体とのsiRNA配列の部分的相同性は、標的遺伝子の近傍に位置する非標的遺伝子の発現を変化させることが考えられる。すなわちRNA干渉は、遺伝子発現の複雑な変化を生じさせ、得られる表現型の解釈は慎重にしなければならない。 20

【0225】

siRNA特異的オフターゲット作用にもかかわらず、同じ標的に対する複数のsiRNA2本鎖を分析すると、標的タンパク質の喪失を示す可能性のある共通の発現シグナチャが同定できるであろう。目的の標的遺伝子をサイレンシングするために複数のsiRNA2本鎖を取り込むことは、観察された表現型と発現パターンが標的遺伝子サイレンシングに関連していることの自信を上昇させるであろう。こうして、RNAiによる遺伝子サイレンシングと組合せた発現プロファイリングは、培養哺乳動物細胞中の遺伝子機能を同定し解析するための強力な手段となるであろう。 30

【0226】

実施例2

6.2. 実施例2: siRNAプールはサイレンシング特異性を上昇させる

本例は、siRNAプールを使用することにより、オフターゲットイベントの数は増加 (付加) するが、その強度制御は低下するらしいことを証明する。オフターゲット遺伝子およびオンターゲット遺伝子のサイレンシングは同様の用量応答で変化するため、単一siRNAの濃度の低下は、特異性を改良しない。これに対してsiRNAのプールを増加させることは、オンターゲットサイレンシングを維持し、オフターゲット遺伝子サイレンシングの数と強度を低下させる。これは、限定量のリスク (RISC) との結合についてのsiRNA間の競合が原因かも知れない。これらの例では、すべてのsiRNAが標的サイレンシングにおいて有効であることがすでに測定されている。これらのsiRNAは同じ標的遺伝子転写体を標的とし、従ってオンターゲットサイレンシングに寄与するがそれぞれは明確なオフターゲット活性を有するため、プール中のsiRNAの数を増加させることによりオフターゲット活性が 40 50

希釈される。その結果、オンターゲット：オフターゲット遺伝子サイレンシングの比が上昇し、特異性が上昇する。これは、プールサイズをさらに大きくすることが特異性のさらなる上昇につながることを示唆する。これは、遺伝子サイレンシングの特異性を上昇させ、これは標的を証明する試みにおいて大いに有益であろう。

【0227】

図10は、3つのsiRNAのsiRNAプールを使用して得られた結果を示す。個々のsiRNAを、33nM濃度でHeLa細胞中にトランスフェクトした。3つのsiRNAのプールを100nM(それぞれ33nM)でトランスフェクトした。トランスフェクションの24時間後RNAを抽出し、モックトランスフェクトした細胞からのRNAに対してプロファイリングした。右のパネルは、オンターゲット遺伝子サイレンシングがプール中で維持されることを示す。オフターゲットイベントの数は、個々のsiRNAからのオフターゲットイベントのほぼ合計であるが、オフターゲットサイレンシングの強度は低下した。

10

【0228】

図11は、siRNAプールがサイレンシング特異性を上昇させたことを示す。単一のsiRNA濃度の低下は、特異性を改善しない(上のパネル)。これに対して、プールサイズを上昇させることはオンターゲットサイレンシングを維持したが、オフターゲット遺伝子サイレンシングの数と強度を低下させた。これは、リスク(RISC)との結合についてのsiRNA間の競合のためかも知れない。その結果、オンターゲット：オフターゲット遺伝子サイレンシングの比は上昇し、特異性が上昇した。これは、さらに大きなプールサイズまで上昇させると特異性がさらに上昇することを示唆し、これは標的評価の試みで非常に有用であろう。

20

【0229】

図12は、9つのsiRNAのsiRNAプールを使用して得られた結果を示す。個々のsiRNAの濃度は11nMであり、プールはそれぞれ11nMの9つのメンバーを含有する。siRNAプール中のsiRNAの数が増加すると、おそらくリスク(RISC)との結合についてのsiRNA間の競合のために、プール中のシグナチャ遺伝子の数が低下した。その結果、オフターゲットサイレンシングの強度が低下し、オンターゲット：オフターゲット遺伝子サイレンシングの比率が上昇し、特異性が上昇した。

【0230】

実施例3

30

6.3. 実施例3：RNA干渉における鎖優先の評価

実施例1において、siRNAがオンターゲットならびにオフターゲット活性を有することが証明された。いくつかのオフターゲット遺伝子は、目的の標的遺伝子と同じ速度でダウンレギュレートされることが証明され、これらの遺伝子がsiRNAにより直接制御されることを意味する。これらの協調的にダウンレギュレートされたオフターゲットシグナチャ遺伝子は、2つの方法でこれらを制御するsiRNAとアラインメントすることがわかった：

1)同一性の中央連続区間：オフターゲット遺伝子とすべて同一のsiRNA2本鎖の中央部分の11個以上の塩基。

【0231】

2)同一性の3'ベースの連続区間：オフターゲット遺伝子とすべて同一のsiRNA2本鎖の3'末端の3塩基以内で終止する8個以上の塩基。

40

【0232】

アラインメントは、そのダウンレギュレーション速度を調べたsiRNAのセンス鎖またはアンチセンス鎖のいずれかと同一であることがわかった。そのアラインメントがセンス鎖と同一のsiRNAについては、アンチセンス鎖がオンターゲットおよびオフターゲット制御を引き起こし；そのアラインメントがアンチセンス鎖と同一のsiRNAについては、アンチセンス鎖の活性以外にセンス鎖がオフターゲット制御を引き起こしていると推測される。

【0233】

この例は、1つの時点のシグナチャのみが利用できる時に、siRNA2本鎖のどちらがオフターゲット遺伝子をサイレンシングしているかを評価するための2つの方法を報告する。

50

【0234】

第1の方法では、1つの時点のシグナチャが、1つのsiRNA鎖対他のsiRNA鎖とのより大きなアラインメントの程度について、すべてのシグナチャ遺伝子を投票することにより評価された。これはまた、目的の標的遺伝子と同じ速度でダウンレギュレートされる遺伝子について鎖優先を評価するために使用することもできる。

【0235】

この方法では、シグナチャ遺伝子配列を、siRNAの両方の配列とアラインメントさせた。各鎖とのアラインメントを、最も長い連続的同一区間について調べた。これらの連続的同一区間の長さを比較した。遺伝子はその鎖に対して最も長い連続的同一区間を有した鎖に、遺伝子が投票したと見なした。投票は、2つの同一区間の長い方の過剰の塩基数により重み付けをした。この理由は、例えば1つの鎖が15塩基の同一区間を有し、他の鎖が5塩基の同一区間を有するケースは、1つの鎖が8塩基の同一区間を有し、他の鎖が7塩基の同一区間を有するケースより、より大きく重み付けをすべきであるということである。

10

【0236】

従って票は、ある鎖中で過剰の連続的同一区間中の塩基である。総票は、優勢であったすべての遺伝子中の勝った鎖の連続的同一性中の総過剰長さである。

【0237】

この方法のバックグラウンドは、siRNA制御シグナチャを測定した複数の遺伝子、例えばマイクロアレイを使用して測定したすべての遺伝子に対するsiRNAの両方の鎖のアラインメントを調べ、鎖を上記したように投票させ、siRNAの各鎖について遺伝子当たりの平均過剰重みを計算することにより、測定された。

20

【0238】

第2の方法では、単一時点シグナチャを、目的の標的遺伝子と同じ速度でダウンレギュレートした遺伝子とのsiRNAのアラインメントから開発された3'偏りモデルとシグナチャ遺伝子アラインメントとを比較することにより評価した。

【0239】

この方法では、シグナチャ遺伝子を分析して、オフターゲット遺伝子アラインメントの3'偏りモデルに一致する割合を測定した(アラインメント2)上記参照)。シグナチャ遺伝子配列をsiRNAの両方の鎖とアラインメントさせた。各鎖とのアラインメントを、連続的同一区間について調べた。これらは：

30

a)少なくとも7塩基の連続的同一区間を有した；および

b)特定の鎖の3'末端の3塩基以内で終止した、

場合にアラインメントを印付けをした。

【0240】

センス鎖との同一性およびアンチセンス鎖との同一性を有する上記基準を満足するアラインメントの印(tally)は、以下のように比較される

$SLR = \log(\text{センス同一印} / \text{アンチセンス同一印})$

SLRの有意性は、その上でsiRNA制御シグナチャが測定されたチップ上のすべての配列についてセンス同一印とアンチセンス同一印を有するシグナチャについて、センス同一印とアンチセンス同一印を比較することにより評価される。

40

【0241】

有意性は超幾何分布から計算され、ここで

x、サンプル - 選択 = 1つの鎖のシグナチャ印

n、サンプル - 合計 = 両方の鎖のシグナチャ印

M、親 - 選択 = 同じ鎖のチップ印

N、親 - 合計 = 両方の鎖のチップ印

そして、超幾何分布の式は以下の通りである：

$$h(x, n, M, N) = \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}}$$

【0242】

超幾何分布は対称ではないため、センス鎖とアンチセンス鎖選択超幾何分布の両方のテイルが使用された。センス同一過剰に対応する2つのテイルを平均して過剰センス同一性のp値が得られ、アンチセンス同一過剰に対応する両方のテイルを平均して過剰アンチセンス同一性のp値が得られた。 10

【0243】

これらの両方の鎖の鎖優先は、標的遺伝子と同じ速度でダウンレギュレートされた遺伝子の鎖優先分析の結果と一致することがわかった。

【0244】

特に最初の方法は、12時間、24時間、および48時間シグナチャの分析に有効であることがわかった。24時間シグナチャが一般に最も有意な結果を与えた。

【0245】

実際、siRNAの鎖優先性の測定についてのシグナチャ遺伝子のかなりの割合またはシグナチャ全体の有用性は、細胞中の遺伝子発現に対するsiRNAの直接作用がシグナチャ遺伝子のかなりの割合により示されることを意味する。例えば両方法は、12時間IGF1R-73シグナチャの約35~40%の遺伝子(すなわち約12遺伝子)が鎖の偏りの推定に関連すると推定した。一方24時間IGF1R-73シグナチャの約35~40%の(すなわち約80遺伝子)もまた関連した。 20

【0246】

表IIは、377個のsiRNAのライブラリーを示す。表II中のBioID数はまた、siRNAを特定するためにこの開示でも使用される。例えばsiRNA MAPK14-193は表II中のBioID No. 193を有するsiRNAに対応する。

【表 2】

表 II 377 個の siRNA のライブラリー

| BioID | アクセッション 番号 | スタート位置 | 19mer 配列 | %サイレンシング | Set 1 | Set 2 | 配列番号 |
|-------|---------------|--------|---------------------|----------|----------|----------|------|
| 31 | NM_000075 | 437 | TGTTGTCCGGCTGATGGAC | 27.0 | Training | Training | 39 |
| 36 | NM_001813 | 1036 | ACTCTTACTGCTCTCCAGT | 86.1 | Test | Training | 40 |
| 37 | NM_001813 | 1278 | CTTAACACGGATGCTGGTG | 60.1 | Test | Training | 41 |
| 38 | NM_001813 | 3427 | GGAGAGCTTTCTAGGACCT | 88.0 | Test | Training | 42 |
| 39 | NM_004073 | 192 | AGTCATCCCGCAGAGCCGC | 55.0 | Training | Training | 43 |
| 40 | NM_004073 | 1745 | ATCGTAGTGCTTGTACTTA | 70.0 | Training | Training | 44 |
| 41 | NM_004073 | 717 | GGAGACGTACCGCTGCATC | 65.0 | Training | Training | 45 |
| 42 | AK092024 | 437 | GCAGTGATTGCTCAGCAGC | 93.0 | Training | Training | 46 |
| 43 | NM_030932 | 935 | GAGTTTACCGACCACCAAG | 81.0 | Training | Training | 47 |
| 44 | NM_030932 | 1186 | TGCGGATGCCATTCACTGG | 35.0 | Training | Training | 48 |
| 45 | NM_030932 | 1620 | CACGGTTGGCAGAGTCTAT | 73.0 | Training | Training | 49 |
| 49 | U53530 | 169 | GCAAGTTGAGCTCTACCGC | 59.0 | Training | Training | 50 |
| 50 | U53530 | 190 | TGGCCAGCGCTTACTGGAA | 75.0 | Training | Training | 51 |
| 64 | NM_006101 | 1623 | GTTCAAAGCTGGATGATC | 79.0 | Tcst | Training | 52 |
| 65 | NM_006101 | 186 | GGCCTCTATACCCCTCAA | 74.4 | Test | Training | 53 |
| 66 | NM_006101 | 968 | AGAACCGAATCGTCTAGAG | 80.3 | Test | Training | 54 |
| 67 | NM_000859 | 253 | CACGATGCATAGCCATCCT | 25.0 | Training | Training | 55 |
| 68 | NM_000859 | 1075 | CAGAGACAGAATCTACT | 45.0 | Training | Training | 56 |
| 69 | NM_000859 | 1720 | CAACAGAAGTTGTCTTGT | 50.0 | Training | Training | 57 |
| 70 | NM_000859 | 2572 | TTGTGTGTGGACCGTAAT | 80.0 | Training | Training | 58 |
| 71 | NM_000875 | 276 | GCTCACGGTCATTACCGAG | 63.9 | Training | Training | 59 |
| 72 | NM_000875 | 441 | CCTGAGGAACATTACTCGG | 0.0 | Training | Training | 60 |
| 73 | NM_000875 | 483 | TGCTGACCTCTGTTACCTC | 50.0 | Training | Training | 61 |
| 74 | NM_000875 | 777 | CGACACGGCCTGTGTAGCT | 58.0 | Training | Training | 62 |
| 75 | NM_000875 | 987 | CGGCAGCCAGAGCATGTAC | 63.0 | Training | Training | 63 |
| 76 | NM_000875 | 1320 | CCAGAACTTGCAGCAACTG | 70.0 | Training | Training | 64 |
| 81 | NM_000875 | 351 | CCTCACGGTCATCCGCGGC | 0.0 | Training | Training | 65 |
| 83 | NM_000875 | 387 | CTAGCCCTGGTCATCTTC | 32.0 | Training | Training | 66 |
| 84 | NM_000875 | 417 | TCTCAAGGATATTGGGCTT | 54.0 | Training | Training | 67 |
| 85 | NM_000875 | 423 | GGATATTGGGCTTTACAAC | 71.0 | Training | Training | 68 |
| 86 | NM_000875 | 450 | CATTACTCGGGGGCCATC | 53.0 | Training | Training | 69 |
| 87 | NM_000875 | 481 | AATGCTGACCTCTGTTACC | 54.6 | Training | Training | 70 |

10

20

30

| | | | | | | | |
|-----|-----------|------|----------------------|------|----------|----------|-----|
| 117 | NM_004523 | 1689 | CTGGATCGTAAGAAGGCAG | 74.7 | Training | Test | 71 |
| 118 | NM_004523 | 484 | TGGAAGGTGAAAGGTCACC | 16.0 | Training | Test | 72 |
| 119 | NM_004523 | 802 | GGACAACTGCAGCTACTCT | 84.1 | Training | Test | 73 |
| 139 | NM_002358 | 219 | TACGGACTCACCTTGCTTG | 83.0 | Training | Training | 74 |
| 144 | NM_001315 | 779 | GTATATACATTTCAGCTGAC | 78.5 | Training | | 75 |
| 145 | NM_001315 | 1080 | GGAACACCCCCGCTTATC | 27.2 | Training | | 76 |
| 146 | NM_001315 | 1317 | GTGGCCGATCCTTATGATC | 81.3 | Training | | 77 |
| 152 | NM_001315 | 607 | ATGTGATTGGTCTGTTGGA | 95.0 | Training | | 78 |
| 153 | NM_001315 | 1395 | GTCATCAGCTTTGTGCCAC | 92.0 | Training | | 79 |
| 154 | NM_001315 | 799 | TAATTCACAGGGACCTAAA | 82.0 | Training | | 80 |
| 155 | NM_001315 | 1277 | TGCCTACTTTGCTCAGTAC | 95.0 | Training | | 81 |
| 193 | NM_001315 | 565 | CCTACAGAGAACTGCGGTT | 90.0 | Training | | 82 |
| 190 | NM_001315 | 763 | TTCTCCGAGGTCTAAAGTA | 87.0 | Training | | 83 |
| 192 | NM_001315 | 1314 | CCAGTGGCCGATCCTTATG | 89.0 | Training | | 84 |
| 194 | NM_001315 | 1491 | GGCCTTTTCACGGGAACTC | 97.0 | Training | | 85 |
| 201 | NM_016195 | 2044 | CTGAAGAAGCTACTGCTTG | 80.3 | Test | Training | 86 |
| 202 | NM_016195 | 4053 | GACATGCGAATGACTAG | 75.9 | Test | Training | 87 |
| 203 | NM_016195 | 3710 | AGAGGAACTCTCTGCAAGC | 84.7 | Test | Training | 88 |
| 204 | NM_014875 | 4478 | AAACTGGGAGGCTACTTAC | 93.0 | Test | Training | 89 |
| 205 | NM_014875 | 1297 | ACTGACAACAAAGTGCAGC | 37.0 | Test | Training | 90 |
| 206 | NM_014875 | 5130 | CTCACATTGTCCACCAGGA | 91.6 | Test | Training | 91 |
| 210 | NM_004523 | 4394 | GACCTGTGCCTTTTAGAGA | 63.7 | Training | Test | 92 |
| 211 | NM_004523 | 2117 | GACTTCATTGACAGTGCC | 71.0 | Training | Test | 93 |
| 212 | NM_004523 | 799 | AAAGGACAACCTGCAGCTAC | 49.0 | Training | Test | 94 |
| 213 | NM_000314 | 2753 | TGGAGGGGAATGCTCAGAA | 40.0 | Training | Training | 95 |
| 214 | NM_000314 | 2510 | TAAAGATGGCACTTTCCCG | 79.0 | Training | Training | 96 |
| 215 | NM_000314 | 2935 | AAGGCAGCTAAAGGAAGTG | 55.0 | Training | Training | 97 |
| 234 | NM_007054 | 963 | TATTGGGCCAGCAGATTAC | 76.9 | Training | Training | 98 |
| 235 | NM_007054 | 593 | TTATGACGCTAGGCCACAA | 74.4 | Training | Training | 99 |
| 236 | NM_007054 | 1926 | GGAGAAAGATCCCTTTGAG | 78.3 | Training | Training | 100 |
| 237 | NM_006845 | 324 | ACAAAAACGGAGATCCGTC | 72.2 | Training | Training | 101 |
| 238 | NM_006845 | 2206 | ATAAGCAGCAAGAAACGGC | 30.9 | Training | Training | 102 |
| 239 | NM_006845 | 766 | GAATTTGCGGCTACTTTGG | 65.8 | Training | Training | 103 |
| 240 | NM_005163 | 454 | CGCACCTTCCATGTGGAGA | 86.8 | Training | Training | 104 |
| 241 | NM_005163 | 1777 | AGACGTTTTTGTGCTGTGG | 76.0 | Training | Training | 105 |
| 242 | NM_005163 | 1026 | GCTGGAGAACCTCATGCTG | 87.8 | Training | Training | 106 |
| 243 | NM_005733 | 2139 | CTCTACCACTGAAGAGTTG | 90.7 | Training | Training | 107 |
| 244 | NM_005733 | 1106 | AAGTGGGTCGTAAGAACCA | 82.5 | Training | Training | 108 |
| 245 | NM_005733 | 696 | GAAGCTGTCCCTGCTAAAT | 93.4 | Training | Training | 109 |
| 246 | NM_001813 | 3928 | GAAGAGATCCCAGTGCTTC | 86.8 | Test | Training | 110 |
| 247 | NM_001813 | 4456 | TCTGAAAGTGACCAGCTCA | 82.5 | Test | Training | 111 |
| 248 | NM_001813 | 2293 | GAAAATGAAGCTTTGCGGG | 78.4 | Test | Training | 112 |
| 249 | NM_005030 | 1135 | AAGAAGAACCAGTGGTTCG | 83.0 | Test | Test | 113 |
| 250 | NM_005030 | 572 | CCGAGTTATTCATCGAGAC | 93.6 | Test | Test | 114 |
| 251 | NM_005030 | 832 | AAGAGACCTACCTCCGGAT | 85.0 | Test | Test | 115 |
| 255 | NM_001315 | 3050 | AATATCCTCAGGGGTGGAG | 36.0 | Training | | 116 |
| 256 | NM_001315 | 1526 | GTGCCTCTTGTGTCAGAGA | 88.0 | Training | | 117 |
| 257 | NM_001315 | 521 | GAAGCTCTCCAGACCATTT | 96.0 | Training | | 118 |
| 261 | NM_006218 | 456 | AGAAGCTGTGGATCTTAGG | 65.3 | Test | Training | 119 |
| 262 | NM_006218 | 3144 | TGATGCACATCATGGTGGC | 68.9 | Test | Training | 120 |
| 263 | NM_006218 | 2293 | CTAGGAAACCTCAGGCTTA | 94.7 | Test | Training | 121 |
| 264 | NM_000075 | 1073 | GCGAATCTCTGCCTTTCGA | 79.0 | Training | Training | 122 |
| 265 | NM_000075 | 685 | CAGTCAAGCTGGCTGACTT | 78.0 | Training | Training | 123 |
| 266 | NM_000075 | 581 | GGATCTGATGCGCCAGTTT | 77.0 | Training | Training | 124 |

| | | | | | | | | |
|-----|-----------|------|----------------------|------|----------|----------|-----|----|
| 288 | NM_020242 | 1829 | GCACAACCTCCTGCAAATTC | 87.4 | Training | Training | 125 | |
| 289 | NM_020242 | 3566 | GATGGAAGAGCCTCTAAGA | 82.7 | Training | Training | 126 | |
| 290 | NM_020242 | 2631 | ACGAAAAGCTGCTTGAGAG | 73.4 | Training | Training | 127 | |
| 291 | NM_004073 | 570 | GAAGACCATCTGTGGCACC | 65.0 | Training | Training | 128 | |
| 292 | NM_004073 | 1977 | TCAGGGACCAGCTTTACTG | 60.0 | Training | Training | 129 | |
| 293 | NM_004073 | 958 | GTTACCAAGAGCCTCTTTG | 75.0 | Training | Training | 130 | |
| 294 | NM_005026 | 3279 | AACCAAAGTGAAGTGGCTG | 56.3 | Training | Training | 131 | |
| 295 | NM_005026 | 2121 | GATCGGCCACTTCCTTTTC | 70.9 | Training | Training | 132 | |
| 296 | NM_005026 | 4004 | AGAGATCTGGGCCTCATGT | 67.3 | Training | Training | 133 | |
| 303 | NM_000051 | 5373 | AGTTCGATCAGCAGCTGTT | 60.9 | Training | Training | 134 | |
| 304 | NM_000051 | 3471 | TAGATTGTTCCAGGACACG | 71.2 | Training | Training | 135 | 10 |
| 305 | NM_000051 | 7140 | GAAGTTGGATGCCAGCTGT | 56.3 | Training | Training | 136 | |
| 309 | NM_004064 | 1755 | TGGTGATCACTCCAGGTAG | 25.3 | Training | Training | 137 | |
| 310 | NM_004064 | 1505 | TGTCCCTTTCAGAGACAGC | 5.0 | Training | Training | 138 | |
| 311 | NM_004064 | 1049 | GACGTCAAACGTAACAGC | 50.2 | Training | Training | 139 | |
| 312 | NM_006219 | 1049 | AAGTTCATGTCCAGGGCTGG | 76.6 | Test | Training | 140 | |
| 313 | NM_006219 | 2631 | CAAAGATGCCCTTCTGAAC | 88.9 | Test | Training | 141 | |
| 314 | NM_006219 | 453 | AATGCGCAAATTCAGCGAG | 32.9 | Test | Training | 142 | |
| 339 | NM_003600 | 437 | GCACAAAAGCTTGTCTCCA | 96.0 | Test | Training | 143 | |
| 340 | NM_003600 | 1071 | TTGCAGATTTTGGGTGGTC | 37.0 | Test | Training | 144 | |
| 341 | NM_003600 | 1459 | ACAGTCTTAGGAATCGTGC | 61.1 | Test | Training | 145 | |
| 342 | NM_004958 | 1476 | AGGACTTCGCCATAAGAG | 61.8 | Test | Training | 146 | |
| 343 | NM_004958 | 5773 | CAACCTCCAGGATACACTC | 80.9 | Test | Training | 147 | 20 |
| 344 | NM_004958 | 7886 | CCAACCTTCTAGCTGCTGT | 71.1 | Test | Training | 148 | |
| 348 | NM_004856 | 1999 | GAATGTGAGCGTAGAGTGG | 92.2 | Training | Training | 149 | |
| 349 | NM_004856 | 1516 | CCATTGGTTACTGACGTGG | 87.7 | Training | Training | 150 | |
| 350 | NM_004856 | 845 | AACCCAAACCTCCACAATC | 71.8 | Training | Training | 151 | |
| 369 | XM_294563 | 117 | GAAAGAAGCAGTTGACCTC | 59.9 | Training | Training | 152 | |
| 370 | XM_294563 | 2006 | CTAAAAGCTGGGTGGACTC | 69.4 | Training | Training | 153 | |
| 371 | XM_294563 | 389 | GAAAGCACCTCTTTGTGTG | 64.2 | Training | Training | 154 | |
| 399 | NM_000546 | 1286 | TGAGGCCTTGGAACTCAAG | 17.8 | | | 155 | |
| 400 | NM_000546 | 2066 | CCTCTTGGTCGACCTTAGT | 74.5 | | | 156 | |
| 401 | NM_000546 | 1546 | GCACCCAGGACTTCCATTT | 93.2 | | | 157 | |
| 417 | NM_001184 | 3790 | GAAACTGCAGCTATCTTCC | 75.8 | Training | Training | 158 | |
| 418 | NM_001184 | 7717 | GTTACAATGAGGCTGATGC | 73.0 | Training | Training | 159 | 30 |
| 419 | NM_001184 | 5953 | TCACGACTCGCTGAACTGT | 78.8 | Training | Training | 160 | |
| 453 | NM_005978 | 323 | GACCGACCCTGAAGCAGAA | 91.3 | Test | Test | 161 | |
| 454 | NM_005978 | 254 | TTCCAGGAGTATGCTGTTT | 74.4 | Test | Test | 162 | |
| 455 | NM_005978 | 145 | GGAACCTTCTGCACAAGGAG | 96.5 | Test | Test | 163 | |
| 465 | NM_000551 | 495 | TGTTGACGGACAGCCTATT | 75.5 | Test | Training | 164 | |
| 466 | NM_000551 | 1056 | GGCATTGGCATCTGCTTTT | 89.7 | Test | Training | 165 | |
| 467 | NM_000551 | 3147 | GTGAATGAGACACTCCAGT | 82.2 | Test | Training | 166 | |
| 468 | NM_002658 | 1944 | GAGCTGGTGTCTGATTGTT | 82.8 | Test | Training | 167 | |
| 469 | NM_002658 | 1765 | GTGTAAGCAGCTGAGGTCT | 44.4 | Test | Training | 168 | |
| 470 | NM_002658 | 232 | CTGCCCAAAGAAATTCGGA | 47.8 | Test | Training | 169 | |
| 507 | NM_003391 | 792 | ATTTGCCCGCGCATTGTG | 27.2 | Test | Training | 170 | |
| 508 | NM_003391 | 2171 | AGAAGATGAATGGTCTGGC | 69.4 | Test | Training | 171 | 40 |
| 509 | NM_003391 | 981 | AACGGGCGATTATCTCTGG | 43.3 | Test | Training | 172 | |
| 540 | NM_002387 | 3490 | GACTTAGAGCTGGGAATCT | 83.7 | Test | Training | 173 | |
| 541 | NM_002387 | 4098 | AGTTGAGGAGTTTCTGCA | 86.1 | Test | Training | 174 | |
| 542 | NM_002387 | 1930 | GGATTATATCCAGCAGCTC | 82.3 | Test | Training | 175 | |
| 585 | NM_014885 | 509 | GTGGCTGGATTTCATGTTCC | 81.5 | Training | Training | 176 | |
| 586 | NM_014885 | 798 | CAAGGCATCCGTTATATCT | 84.7 | Training | Training | 177 | |
| 587 | NM_014885 | 270 | ACCAGGATTTGGAGTGGAT | 84.7 | Training | Training | 178 | |

| | | | | | | |
|------|-----------|------|-----------------------|------|----------|-----|
| 639 | NM_001274 | 250 | CTGAAGAAGCAGTCGCAGT | 77.7 | | 179 |
| 640 | NM_001274 | 858 | ATCGATTCTGCTCCTCTAG | 86.2 | | 180 |
| 641 | NM_001274 | 1332 | TGCCTGAAAGAGACTTGTG | 85.4 | | 181 |
| 651 | NM_001259 | 807 | TCTTGGACGTGATTGGACT | 89.8 | Training | 182 |
| 652 | NM_001259 | 1036 | AGAAAACCTGGATTCCCAC | 88.9 | Training | 183 |
| 653 | NM_001259 | 556 | ACCACAGAACATTCTGGTG | 89.3 | Training | 184 |
| 672 | NM_003161 | 2211 | GAAAGCCAGACAACCTTCTG | 87.1 | Test | 185 |
| 673 | NM_003161 | 1223 | CTCTCAGTGAAAGTGCCAA | 91.2 | Test | 186 |
| 674 | NM_003161 | 604 | GACACTGCCTGCTTTTACT | 98.1 | Test | 187 |
| 678 | NM_004972 | 3526 | AAGAACCTGGTGAAAGTCC | 57.2 | Training | 188 |
| 679 | NM_004972 | 4877 | GAAGTGCAGCAGGTTAAGA | 54.8 | Training | 189 |
| 680 | NM_004972 | 1509 | AGCCGAGTTGTA ACTATCC | 74.9 | Training | 190 |
| 684 | NM_007194 | 1245 | GATCACAGTGGCAATGGAA | 80.9 | | 191 |
| 685 | NM_007194 | 1432 | AAACTCTTGGAAAGTGGTGC | 39.2 | | 192 |
| 686 | NM_007194 | 2269 | ATGAATCCACAGCTCTACC | 44.6 | | 193 |
| 687 | NM_007313 | 3866 | GAATGGAAGCCTGAACTGA | 92.4 | Test | 194 |
| 688 | NM_007313 | 2451 | AGACATCATGGAGTCCAGC | 5.0 | Test | 195 |
| 689 | NM_007313 | 1296 | CAAGTTCTCCATCAAGTCC | 91.1 | Test | 196 |
| 711 | NM_139049 | 129 | GGAATAGTATGCGCAGCTT | 92.5 | Test | 197 |
| 712 | NM_139049 | 369 | GTGATT CAGATGGAGCTAG | 89.0 | Test | 198 |
| 713 | NM_139049 | 969 | CACCCGTACATCAATGTCT | 77.0 | Test | 199 |
| 858 | NM_001253 | 522 | TCATTGGAAGAACAGCGGC | 0.0 | Test | 200 |
| 859 | NM_001253 | 2571 | AAGAAGACGTT CAGCGACA | 93.5 | Test | 201 |
| 860 | NM_001253 | 911 | AAAAAGCCTGCCCTTGGTT | 88.1 | Test | 202 |
| 1110 | NM_006101 | 1847 | CTTGCAACGTCTGTTAGAG | 72.3 | Test | 203 |
| 1111 | NM_006101 | 999 | CTGAAGGCTTCCTTACAAG | 82.9 | Test | 204 |
| 1112 | NM_006101 | 1278 | CAGAAGTTGTGGAATGAGG | 79.1 | Test | 205 |
| 1182 | NM_016231 | 1302 | GCAATGAGGACAGCTTGTG | 79.8 | Test | 206 |
| 1183 | NM_016231 | 1829 | TGTAGCTTCCACTGGAGT | 79.3 | Test | 207 |
| 1184 | NM_016231 | 1019 | TCTCCTTGTGAACAGCAAC | 62.5 | Test | 208 |
| 1212 | NM_001654 | 1072 | AGTGAAGAACCTGGGGTAC | 79.3 | Test | 209 |
| 1213 | NM_001654 | 595 | GTTCCACCAGCATTGTTCC | 86.2 | Test | 210 |
| 1214 | NM_001654 | 1258 | GAATGAGATGCAGGTGCTC | 86.9 | Test | 211 |
| 1287 | NM_005417 | 2425 | CAATTCGTGGGAGGCATCA | 73.9 | Test | 212 |
| 1288 | NM_005417 | 1077 | GGGGAGTTTGTCTGGACTTT | 66.4 | Test | 213 |
| 1289 | NM_005417 | 3338 | GCAGTGCCTGCCTATGAAA | 68.2 | Test | 214 |
| 1290 | NM_001982 | 3223 | CTAGACCCTAGACCCTAGACT | 63.5 | Test | 215 |
| 1291 | NM_001982 | 3658 | GAGGATGTCAACGGTTATG | 49.4 | Test | 216 |
| 1292 | NM_001982 | 2289 | CAAAGTCTTGCCAGAATC | 45.3 | Test | 217 |
| 1293 | NM_005400 | 249 | GATCGAGCTGGCTGTCTTT | 85.4 | Test | 218 |
| 1294 | NM_005400 | 1326 | GGTCTTAAAGAAGGACGTC | 63.4 | Test | 219 |
| 1295 | NM_005400 | 1848 | TGAGGACGACCTATTTGAG | 0.0 | Test | 220 |
| 1317 | NM_002086 | 465 | TGAGCTGGTGGATTATCAC | 85.5 | Test | 221 |
| 1318 | NM_002086 | 183 | CTGGTACAAGGCAGAGCTT | 95.5 | Test | 222 |
| 1319 | NM_002086 | 720 | CCGGAACGTCTAAGAGTCA | 92.3 | Test | 223 |
| 1332 | NM_006219 | 2925 | TACAGAAAAGTTTGCCCGG | 20.1 | Test | 224 |
| 1333 | NM_006219 | 2346 | AATGAAGCCTTTGTGGCTG | 22.4 | Test | 225 |
| 1334 | NM_006219 | 2044 | GTGCACATTCCTGCTGTCT | 79.0 | Test | 226 |
| 1335 | NM_003600 | 1618 | CCTCCCTATTCAGAAAGCT | 84.2 | Test | 227 |
| 1336 | NM_003600 | 650 | GACTTTGAAATTGGTCGCC | 52.1 | Test | 228 |
| 1337 | NM_003600 | 538 | CACCCAAAAGAGCAAGCAG | 96.3 | Test | 229 |
| 1338 | XM_294563 | 2703 | TAAGCCTGGTGGTGATCTT | 78.1 | Training | 230 |
| 1339 | XM_294563 | 1701 | AAGGTCTTTACGCCAGTAC | 29.5 | Training | 231 |
| 1340 | XM_294563 | 789 | GGAATGTATCCGAGCACTG | 73.5 | Training | 232 |
| | | | | | | 10 |
| | | | | | | 20 |
| | | | | | | 30 |
| | | | | | | 40 |

| | | | | | | | | |
|------|-----------|------|------|------|----------|----------|-----|----|
| 1386 | NM_033360 | 493 | GGA | 91.0 | Test | Training | 233 | |
| 1387 | NM_033360 | 897 | GC | 84.8 | Test | Training | 234 | |
| 1388 | NM_033360 | 704 | GAA | 0.0 | Test | Training | 235 | |
| 1389 | NM_024408 | 4735 | CTTT | 91.6 | Test | Training | 236 | |
| 1390 | NM_024408 | 2674 | CCA | 96.4 | Test | Training | 237 | |
| 1391 | NM_024408 | 5159 | GACT | 95.8 | Test | Training | 238 | |
| 1392 | NM_000435 | 6045 | GCT | 0.0 | Test | Training | 239 | |
| 1393 | NM_000435 | 5495 | TGCC | 0.0 | Test | Training | 240 | |
| 1394 | NM_000435 | 4869 | TGAT | 0.0 | Test | Training | 241 | |
| 1410 | AF308602 | 770 | ATAT | 36.7 | Test | Training | 242 | |
| 1411 | AF308602 | 3939 | AGGC | 81.3 | Test | Training | 243 | 10 |
| 1412 | AF308602 | 1644 | CAC | 81.3 | Test | Training | 244 | |
| 1581 | NM_005633 | 3593 | TAT | 70.8 | Test | Training | 245 | |
| 1582 | NM_005633 | 364 | ATT | 1.4 | Test | Training | 246 | |
| 1583 | NM_005633 | 3926 | CTT | 66.9 | Test | Training | 247 | |
| 1620 | NM_002388 | 1097 | GTCT | 95.0 | Test | Training | 248 | |
| 1621 | NM_002388 | 286 | AGG | 94.6 | Test | Training | 249 | |
| 1622 | NM_002388 | 2268 | TCC | 92.5 | Test | Training | 250 | |
| 1629 | NM_012193 | 3191 | TTGG | 80.0 | Test | Test | 251 | |
| 1630 | NM_012193 | 5335 | CCAT | 85.0 | Test | Test | 252 | |
| 1631 | NM_012193 | 2781 | GTT | 43.1 | Test | Test | 253 | |
| 1632 | NM_004380 | 3708 | GAC | 85.3 | Test | Training | 254 | |
| 1633 | NM_004380 | 339 | TGG | 81.1 | Test | Training | 255 | 20 |
| 1634 | NM_004380 | 5079 | GCAC | 79.0 | Test | Training | 256 | |
| 1641 | NM_017412 | 2331 | CAG | 97.3 | Test | Training | 257 | |
| 1643 | NM_017412 | 2783 | ATGT | 95.7 | Test | Training | 258 | |
| 1695 | NM_001903 | 2137 | TGAC | 38.4 | Test | Training | 259 | |
| 1696 | NM_001903 | 655 | CGT | 97.9 | Test | Training | 260 | |
| 1697 | NM_001903 | 3117 | TGAC | 40.1 | Test | Training | 261 | |
| 1815 | NM_020168 | 3064 | GAG | 85.0 | Training | Training | 262 | |
| 1816 | NM_020168 | 681 | CGAC | 86.1 | Training | Training | 263 | |
| 1817 | NM_020168 | 1917 | TGAG | 72.1 | Training | Training | 264 | |
| 2502 | NM_000271 | 237 | GAG | 87.0 | Training | Training | 265 | |
| 2503 | NM_000271 | 559 | TACT | 76.0 | Training | Training | 266 | |
| 2504 | NM_000271 | 1783 | AACT | 39.0 | Training | Training | 267 | 30 |
| 2505 | NM_000271 | 2976 | GCC | 84.0 | Training | Training | 268 | |
| 2512 | NM_005030 | 245 | GGG | 88.6 | Test | Test | 269 | |
| 2513 | NM_005030 | 1381 | CAC | 90.5 | Test | Test | 270 | |
| 2514 | NM_005030 | 834 | GAG | 91.0 | Test | Test | 271 | |
| 2521 | NM_000314 | 1316 | CCC | 93.0 | Training | Training | 272 | |
| 2522 | NM_000314 | 1534 | CTAT | 89.0 | Training | Training | 273 | |
| 2523 | NM_000314 | 2083 | CAG | 90.0 | Training | Training | 274 | |
| 2524 | NM_006622 | 1928 | CAG | 65.0 | Training | Training | 275 | |
| 2525 | NM_006622 | 586 | TGT | 73.0 | Training | Training | 276 | |
| 2526 | NM_006622 | 1252 | AACC | 70.0 | Training | Training | 277 | |
| 2527 | NM_139164 | 200 | CTG | 79.0 | Training | Training | 278 | |
| 2528 | NM_139164 | 568 | GACA | 71.0 | Training | Training | 279 | 40 |
| 2529 | NM_139164 | 488 | GTCT | 66.0 | Training | Training | 280 | |
| 2530 | NM_139164 | 578 | ACC | 82.0 | Training | Training | 281 | |
| 2546 | NM_014875 | 1090 | TAG | 63.5 | Test | Training | 282 | |
| 2547 | NM_014875 | 1739 | AGAG | 73.2 | Test | Training | 283 | |
| 2548 | NM_014875 | 3563 | GACC | 87.1 | Test | Training | 284 | |
| 2602 | NM_002387 | 2655 | TAG | 71.0 | Test | Training | 285 | |
| 2603 | NM_002387 | 1418 | ACAG | 43.5 | Test | Training | 286 | |

| | | | | | | | | |
|------|-----------|-------|----------------------|------|------|----------|-----|----|
| 2604 | NM_002387 | 941 | GAGAATGAGAGCCTGACTG | 81.0 | Test | Training | 287 | |
| 2605 | NM_016231 | 1683 | GGAAACAGAGTGCCTCTCT | 55.3 | Test | Training | 288 | |
| 2606 | NM_016231 | 915 | CCACTCAGCTCAGATCATG | 82.3 | Test | Training | 289 | |
| 2607 | NM_016231 | 737 | TCTGGTCTCTTGCAAAAGG | 30.3 | Test | Training | 290 | |
| 2611 | NM_004380 | 4230 | ATTTTTGCGGCGCCAGAAT | 79.0 | Test | Training | 291 | |
| 2612 | NM_004380 | 2197 | GAAAAACGGAGGTCGCGTT | 85.9 | Test | Training | 292 | |
| 2613 | NM_004380 | 5701 | GAAAACAAATGCCCGTGC | 55.4 | Test | Training | 293 | |
| 2614 | NM_005978 | 276 | TGGCACTCATCACTGTCAT | 91.8 | Test | Test | 294 | |
| 2615 | NM_005978 | 229 | TGAGAACAGTGACCAGCAG | 91.9 | Test | Test | 295 | |
| 2616 | NM_005978 | 369 | GGGCCCAGGACTGTTGATG | 94.5 | Test | Test | 296 | |
| 2617 | NM_017412 | 3128 | AGAGATGGGCATTGTTTCC | 94.3 | Test | Training | 297 | 10 |
| 2618 | NM_017412 | 814 | GCTCATGGAGATGTTTGGT | 88.7 | Test | Training | 298 | |
| 2619 | NM_017412 | 1459 | AGCATTGCTGTTTCACGCC | 93.1 | Test | Training | 299 | |
| 2620 | NM_001654 | 1902 | TTGAGCTGCTGCAACGGTC | 67.2 | Test | Training | 300 | |
| 2621 | NM_001654 | 1006 | GTCCCCACATTCCAAGTCA | 90.0 | Test | Training | 301 | |
| 2622 | NM_001654 | 2327 | CCTCTCTGGAATTTGTGCC | 85.7 | Test | Training | 302 | |
| 2623 | NM_002658 | 202 | CAAGTACTTCTCCAACATT | 87.2 | Test | Training | 303 | |
| 2624 | NM_002658 | 181 | TGGAGGAACATGTGTGCC | 0.0 | Test | Training | 304 | |
| 2625 | NM_002658 | 436 | TTACTGCAGGAACCCAGAC | 0.0 | Test | Training | 305 | |
| 2629 | NM_006218 | 1334 | TGGCTTTGAATCTTTGGCC | 3.5 | Test | Training | 306 | |
| 2630 | NM_006218 | 2613 | AGGTGCACTGCAGTTCAAC | 53.8 | Test | Training | 307 | |
| 2631 | NM_006218 | 1910 | TTCAGCTAGTACAGGTCCT | 78.0 | Test | Training | 308 | |
| 2632 | NM_003161 | 1834 | TTGATTCCCTCGCGACATCT | 88.3 | Test | Training | 309 | 20 |
| 2633 | NM_003161 | 1555 | GCTTTTCCCATGATCTCCA | 90.7 | Test | Training | 310 | |
| 2634 | NM_003161 | 217 | CTTGGCATGGAACATTGTG | 61.4 | Test | Training | 311 | |
| 2635 | NM_003391 | 2072 | GCCTCAGAAAGGGATTGCT | 79.1 | Test | Training | 312 | |
| 2636 | NM_003391 | 1318 | GCTCTGGATGTGCACACAT | 60.5 | Test | Training | 313 | |
| 2637 | NM_003391 | 1734 | GTGTCTCAAAGGAGCTTTC | 87.1 | Test | Training | 314 | |
| 2641 | AF308602 | 4260 | ATTCAACGGGCTCTGTGC | 0.0 | Test | Training | 315 | |
| 2642 | AF308602 | 1974 | GATCGATGGCTACGAGTGT | 84.0 | Test | Training | 316 | |
| 2643 | AF308602 | 5142 | CATCCCCTACAAGATCGAG | 41.6 | Test | Training | 317 | |
| 2644 | NM_024408 | 8232 | GCAACTTTGGTCTCCTTTC | 91.0 | Test | Training | 318 | |
| 2645 | NM_024408 | 10503 | GCAATTGGCTGTGATGCTC | 86.6 | Test | Training | 319 | |
| 2646 | NM_024408 | 8643 | GAGACAAGTTAACTCGTGC | 89.4 | Test | Training | 320 | |
| 2647 | NM_007313 | 4222 | TCCTGGCAAGAAAGCTTGA | 65.6 | Test | Training | 321 | 30 |
| 2648 | NM_007313 | 3237 | AAACCTCTACACGTTCTGC | 53.5 | Test | Training | 322 | |
| 2649 | NM_007313 | 302 | CTAAAGGTGAAAAGCTCCG | 67.8 | Test | Training | 323 | |
| 2650 | NM_000551 | 631 | GATCTGGAAGACCACCCAA | 70.9 | Test | Training | 324 | |
| 2651 | NM_000551 | 4678 | CAGAACCCAAAAGGGTAAG | 0.0 | Test | Training | 325 | |
| 2652 | NM_000551 | 4382 | AGGAAATAGGCAGGGTGTG | 4.3 | Test | Training | 326 | |
| 2653 | NM_001903 | 1888 | AGCAGTGCTGATGATAAGG | 89.1 | Test | Training | 327 | |
| 2654 | NM_001903 | 2606 | AAGCCATTGGTGAAGAGAG | 91.9 | Test | Training | 328 | |
| 2655 | NM_001903 | 1583 | TGTGTCAATTGCTCTCCAAG | 90.3 | Test | Training | 329 | |
| 2656 | NM_002388 | 842 | GCAGATGAGCAAGGATGCT | 86.8 | Test | Training | 330 | |
| 2657 | NM_002388 | 1754 | GTACATCCATGTGGCCAAA | 94.6 | Test | Training | 331 | |
| 2658 | NM_002388 | 2642 | TGGGTCATGAAAGCTGCCA | 93.1 | Test | Training | 332 | |
| 2662 | NM_005633 | 3251 | GAACACCGTTAACACCTCC | 31.2 | Test | Training | 333 | 40 |
| 2663 | NM_005633 | 2899 | ATAACAGGAGAGATCCAGC | 21.7 | Test | Training | 334 | |
| 2664 | NM_005633 | 2607 | TGGTGTCCITGAGGTTGTC | 75.1 | Test | Training | 335 | |
| 2665 | NM_033360 | 329 | ACCTGTCTCTTGATATTC | 81.4 | Test | Training | 336 | |
| 2666 | NM_033360 | 529 | TAAATGTGATTTGCCTTCT | 47.8 | Test | Training | 337 | |
| 2667 | NM_033360 | 585 | GAAGTTATGGAATTCCTTT | 94.2 | Test | Training | 338 | |
| 2668 | NM_139049 | 745 | CACCATGTCCTGAATTCAT | 80.7 | Test | Training | 339 | |
| 2669 | NM_139049 | 433 | TCAAGCACCTTCATTCTGC | 42.6 | Test | Training | 340 | |

| | | | | | | | | |
|------|-----------|------|----------------------|-------|----------|----------|-----|----|
| 2670 | NM_139049 | 550 | CGAGTTTTATGATGACGCC | 79.9 | Test | Training | 341 | |
| 2671 | NM_002086 | 555 | ATACGTCCAGGCCCTCTTT | 87.9 | Test | Test | 342 | |
| 2672 | NM_002086 | 392 | TGCAGCACTTCAAGGTGCT | 36.9 | Test | Test | 343 | |
| 2673 | NM_002086 | 675 | CGGGCAGACCCGGCATGTTT | 92.6 | Test | Test | 344 | |
| 2674 | NM_004958 | 5024 | GACATGAGAACCTGGCTCA | 77.8 | Test | Training | 345 | |
| 2675 | NM_004958 | 2155 | CTTGCAGGCCTTGTGTTGTG | 83.2 | Test | Training | 346 | |
| 2676 | NM_004958 | 6955 | TAATACAGCTGGGGACGAC | 52.3 | Test | Training | 347 | |
| 2677 | NM_012193 | 467 | AGAACCTCGGCTACAACGT | 71.5 | Test | Test | 348 | |
| 2678 | NM_012193 | 473 | TCGGCTACAACGTGACCAA | 51.3 | Test | Test | 349 | |
| 2679 | NM_012193 | 449 | TCCGCATCTCCATGTGCCA | 37.5 | Test | Test | 350 | |
| 2680 | NM_005400 | 665 | TCACAAAGTGTGCTGGGTT | 43.9 | Test | Training | 351 | 10 |
| 2681 | NM_005400 | 2178 | CCAGGAGGAATTCAAAGGT | 41.6 | Test | Training | 352 | |
| 2682 | NM_005400 | 1022 | GCTCACCATCTGAGGAAGA | 64.2 | Test | Training | 353 | |
| 2686 | NM_001982 | 948 | TGACAGTGGAGCCTGTGTA | 65.8 | Test | Training | 354 | |
| 2687 | NM_001982 | 1800 | CTTTCTGAATGGGGAGCCT | 61.7 | Test | Training | 355 | |
| 2688 | NM_001982 | 2860 | TACACACACCAGAGTGATG | 0.0 | Test | Training | 356 | |
| 2692 | NM_016195 | 5331 | ATGAAGGAGAGTGATCACC | 10.5 | Test | Training | 357 | |
| 2693 | NM_016195 | 4829 | AATGGCAGTGAACACCCT | 67.3 | Test | Training | 358 | |
| 2694 | NM_016195 | 1480 | AAGTTTGTGTCCCAGACAC | 80.5 | Test | Training | 359 | |
| 2695 | NM_000435 | 2107 | AATGGCTTCCGCTGCCTCT | 0.0 | Test | Training | 360 | |
| 2696 | NM_000435 | 5193 | GAACATGGCCAAGGGTGAG | 15.5 | Test | Training | 361 | |
| 2697 | NM_000435 | 7273 | GAGTCTGGGACCTCCTTCT | 0.0 | Test | Training | 362 | |
| 2802 | NM_004523 | 46 | CCAGGGGAGACTCCGGCCCC | 6.7 | Training | Test | 363 | 20 |
| 2803 | NM_004523 | 132 | GGGACCGTCATGGCGTCGC | 8.2 | Training | Test | 364 | |
| 2804 | NM_004523 | 221 | ATTTAATTTGGCAGAGCGG | 0.0 | Training | Test | 365 | |
| 2805 | NM_004523 | 322 | GCTCAAGGAAAACATACAC | 76.2 | Training | Test | 366 | |
| 2806 | NM_004523 | 365 | TACTAAACAGATTGATGTT | 77.9 | Training | Test | 367 | |
| 2807 | NM_004523 | 581 | TACTGATAATGGTACTGAA | 93.8 | Training | Test | 368 | |
| 2808 | NM_004523 | 716 | AGGAGTGATAATTAAGGT | 84.8 | Training | Test | 369 | |
| 2809 | NM_004523 | 852 | GTTTTCTCTGTTACAATAC | 85.4 | Training | Test | 370 | |
| 2810 | NM_004523 | 995 | TGGAAATATAAATCAATCC | 0.0 | Training | Test | 371 | |
| 2811 | NM_004523 | 1085 | ACTAACTAGAATCCTCCAG | 0.0 | Training | Test | 372 | |
| 2812 | NM_004523 | 1174 | AAACTCTGAGTACATTGGA | 81.9 | Training | Test | 373 | |
| 2813 | NM_004523 | 1375 | TAACTGTTCAAGAAGAGCA | 14.1 | Training | Test | 374 | |
| 2814 | NM_004523 | 1570 | AAGAAGAATATATCACATC | 0.0 | Training | Test | 375 | 30 |
| 2815 | NM_004523 | 1706 | AGTTGACCAACACAATGCA | 86.0 | Training | Test | 376 | |
| 2816 | NM_004523 | 2197 | TACATGAACTACAAGAAAA | 90.0 | Training | Test | 377 | |
| 2817 | NM_004523 | 2858 | GACTAAGCTTAATTGCTTT | 87.0 | Training | Test | 378 | |
| 2818 | NM_004523 | 3089 | GGGGCAGTATACTGAAGAA | 64.5 | Training | Test | 379 | |
| 2819 | NM_004523 | 3878 | TTCTTGATATTATTAAGT | 0.0 | Training | Test | 380 | |
| 2820 | NM_004523 | 4455 | TCTATAATTTATATTCTTT | 9.3 | Training | Test | 381 | |
| 2821 | NM_004523 | 4648 | TACAAAGAATAAAATTTTCT | 23.5 | Training | Test | 382 | |
| 2823 | NM_005030 | 45 | CAGCGCAGCTTCGGGAGCA | 72.1 | Training | Test | 383 | |
| 2824 | NM_005030 | 131 | CGGAGTTGCAGCTCCCGGA | 85.7 | Training | Test | 384 | |
| 2825 | NM_005030 | 303 | GGCAAGATTGTGCCCTAAGT | 80.1 | Training | Test | 385 | |
| 2826 | NM_005030 | 346 | GGGAGAAGATGTCCATGGA | 100.0 | Training | Test | 386 | |
| 2827 | NM_005030 | 432 | GACTTCGTGTTTCGTGGTGT | 89.3 | Training | Test | 387 | 40 |
| 2828 | NM_005030 | 519 | GCCCCGATACTACCTACGGC | 86.2 | Training | Test | 388 | |
| 2829 | NM_005030 | 648 | GGACTGGCAACCAAAGTCG | 86.7 | Training | Test | 389 | |
| 2830 | NM_005030 | 777 | TGTATCATGTATACCTTGT | 84.3 | Training | Test | 390 | |
| 2831 | NM_005030 | 821 | TTCTTGCCATAAAGAGACC | 26.8 | Training | Test | 391 | |
| 2832 | NM_005030 | 907 | TCCAGAAGATGCTTCAGAC | 90.8 | Training | Test | 392 | |
| 2833 | NM_005030 | 952 | ACGAGCTGCTTAATGACGA | 87.7 | Training | Test | 393 | |
| 2834 | NM_005030 | 1038 | TCGATTGCTCCAGCAGCC | 31.4 | Training | Test | 394 | |

| | | | | | | | |
|------|-----------|------|----------------------|------|----------|------|-----|
| 2835 | NM_005030 | 1082 | CACAGTCCTCAATAAAGGC | 62.9 | Training | Test | 395 |
| 2836 | NM_005030 | 1214 | CAATGCCTCCAAGCCCTCG | 0.0 | Training | Test | 396 |
| 2837 | NM_005030 | 1300 | AGTGGGTGGACTATTCGGA | 84.9 | Training | Test | 397 |
| 2838 | NM_005030 | 1515 | TACATGAGCGAGCACTTGC | 20.3 | Training | Test | 398 |
| 2839 | NM_005030 | 1860 | CTCAAGGCCTCCTAATAGC | 74.2 | Training | Test | 399 |
| 2840 | NM_005030 | 1946 | CCGCGGTGCCATGTCTGCA | 79.7 | Training | Test | 400 |
| 2841 | NM_005030 | 2075 | CCCCTCCCCCTCAACCCCA | 34.6 | Training | Test | 401 |
| 3041 | NM_014875 | 4629 | ATTTTCTAGAAAACGGTAA | 91.8 | | | 402 |
| 3042 | NM_014875 | 77 | GAGGGGCGAAGTTTCGGCA | 71.2 | | | 403 |
| 3043 | NM_014875 | 243 | CTGGGACCGGGAAGCCGGA | 0.0 | | | 404 |
| 3044 | NM_014875 | 5094 | CTTCTACTTCTGTTGGCAG | 85.9 | | | 405 |
| 3045 | NM_014875 | 4354 | ACTTACTATTTCAGACTGCA | 85.7 | | | 406 |
| 3046 | NM_014875 | 524 | GCCCTCACCCACAGTAGCC | 68.1 | | | 407 |
| 3047 | NM_014875 | 5349 | CAGAGGAATGCACACCCAG | 73.6 | | | 408 |
| 3048 | NM_014875 | 4824 | GATTGATTAGATCTCTTGA | 91.3 | | | 409 |
| 3049 | NM_014875 | 3014 | GTGAGTATTATCCCAGTTG | 41.5 | | | 410 |
| 3050 | NM_014875 | 2959 | ATCTGGGGTGCTGATTGCT | 46.3 | | | 411 |
| 3051 | NM_014875 | 1514 | GTGACAGTGGCAGTACGCG | 67.7 | | | 412 |
| 3052 | NM_014875 | 1114 | TCAGACTGAAGTTGTTAGA | 80.8 | | | 413 |
| 3053 | NM_014875 | 2079 | GTTGGCTAGAATTGGGAAA | 91.8 | | | 414 |
| 3054 | NM_014875 | 3560 | GAAGACCATAGCATCCGCC | 74.8 | | | 415 |

Test: テスト Training: トレーニング

10

20

【 0 2 4 7 】

表IIIは、IGF1R-73早期動力学遺伝子とIGF1R-73のアンチセンス鎖とのアラインメントを示す。表IVAとBは、IGF1R-73早期動力学遺伝子とIGF1R-73のそれぞれセンス鎖とアンチセンス鎖とのアラインメントを示す。表VAとBは、IGF1R-193早期動力学遺伝子とMAPK14-193のそれぞれセンス鎖とアンチセンス鎖とのアラインメントを示す。これらの表において、大文字は相補的塩基、小文字は非相補的塩基、灰色の影はsiRNA鎖上のアラインメントのギャップの位置、Contig start (コンティグスタート) は最も長い連続的アラインメント区間の開始位置、そしてContig length (コンティグ長さ) は、最も長い連続的アラインメント区間を示す。

【表 3】

表 III IGF1R-73 早期動力学遺伝子と IGF1R-73 のアンチセンス鎖とのアラインメント

| | | IGF1R-73 アンチセンス鎖配列 | | | | | | | | | | | | | | | | | | | |
|----------|----|--------------------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|----|
| | | G | A | G | G | T | A | A | C | A | G | A | G | G | T | C | A | G | C | A | |
| NM_0060 | 70 | t | A | G | G | c | t | t | g | t | t | g | G | G | T | C | A | G | C | g | |
| NM_0062 | 91 | a | g | c | a | T | c | t | t | g | G | A | c | G | T | C | A | G | C | A | 10 |
| NM_0018 | 14 | t | A | t | a | a | A | g | a | A | G | A | G | G | g | C | A | G | C | A | |
| NM_0143 | 20 | c | A | G | G | T | A | A | C | A | G | A | G | G | a | C | A | G | t | A | |
| AF167706 | | c | A | G | a | a | g | c | C | A | G | A | G | c | T | C | A | G | t | A | |
| AF179224 | | t | g | G | c | c | t | t | g | g | c | A | G | G | T | C | A | G | C | A | |
| AK000808 | | a | g | a | G | g | A | g | g | A | t | g | G | G | T | C | A | G | C | c | |
| NM_0057 | 65 | t | g | a | G | T | A | A | C | A | G | A | G | G | a | C | A | G | C | t | |
| NM_0190 | 86 | G | A | G | G | g | A | A | C | A | G | A | G | G | T | C | A | G | g | A | 20 |

【表 4 A】

表 IVA IGF1R-73 のセンス鎖との同一性を示す IGF1R-73 12 時間シグネチャー遺伝子のアラインメント

| アクセッション | IGF1R-73 センス鎖配列 | | | | | | | | | | | | | | | | コンテイング | | | | | |
|---------------|-----------------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|--------|---|---|------|----|----|
| | T | G | C | T | G | A | C | C | T | C | T | G | T | T | A | C | C | T | C | スタート | 長さ | |
| Contig45859_R | | | | | | | | | | | | | | | | | | | | | | |
| C | - | - | - | - | - | A | t | a | a | t | c | a | T | T | A | C | C | T | t | 8 | 6 | |
| NM_005638 | T | a | C | T | G | g | C | a | T | C | T | G | T | T | A | t | C | T | a | 9 | 7 | |
| NM_003369 | a | G | C | T | G | A | g | C | T | C | T | G | c | T | g | t | a | c | t | 2 | 5 | 10 |
| Z49105 | c | c | C | c | t | c | a | C | T | C | T | G | T | T | t | C | C | T | g | 8 | 7 | |
| AF070648 | a | t | t | a | c | c | a | a | c | C | T | G | T | T | A | C | C | T | a | 10 | 9 | |
| NM_003721 | c | G | C | T | G | A | C | C | T | C | a | c | c | a | c | C | g | a | a | 2 | 9 | |
| AF167706 | g | G | g | a | G | A | g | t | g | C | T | G | T | g | A | C | C | T | C | 15 | 5 | |
| AF072928 | a | G | C | a | G | A | - | C | T | C | T | G | c | T | A | C | C | c | g | 8 | 5 | |
| AK000745 | T | G | t | T | a | A | C | a | T | C | T | G | T | T | t | C | a | g | g | 9 | 6 | |
| NM_014637 | T | G | C | T | G | c | C | C | T | C | a | t | a | g | c | a | g | a | g | 1 | 5 | |
| AK001846 | T | G | g | g | a | g | C | t | T | C | T | G | T | T | A | a | C | T | t | 9 | 7 | |
| NM_005765 | a | t | g | g | a | t | t | g | a | a | T | G | T | T | A | C | C | T | g | 11 | 8 | |
| NM_006358 | a | a | C | T | G | A | C | a | g | C | T | G | c | c | A | C | C | T | t | 3 | 5 | |
| NM_004157 | g | c | C | T | G | A | C | C | T | C | g | t | c | g | A | a | t | T | C | 3 | 8 | 20 |
| NM_014814 | a | G | C | T | G | A | a | C | T | C | T | t | c | c | t | t | g | a | C | 2 | 5 | |
| NM_019086 | T | t | t | T | t | c | t | C | a | C | T | G | T | T | A | C | t | g | a | 10 | 7 | |
| NM_003011 | c | a | C | c | G | A | g | C | T | C | T | G | T | g | g | g | a | a | a | 8 | 6 | |
| NM_017443 | a | c | C | T | t | A | C | T | C | T | G | c | T | A | C | t | g | t | 9 | 5 | | |
| NM_006804 | a | G | C | T | G | A | C | C | c | g | a | G | a | c | t | t | g | g | a | 2 | 7 | |
| AK000808 | c | G | C | g | t | l | a | a | c | C | T | G | T | T | c | C | C | T | t | 10 | 5 | |
| NM_016031 | g | G | C | T | a | c | C | C | T | C | T | G | a | T | g | g | g | g | t | 7 | 6 | |
| Contig1462_R | | | | | | | | | | | | | | | | | | | | | | |
| C | a | a | a | g | a | A | g | C | T | C | T | G | T | c | A | C | a | a | t | 8 | 6 | |
| NM_005321 | c | c | a | g | t | g | C | C | T | C | T | G | c | T | t | C | C | g | g | 7 | 6 | |
| NM_013397 | c | a | t | T | G | A | C | C | T | C | T | G | g | c | A | C | C | T | C | 3 | 7 | |
| NM_003765 | c | G | a | a | G | A | C | C | c | C | T | t | T | T | t | t | g | T | a | 5 | 4 | 30 |
| Contig47067_R | | | | | | | | | | | | | | | | | | | | | | |
| C | g | a | a | T | c | t | t | a | a | g | T | G | T | T | A | C | a | g | g | 11 | 6 | |
| Contig49512_R | | | | | | | | | | | | | | | | | | | | | | |
| C | T | G | C | T | G | A | t | g | c | t | T | c | c | a | g | t | a | g | a | 1 | 6 | |
| Contig47994_R | | | | | | | | | | | | | | | | | | | | | | |
| C | T | t | C | c | t | A | g | t | T | C | T | G | T | T | t | t | C | T | g | 9 | 6 | |
| Contig48185_R | | | | | | | | | | | | | | | | | | | | | | |
| C | a | a | g | a | G | c | C | C | T | - | T | t | T | T | A | C | C | T | g | 13 | 6 | |

【表 4 B】

表 IVB IGF1R-73 のアンチセンス鎖との同一性を示す IGF1R-73 12 時間シグネチャー遺伝子のアラインメント

| アクセッション | IGF1R-73 アンチセンス鎖配列 | | | | | | | | | | | | | | | コンティグ | | | | | | |
|---------------|--------------------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|-------|---|---|---|------|----|--|
| | G | A | G | G | T | A | A | C | A | G | A | G | G | T | C | A | G | C | A | スタート | 長さ | |
| NM_019086 | G | A | G | G | g | A | A | C | A | G | A | G | G | T | C | A | G | g | A | 6 | 12 | |
| NM_005765 | t | g | a | G | T | A | A | C | A | G | A | G | G | a | C | A | G | C | t | 4 | 10 | |
| NM_016031 | c | A | c | a | c | A | g | a | A | G | A | G | G | T | C | A | G | C | A | 9 | 11 | |
| NM_014637 | c | t | a | a | g | g | A | C | A | G | A | G | G | T | C | A | G | a | t | 7 | 11 | |
| NM_003011 | G | A | G | t | T | A | g | g | A | G | A | G | G | T | C | A | t | C | A | 9 | 8 | |
| AK000745 | t | g | a | a | a | t | i | a | A | G | A | G | G | T | C | A | G | C | A | 9 | 11 | |
| AF070648 | c | A | a | c | T | g | A | a | t | G | A | G | G | T | C | A | G | C | A | 10 | 10 | |
| NM_005638 | G | A | G | G | T | g | A | C | A | G | A | G | c | a | g | A | t | t | c | 7 | 6 | |
| NM_003369 | G | A | G | G | T | t | A | C | t | G | c | a | G | T | C | A | G | a | t | 1 | 5 | |
| AF072928 | c | g | a | a | c | t | A | t | A | G | A | G | G | g | C | A | G | C | A | 9 | 5 | |
| NM_014814 | t | A | c | t | a | A | A | C | A | G | A | G | t | T | C | A | a | a | A | 6 | 7 | |
| AF167706 | c | A | G | a | a | g | c | C | A | G | A | G | c | T | C | A | G | t | A | 8 | 5 | |
| AK001846 | t | A | c | a | a | t | g | t | A | t | t | G | G | T | C | A | G | C | A | 12 | 8 | |
| NM_017443 | c | A | G | a | T | c | A | C | A | G | A | G | G | - | C | A | G | C | c | 7 | 7 | |
| NM_006358 | t | c | t | G | g | g | t | C | A | a | A | G | G | T | C | A | a | C | A | 11 | 6 | |
| NM_003765 | G | g | a | a | T | A | A | C | A | G | A | G | a | g | a | t | a | C | t | 5 | 8 | |
| Z49105 | a | t | t | t | a | t | g | a | A | G | A | G | a | T | C | A | G | C | g | 14 | 5 | |
| NM_004157 | l | g | a | t | g | A | A | C | A | G | A | G | a | T | C | A | G | a | c | 6 | 7 | |
| AK000808 | a | g | a | G | g | A | g | g | A | t | g | G | G | T | C | A | G | C | c | 12 | 7 | |
| NM_003721 | c | c | G | G | T | g | t | C | A | G | A | G | G | g | C | c | c | g | c | 8 | 6 | |
| Contig1462_R | | | | | | | | | | | | | | | | | | | | | | |
| C | a | t | a | t | a | c | A | C | A | t | t | t | G | T | C | A | G | C | A | 13 | 7 | |
| NM_006804 | a | t | a | c | c | A | A | C | A | c | A | G | G | c | a | t | c | C | g | 6 | 4 | |
| NM_005321 | c | A | a | G | g | c | c | a | A | G | A | a | G | c | C | A | G | C | A | 15 | 5 | |
| Contig47067_R | | | | | | | | | | | | | | | | | | | | | | |
| C | t | A | c | a | c | t | t | t | A | a | A | a | G | T | C | A | G | C | A | 13 | 7 | |
| Contig49512_R | | | | | | | | | | | | | | | | | | | | | | |
| C | G | A | G | G | c | A | A | C | A | a | g | a | t | T | a | A | a | a | A | 1 | 4 | |
| Contig47994_R | | | | | | | | | | | | | | | | | | | | | | |
| C | G | t | c | t | g | g | A | a | A | G | A | G | c | T | C | A | t | C | t | 9 | 4 | |
| Contig48185_R | | | | | | | | | | | | | | | | | | | | | | |
| C | G | A | G | a | a | A | A | C | A | G | t | c | a | a | a | t | t | t | A | 6 | 5 | |
| Contig45859_R | | | | | | | | | | | | | | | | | | | | | | |
| C | t | t | a | G | T | t | t | a | A | a | A | G | G | T | C | A | a | t | c | 11 | 6 | |
| NM_013397 | G | A | G | t | T | A | t | C | A | G | A | G | G | g | C | - | - | - | - | 8 | 6 | |

【表 5 A】

表 VA MAPK14-193 のセンス鎖との同一性を示す MAPK14-193 12 時間シグネチャー遺伝子のアラインメント

| アクセッション | MAPK14-193 センス鎖配列 | | | | | | | | | | | | | | コンテイング スタート | コンテイング 長さ | | | | | | |
|---------|-------------------|---|---|---|---|---|---|---|---|---|---|---|---|---|----------------|--------------|---|---|---|----|----|----|
| | C | C | T | A | C | A | G | A | G | A | A | C | T | G | | | | C | G | G | T | T |
| NM_0022 | | | | | | | | | | | | | | | | | | | | | | |
| 71 | g | C | T | A | C | A | G | A | G | A | A | C | T | G | C | a | t | c | T | 2 | 14 | |
| NM_0177 | | | | | | | | | | | | | | | | | | | | | | |
| 48 | t | t | T | A | C | A | G | A | G | A | A | C | T | t | C | G | G | T | a | 3 | 11 | 10 |
| NM_0210 | | | | | | | | | | | | | | | | | | | | | | |
| 33 | t | C | T | A | C | A | G | A | G | A | A | C | T | G | C | a | G | c | c | 2 | 14 | |
| NM_0041 | | | | | | | | | | | | | | | | | | | | | | |
| 65 | a | g | g | c | C | t | c | A | G | A | A | C | T | G | C | G | G | g | T | 8 | 10 | |
| NM_0045 | | | | | | | | | | | | | | | | | | | | | | |
| 83 | g | C | T | A | C | A | G | A | G | g | c | a | g | G | C | c | a | g | c | 2 | 8 | |
| NM_0132 | | | | | | | | | | | | | | | | | | | | | | |
| 42 | a | a | T | A | t | t | t | c | t | t | c | C | T | G | C | G | G | T | T | 12 | 8 | |
| NM_0029 | | | | | | | | | | | | | | | | | | | | | | |
| 46 | g | a | a | g | C | A | G | g | G | A | A | C | T | t | t | G | G | T | g | 9 | 5 | 20 |
| NM_0022 | | | | | | | | | | | | | | | | | | | | | | |
| 00 | t | C | T | A | C | A | G | c | c | A | g | C | T | G | g | a | G | g | c | 2 | 6 | |

【表 5 B】

表 VB MAPK14-193 のアンチセンス鎖との同一性を示す MAPK14-193 12 時間シグネチャー遺伝子のアラインメント

| アクセッション | MAPK14-193 アンチセンス鎖配列 | | | | | | | | | | | | | | コンテイング スタート | コンテイング 長さ | | | | | | |
|----------|----------------------|---|---|---|---|---|---|---|---|---|---|---|---|---|----------------|--------------|---|---|---|----|---|----|
| | A | A | C | C | G | C | A | G | T | T | C | T | C | T | | | | G | T | A | G | G |
| NM_00227 | | | | | | | | | | | | | | | | | | | | | | |
| 1 | g | g | C | C | G | C | c | G | T | T | C | T | C | c | t | a | A | G | a | 8 | 6 | |
| NM_02103 | | | | | | | | | | | | | | | | | | | | | | |
| 3 | A | A | a | C | G | a | A | t | T | T | C | a | C | T | G | T | g | t | a | 13 | 4 | |
| NM_01324 | | | | | | | | | | | | | | | | | | | | | | |
| 2 | g | t | a | a | G | C | A | G | T | c | C | T | g | T | G | a | g | t | G | 5 | 5 | |
| NM_00458 | | | | | | | | | | | | | | | | | | | | | | |
| 3 | g | A | C | a | a | C | A | G | T | T | t | g | C | T | G | T | t | c | a | 6 | 5 | |
| NM_00416 | | | | | | | | | | | | | | | | | | | | | | |
| 5 | c | c | t | C | t | C | g | G | T | T | C | T | C | T | a | g | g | t | c | 8 | 7 | |
| NM_00294 | | | | | | | | | | | | | | | | | | | | | | |
| 6 | g | t | a | C | G | C | - | G | T | T | t | T | C | T | G | g | A | G | c | 12 | 4 | 40 |
| NM_00220 | | | | | | | | | | | | | | | | | | | | | | |
| 0 | c | A | C | C | G | C | A | G | a | c | a | g | a | c | c | c | c | t | c | 2 | 7 | |
| NM_01774 | | | | | | | | | | | | | | | | | | | | | | |
| 8 | c | c | a | g | a | a | A | a | c | T | C | T | C | T | G | c | A | G | a | 10 | 6 | |

鎖優先は、IGF1R-73とMAPK14-193のアラインメントデータに基づく投票法と3'偏り法を使用して評価し、それぞれ表VIとVIIに示す。

【表 6】

表 VI. 鎖偏りを予測する投票法

| | MAPK14-193 | | IGF1R-73 | |
|------------------|------------|-----|----------|-----|
| 時点 (時) | 12 | 24 | 12 | 24 |
| シグナチャ 遺伝子 | 11 | 96 | 29 | 198 |
| アンチセンス 同一性スコア | 3 | 87 | 45 | 266 |
| センス 同一性スコア | 45 | 142 | 23 | 164 |
| 予測スコア | 5 | 47 | 14 | 97 |

10

20

【0248】

予測されるスコアは、シグナチャサイズに対応し、24,975のmRNA配列とアラインメントした377のsiRNA(表II)について遺伝子当たりの平均スコアから計算した。遺伝子当たり平均スコアは0.491であり、標準偏差は0.0454であった。

【0249】

センス同一性スコアは、MAPK14-193の12時間と24時間でアンチセンス同一性スコアを超えた。センス鎖との同一性によるアラインメントはsiRNAのアンチセンス鎖に相補的であり、細胞中でこれとハイブリダイズすると予測されるため、これは、アンチセンス鎖の過剰のオフターゲット活性に対応した。12時間では、センス同一性スコアのみがバックグラウンドより大きく、このsiRNAの直接オフターゲットシグナチャが、アンチセンス鎖の活性のためである可能性があった。24時間では、両方の鎖がバックグラウンドより大きく、これはおそらく二次的制御の開始のためである。しかし過剰のセンス同一性も検出できた。

30

【0250】

アンチセンス同一性スコアは、IGF1R-73の12時間と24時間でセンス同一性スコアを超えた。アンチセンス鎖との同一性によるアラインメントはsiRNAのセンス鎖に相補的であり、細胞中でこれとハイブリダイズすると予測されるため、これは、センス鎖の過剰のオフターゲット活性に対応した。12時間では、アンチセンス同一性スコアのみがバックグラウンドより大きく、このsiRNAの直接オフターゲットシグナチャが、センス鎖の活性のためである可能性があった。センス同一性スコアはわずかにバックグラウンドより大きく、アンチセンス鎖は目的の標的の弱いサイレンシングを示し、同様に活性であることがわかったため、これは予想外ではなかった。24時間では、両方の鎖がバックグラウンドより大きく、これはおそらく二次的制御の寄与により補助されたためであった。しかし過剰のアンチセンス同一性も検出できた。

40

【表 7】

表 VII. 鎖偏りを予測する 3'偏り法

| | MAPK14-193 | | IGF1R-73 | |
|------------------|------------|----|----------|-----|
| 時点 (時) | 12 | 24 | 12 | 24 |
| シグナチャ 遺伝子 | 11 | 96 | 29 | 198 |
| アンチセンス 同一性スコア | 0 | 9 | 10 | 58 |
| センス 同一性スコア | 3 | 19 | 2 | 20 |
| 予測スコア | 1 | 9 | 3 | 19 |

10

20

【0251】

予測されるスコアは、シグナチャサイズに対応し、24,975のmRNA配列とアラインメントした377のsiRNA(表II)について遺伝子当たりの平均スコアから計算した。遺伝子当たり平均スコアは0.0945であり、標準偏差は0.0221であった。

【0252】

センス同一性スコアは、MAPK14-193の12時間と24時間でアンチセンス同一性スコアを超える。センス鎖との同一性によるアラインメントはsiRNAのアンチセンス鎖に相補的であり、細胞中でこれとハイブリダイズすると予測されるため、これは、アンチセンス鎖の過剰のオフターゲット活性に対応する。12時間と24時間の両方で、センス同一性スコアのみがバックグラウンドより大きく、このsiRNAの直接オフターゲットシグナチャが、アンチセンス鎖の活性のためである可能性がある。

30

【0253】

アンチセンス同一性スコアは、IGF1R-73の12時間と24時間でセンス同一性スコアを超える。アンチセンス鎖との同一性によるアラインメントはsiRNAのセンス鎖に相補的であり、細胞中でこれとハイブリダイズすると予測されるため、これは、センス鎖の過剰のオフターゲット活性に対応する。12時間では、アンチセンス同一性スコアのみがバックグラウンドより大きく、このsiRNAの直接オフターゲットシグナチャが、センス鎖の活性のためである可能性がある。12時間ではセンス同一性スコアはバックグラウンドより小さく、この方法により直接シグナチャに寄与するとして検出できない。24時間では、両方の鎖がバックグラウンドより大きく、これはおそらく二次的制御の寄与により補助されたためである。しかし過剰のアンチセンス同一性も検出できた。

40

【0254】

投票法と比較して、3'偏り法は、siRNA鎖の1つの二次的制御または弱い寄与のために、検出に対して比較的非感受性のようである。

【0255】

実施例4

6.4. 実施例4: siRNAの鎖優先の予測のための塩基組成モデル

良いsiRNAと悪いsiRNAとのG/C含量の平均差は、siRNA機能的モチーフと耐性モチーフとを分けるG/C PSSMのモデルを提供する。siRNAの両方の鎖が活性があることは公知のように(例えば、Elbashirら、2001, Genes Dev.15:188-200)、siRNAのセンス鎖とアンチセ

50

ンス鎖の両方のG/C含量が、良いsiRNAと悪いsiRNAのG/C含量の平均差から得られるsiRNA機能的標的モチーフG/C含量のモデルにいかとうまく適合するかを発見することは興味深い。このために、良いsiRNAと悪いsiRNAの逆相補体を調べた。これらの逆相補体は、siRNA A2本鎖のセンス鎖の仮説の完全一致標的部位に対応する。逆相補体を、siRNA2本鎖のアンチセンス鎖の実際の完全一致標的部位で示される実際の良いsiRNAと悪いsiRNAに比較した。

【0256】

図13Aは、19量体のsiRNA2本鎖領域内の、悪いsiRNAの逆相補体のG/C含量と悪いsiRNA自体のG/C含量と差を示す。比較のために、良いsiRNAのG/C含量と悪いsiRNAのG/C含量の差を示す。曲線は5のウィンドウ（または5のウィンドウの一部、配列の末端で）にわたって平滑化される。 10

【0257】

図13Bは、19量体のsiRNA2本鎖領域内の、良いsiRNAの逆相補体のG/C含量と悪いsiRNA自体のG/C含量と差を示す。比較のために、良いsiRNAのG/C含量と悪いsiRNAのG/C含量の差を示す。曲線は5のウィンドウ（または5のウィンドウの一部、配列の末端で）にわたって平滑化される。

【0258】

悪いsiRNAの逆相補体は、良いsiRNAよりも悪いsiRNAとより大きく異なることがわかった。平均すると、悪いsiRNAの逆相補体は、良いsiRNAより5'末端でより強いG/C含量を示し、3'末端では良いsiRNAとG/C含量が同様であった。これに対して良いsiRNAの逆相補体は、良いsiRNAとよりも悪いsiRNAに実質的によりよく似ていることがわかった。平均すると、良いsiRNAの逆相補体は、5'末端で悪いsiRNAとほとんど違いがなく、3'末端では悪いsiRNAよりG/C含量がわずかに少なかった。 20

【0259】

これらの結果は、G/C PSSMが悪いsiRNAとして強いセンス鎖を有するsiRNAを、良いsiRNA Aとして弱いセンス鎖を有するsiRNAと区別していることを意味するようである。そのG/C PSSMスコアが逆相補体のG/C PSSMスコアより大きいsiRNAは、そのセンス鎖より活性が大きいアンチセンス鎖を有すると予測される。これに対してそのG/C PSSMスコアが逆相補体のG/C PSSMスコアより小さいsiRNAは、そのアンチセンス鎖より活性が大きいセンス鎖を有すると予測される。 30

【0260】

効力の上昇は、より大きいアンチセンス鎖活性とより小さいセンス鎖活性に対応することが証明されている。すなわち本発明のG/C PSSMは、優勢なアンチセンス鎖活性のためにより大きな効力を有する良いsiRNA（「アンチセンス鎖活性」siRNA）を、優勢なセンス鎖活性を有するsiRNA（「センス活性」siRNA）と区別するようである。

【0261】

鎖の偏りの予測についてのsiRNAのG/C PSSMとその逆相補体の比較の妥当性を、3'偏り法によるsiRNA発現プロファイルからの鎖の偏りの推定との比較により試験した。

【0262】

siRNAとその逆相補体は、加重マトリックスとして図13Aに記載の19量体内の悪いsiRNAと悪いsiRNAとの平滑化したG/C含量差を使用してスコア化した。各鎖のG/C PSSMは、曲線モデルPSSMのスコア計算法に従う、siRNA鎖G/C含量とG/C含量差マトリックスのドット積であった。 40

【0263】

siRNAは、アンチセンス同一性スコアがセンス同一性スコアを超えた場合、発現プロファイル分析の3'偏り法によりセンス活性と呼んだ。siRNAは、その逆相補体G/C PSSMスコアがそれ自体のG/C PSSMスコアを超えた場合、G/C PSSM法によりセンス活性と呼んだ。

【0264】

図14において、測定したサイレンシング効力により区分けし、発現プロファイルとG/C PSSM法によりセンス活性コールの頻度を比較した。これらの方法は異なる分析に基づくが 50

、一致は極めて良好であった。いずれも、低サイレンシングsiRNA対高サイレンシングsiRNAのより高い比率がセンス活性であると予測される。(siRNA G/C PSSMスコア - 逆相補体 G/C PSSMスコア) 対 \log_{10} (センス同一性スコア / アンチセンス同一性スコア) の相関係数は、図14で分類した61個のsiRNAのセットについて0.59である。

【0265】

実施例5

6.5. 実施例5：センス鎖修飾はセンス活性siRNAの特異性を変化させる

本例は、オフターゲットサイレンシングであるsiRNA鎖の不活性化がsiRNAの特異性を改良し得ることを例示する。

【0266】

図15は、センス鎖修飾が、「センス活性」siRNAの特異性を変化させることを示す。siRNAは、100nMの濃度でHeLa細胞中にトランスフェクトされた。トランスフェクションの12時間後にRNAを抽出し、モックトランスフェクトした細胞からのRNAに対してプロファイリングした。オンターゲットsiRNAはセンス鎖中に2つの2'-O-メチル置換を含有して、この鎖を不活性化し、これが遺伝子サイレンシングで機能することを妨害した。以前はアンチセンス活性であると測定されたsiRNA 71のシグナチャは、センス鎖不活性化により変化しなかった。これに対してsiRNA 73のシグナチャ(両方の鎖が活性である)は、センス鎖不活性化により大きく変化した。この状況で、センス鎖は遺伝子サイレンシングで機能することが妨害され、リスク(RISC)と相互作用することが妨害され、従ってアンチセンス鎖が優勢になった。この結果は、センス鎖の化学的不活性化が、センス活性siRNAのオフターゲットシグナチャを変化させることができるが、それを排除しないことを示す。

10

20

【0267】

図16はオフターゲット制御における鎖の偏りを示す。siRNAをシグナチャ遺伝子とアラインメントさせた。センス鎖またはアンチセンス鎖とのアラインメントに対する偏りを、連続的アラインメント長の比較により測定した。ランダムセットの予測されるスコアを、灰色で示す。siRNA 73について、非修飾siRNAのシグナチャはアンチセンス鎖との同一性を示し、センス鎖によるサイレンシングを示している。このsiRNAのセンス不活性化型では、シグナチャ遺伝子はセンス鎖との同一性を示し、アンチセンス鎖の活性によるサイレンシングを示している。これは両方の鎖がsiRNA 73について機能性であり、センス鎖が優勢であることを証明する。この鎖の不活性化は、アンチセンス鎖が優勢になることを可能にした。siRNA 71について、シグナチャ遺伝子はセンス鎖との同一性を示し、アンチセンス鎖がこの2本鎖中では優先的に活性であり、センス鎖の化学的不活性化がこのシグナチャを変化させないことを証明している。

30

【0268】

図17は、センス鎖修飾が「センス活性」siRNAの力価を上昇させることを示す。siRNA力価曲線は、より低い濃度でサイレンシング効率を維持する能力として定義される力価を評価するのに使用した。siRNAを記載の濃度でHeLa細胞中にトランスフェクトさせた。トランスフェクションの24時間後にRNAを抽出し、オンターゲットサイレンシングをリアルタイムPCRにより測定した。センス活性siRNA(黒の実線)は、アンチセンス活性siRNA(灰色の実線)と比較して力価が低下していることを示した。このアッセイはオンターゲットサイレンシングを測定し、これはアンチセンス鎖の活性の関数である。2本鎖73中のセンス鎖の優勢な活性は、アンチセンス鎖の活性に干渉し、こうして力価と最大効率の両方を制限する。この2本鎖のセンス鎖の化学的不活性化は、おそらくアンチセンス鎖との結合のためにリスク(RISC)を解放することにより、力価と最大効率(黒の点線)の両方を上昇させた。これは、センス鎖の不活性化が、より有効な標的遺伝子サイレンシングについてより強力なsiRNAを達成するための1つの機構であることを示唆する。

40

【0269】

7. 引用した文献

本明細書で引用したすべての文献は、各特許または特許出願が具体的かつ個別に参照することによりその全体が本明細書に組み込まれるかのように、同程度に参照することによ

50

りその全体が本明細書に組み込まれる。

【0270】

当業者に明らかなように、本発明の精神と範囲を逸脱することなく本発明の多くの修飾および変更態様が可能である。本明細書に記載の具体例は例示のためのみであり、本発明は、特許請求の範囲の条項とかかる特許請求の範囲の全範囲によってのみ限定される。

【図面の簡単な説明】

【0271】

【図1A】図1Aは、発現プロファイルにより証明されるsiRNA配列に特異的な遺伝子発現パターンを示す図である。

【図1B】図1Bは、発現プロファイルにより証明されるsiRNA配列に特異的な遺伝子発現パターンを示す図である。 10

【図2】図2はオフターゲット遺伝子サイレンシングの用量効果を示す図である。

【図3A】図3Aは、RNAiによるMAPK14タンパク質とRNAノックダウンの動力的分析を示す図である。

【図3B】図3Bは、RNAiによるMAPK14タンパク質とRNAノックダウンの動力的分析を示す図である。

【図3C】図3Cは、RNAiによるMAPK14タンパク質とRNAノックダウンの動力的分析を示す図である。

【図4A】図4Aは、オフターゲット遺伝子制御への配列類似性の寄与を示す図である。

【図4B】図4Bは、オフターゲット遺伝子制御への配列類似性の寄与を示す図である。 20

【図4C】図4Cは、オフターゲット遺伝子制御への配列類似性の寄与を示す図である。

【図5】図5は、MAPK14とIGF1Rサイレンシングは、明確な遺伝子発現パターンを示す図である。

【図6A】図6Aは、IGF1R siRNAが受容体発現と機能を妨害することを示す図である。

【図6B】図6Bは、IGF1R siRNAが受容体発現と機能を妨害することを示す図である。

【図6C】図6Cは、IGF1R siRNAが受容体発現と機能を妨害することを示す図である。

【図7】図7は、siRNAオリゴによるMAPK14 mRNAのノックダウンに対する単一ヌクレオチドミスマッチの作用を示す図である。

【図8】図8は、本発明の方法を実施するのに有用なコンピュータシステムの例を示す図である。 30

【図9】図9は、RNA干渉の略図である。

【図10】図10は、3つのsiRNAのsiRNAプールを使用して得られた結果を示す図である。

【図11】図11は、siRNAプールがサイレンシング特異性を上昇させることを示す図である。

【図12】図12は、9つのsiRNAのsiRNAプールを使用して得られた結果を示す図である。

【図13A】図13Aは、siRNAとその逆相補体のGC含量と悪いsiRNA (bad siRNA) のGC含量との比較を示し、悪いsiRNAが良いsiRNA (good siRNA) と同様のセンス鎖を有し、良いsiRNAが悪いsiRNAと同様のセンス鎖を有することを示す図である。 40

【図13B】図13Bは、siRNAとその逆相補体のGC含量と悪いsiRNA (bad siRNA) のGC含量との比較を示し、悪いsiRNAが良いsiRNA (good siRNA) と同様のセンス鎖を有し、良いsiRNAが悪いsiRNAと同様のセンス鎖を有することを示す図である。

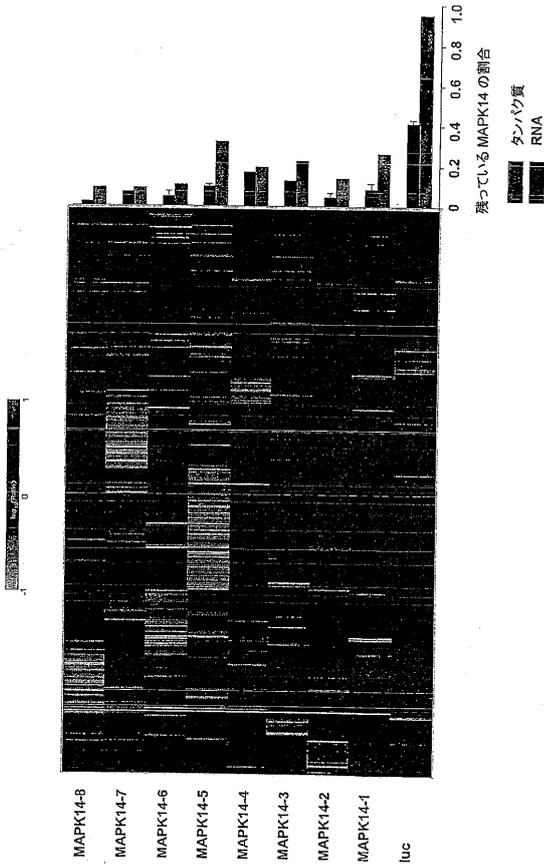
【図14】図14は、あまり有効でないsiRNAが活性なセンス鎖を有することを示す図である。

【図15】図15は、センス鎖修飾が、「センス活性」siRNAの特異性を変化させることを示す図である。

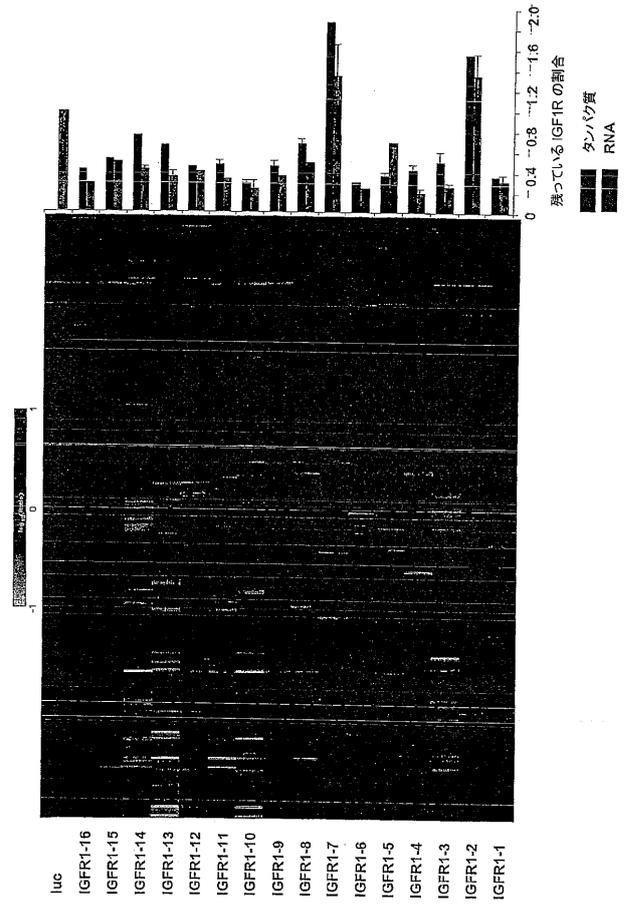
【図16】図16はオフターゲット制御における鎖の偏りを示す図である。

【図17】図17は、センス鎖修飾が「センス活性」siRNAの力価を上昇させることを示す図である。 50

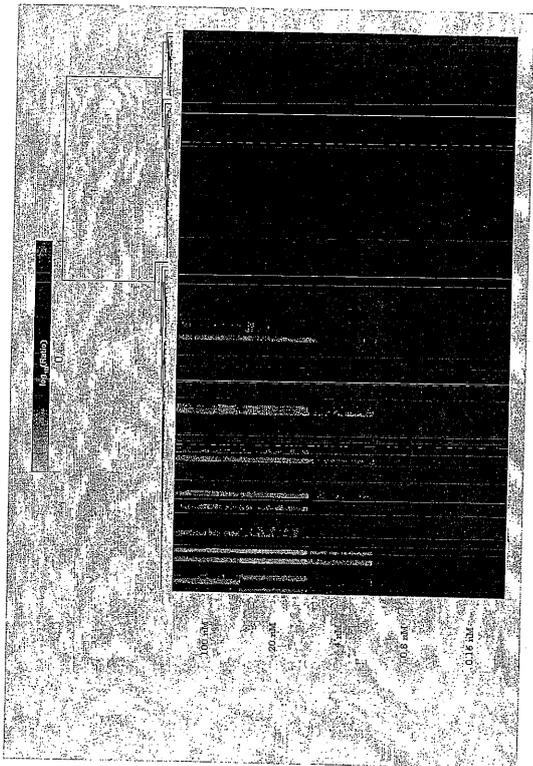
【 図 1 A 】



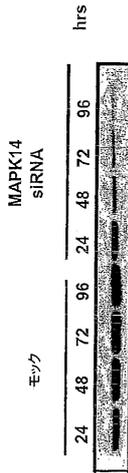
【 図 1 B 】



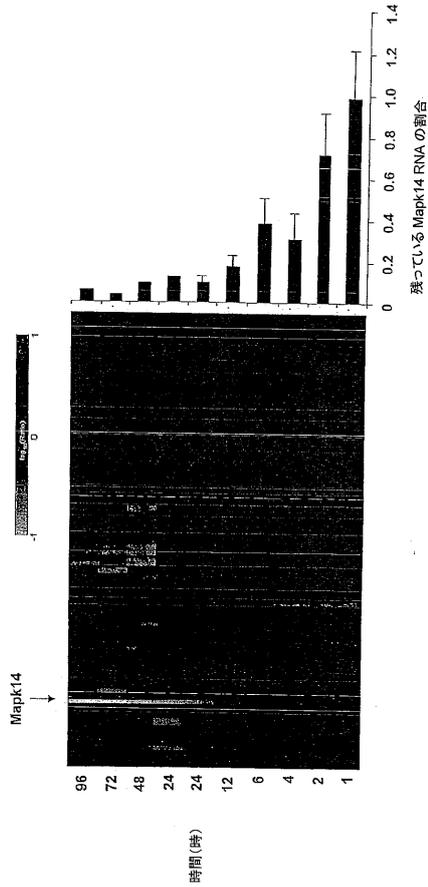
【 図 2 】



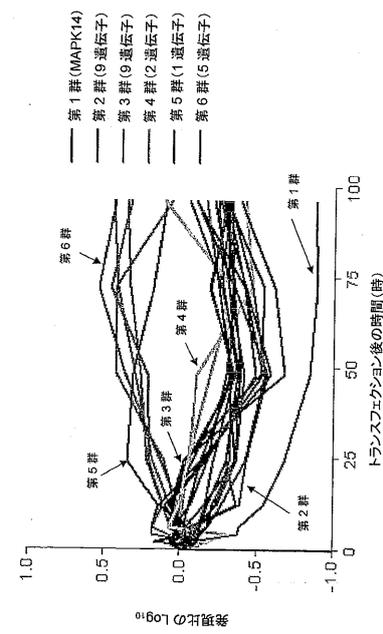
【 図 3 A 】



【 図 3 B 】



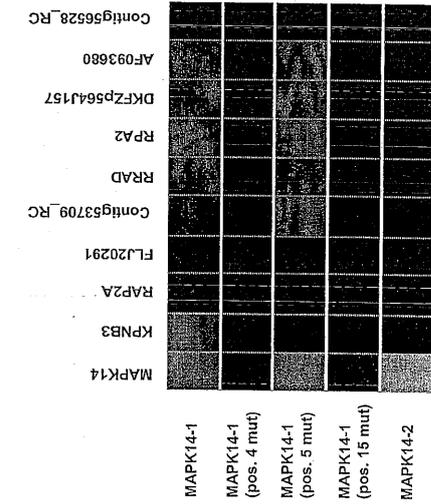
【 図 3 C 】



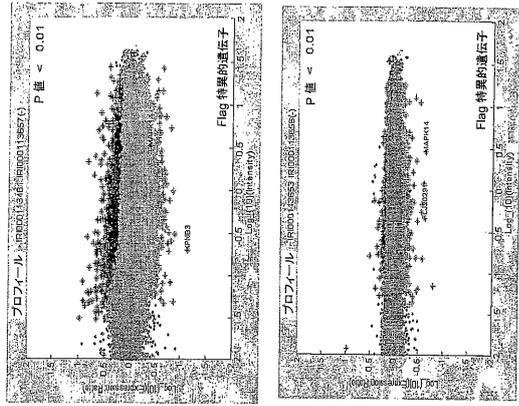
【 図 4 A 】

| 配列 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | |
|---|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|--|
| 遺伝子名 | | | | | | | | | | | | | | | | | | | | |
| NM_001315 (SEQ ID NO:29) MAPK14 | C | C | T | A | C | A | G | A | G | A | A | C | T | G | C | G | G | T | T | |
| NM_002271 (SEQ ID NO:30) KPNB3 | G | C | T | A | C | A | G | A | G | A | A | C | T | G | C | A | T | C | T | |
| Hs13_10109_29_13_1603 RPA2 (SEQ ID NO:31) | T | T | T | A | C | A | G | A | G | A | A | C | T | G | C | A | G | C | C | |
| NM_017748 (SEQ ID NO:32) FLJ20291 | T | T | T | A | C | A | G | A | G | A | A | C | T | G | C | G | G | T | A | |
| A133672 (SEQ ID NO:33) Contig53709_RC | C | C | T | C | A | A | A | G | A | A | C | C | T | G | C | G | G | T | T | |
| NM_004165 (SEQ ID NO:34) RRAD | A | G | G | C | T | C | A | G | A | A | C | T | G | C | G | G | G | T | T | |
| NM_002946 (SEQ ID NO:35) RPA2 | G | A | G | C | A | G | A | G | A | A | C | T | T | G | G | T | T | G | T | |
| NM_018457 (SEQ ID NO:36) DKFZp564J157 | A | T | G | A | G | C | T | T | T | G | A | C | T | G | C | G | G | T | T | |
| NM_013242 (SEQ ID NO:37) AF093660 | A | A | T | A | T | T | T | C | T | C | C | T | G | C | G | G | T | T | T | |
| AW237459 (SEQ ID NO:38) Contig56528_RC | A | G | G | A | G | A | A | T | G | A | C | T | G | C | G | G | T | A | A | |

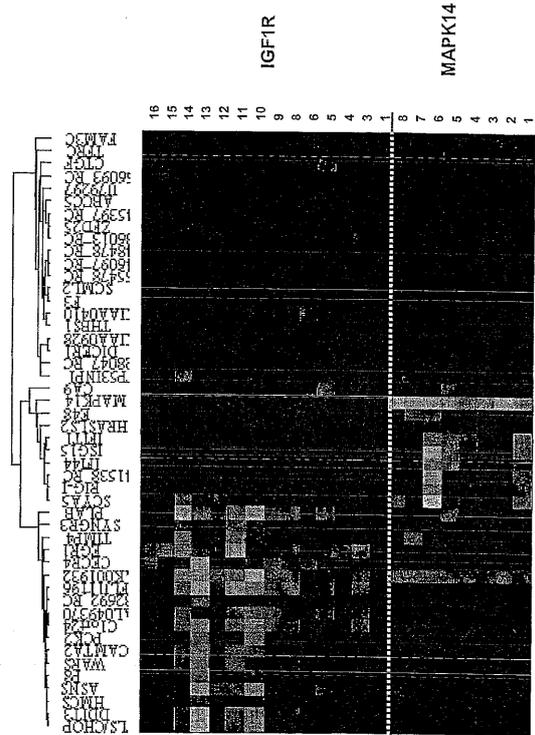
【 図 4 B 】



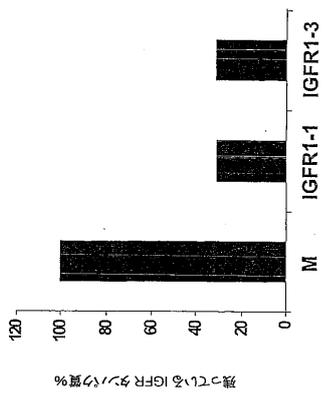
【 図 4 C 】



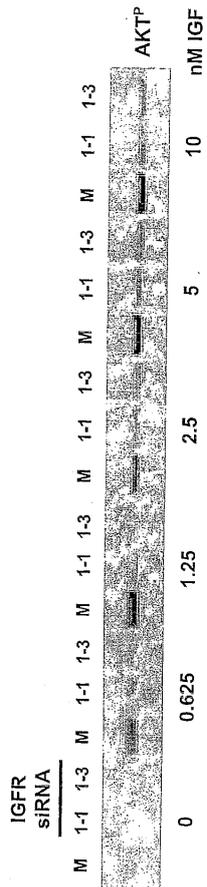
【 図 5 】



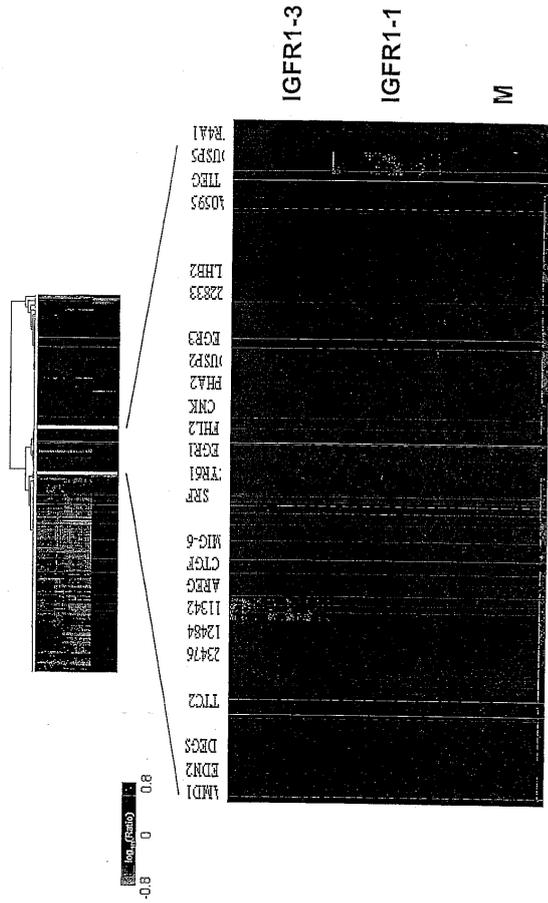
【 図 6 A 】



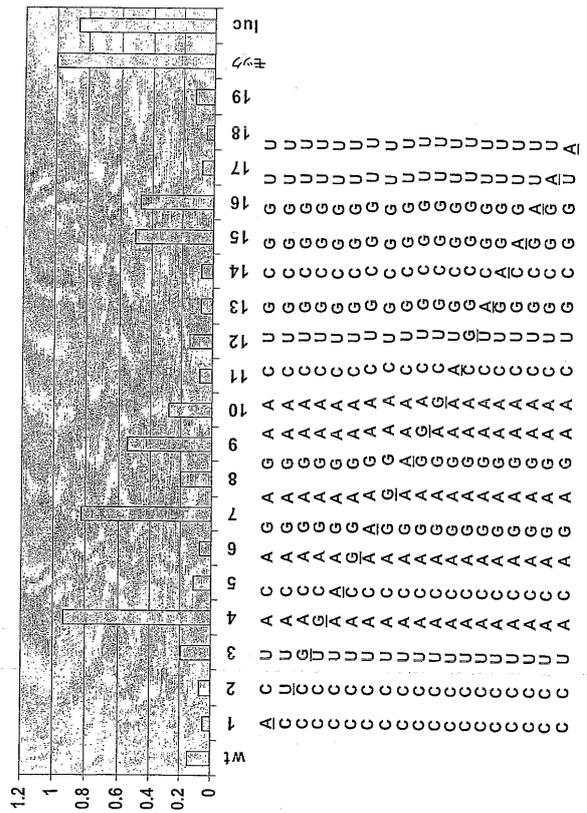
【 図 6 B 】



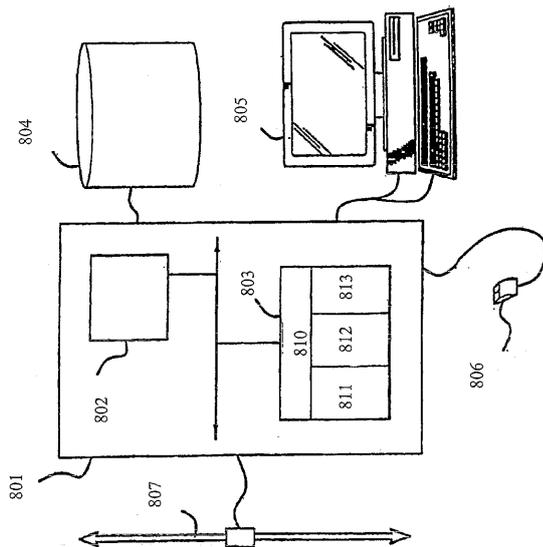
【 図 6 C 】



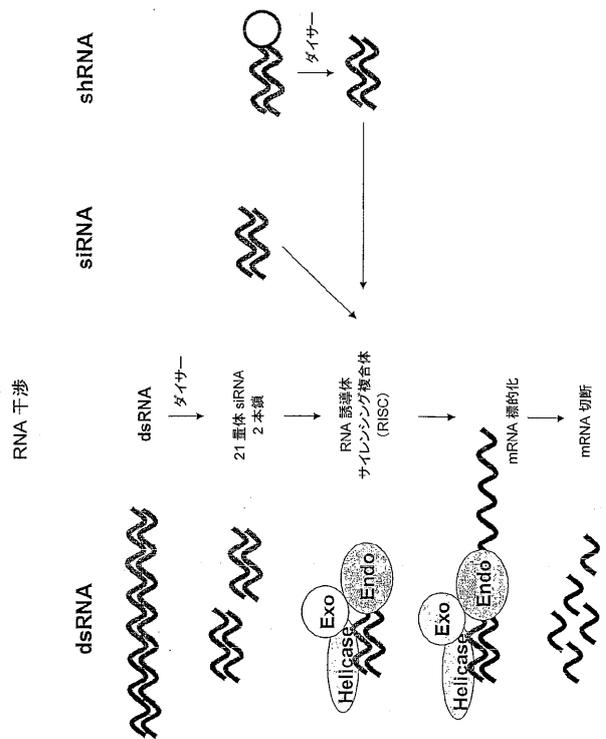
【 図 7 】



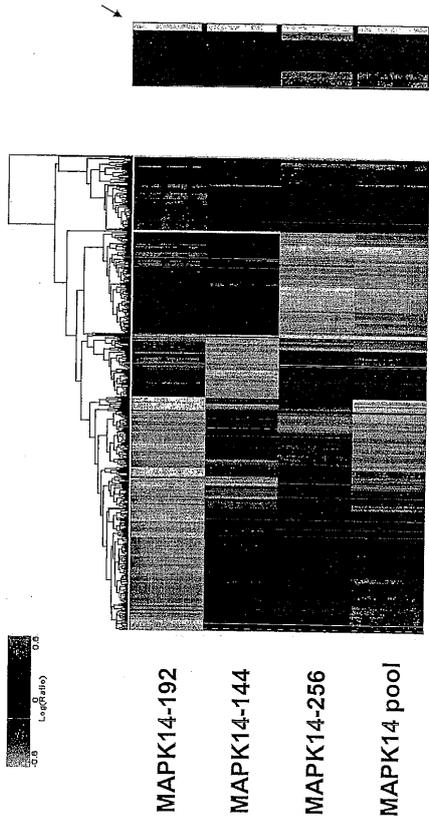
【 図 8 】



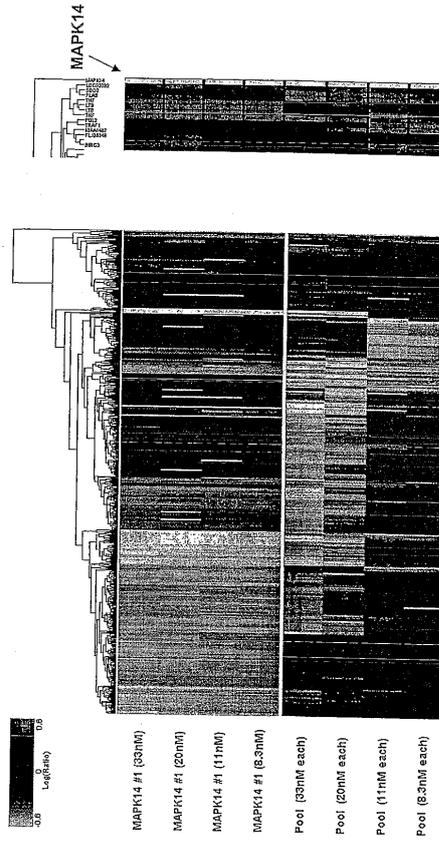
【 図 9 】



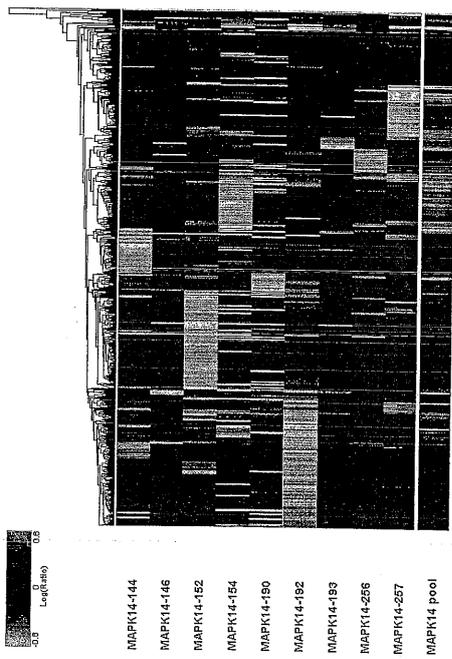
【 図 1 0 】



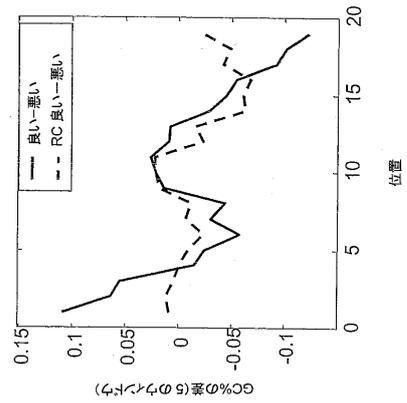
【 図 1 1 】



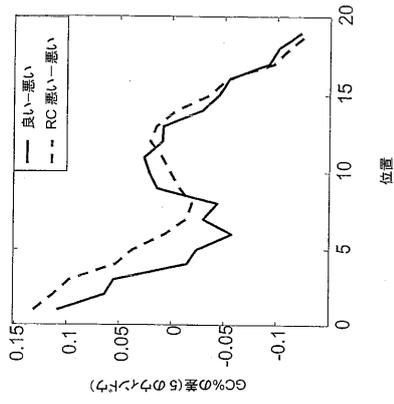
【 図 1 2 】



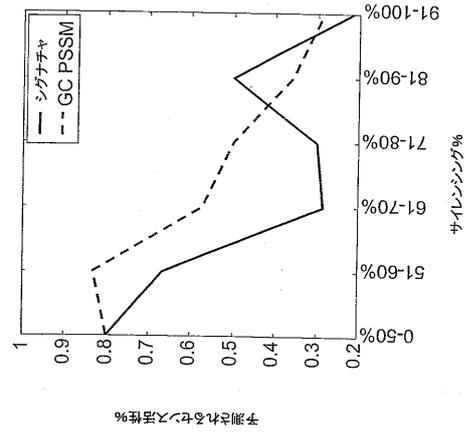
【 図 1 3 A 】



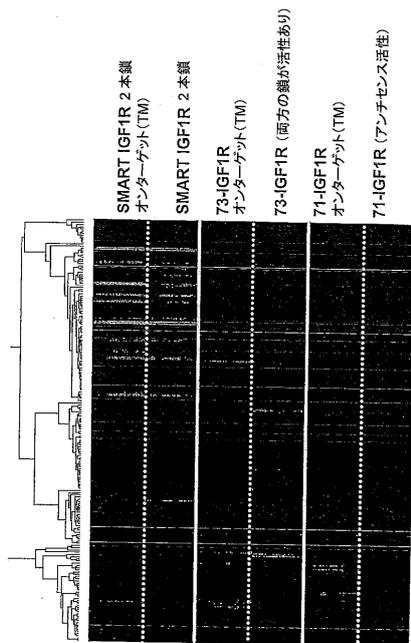
【図 13 B】



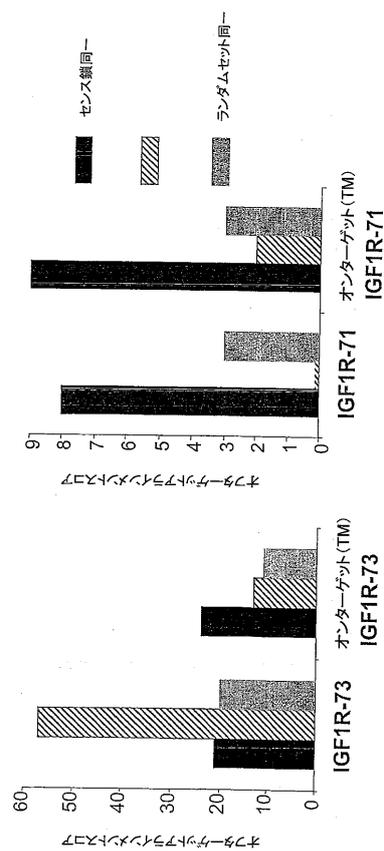
【図 14】



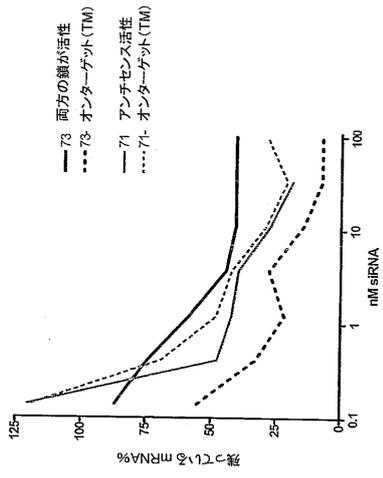
【図 15】



【図 16】



【 図 17 】



【 国際調査報告 】

| INTERNATIONAL SEARCH REPORT | | International application No. PCT/US04/15439 |
|--|---|--|
| A. CLASSIFICATION OF SUBJECT MATTER IPC(7) : C07H 21/04; A61K 48/00 US CL : 536/24.5; 514/44 According to International Patent Classification (IPC) or to both national classification and IPC | | |
| B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) U.S. : 536/24.5; 514/44 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) | | |
| C. DOCUMENTS CONSIDERED TO BE RELEVANT | | |
| Category * | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
| X | WO 02/44321 A2 (TUSCHL et al) 06 June 2002 (06.06.2002), Figure 18. | 1-11 |
| T | GREEN, D.W. et al. Antisense Oligonucleotides: An evolving technology for the modulation of gene expression in human disease. J. Am. Coll. Surg. 2000, Vol. 191, pages 93-105, pages 93-105, see entire document. | 1-11 |
| T | JEN, K.Y. et al. Suppression of gene expression by targeted disruption of messenger RNA: available options and current strategies. Stem Cells 2000, Vol. 18, pages 307-318, see entire document. | 1-11 |
| T | CAPLEN, N.J. et al. RNAi as a gene therapy approach. Gene Therapy 2003, Vol. 3, No. 4, pages 575-586, see entire document. | 1-11 |
| T | PAROO, Z. et al. Challenges for RNAi in vivo. Trends in Biotechnology. 2004, Vol. 22; pages 309-394, see entire document. | 1-11 |
| T | NOVINA, C.D. et al. The RNAi revolution. Nature. 2004, Vol. 430; pages 161-164, see entire document. | 1-11 |
| <input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex. | | |
| * Special categories of cited documents: | | |
| "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed | "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. "&" document member of the same patent family | |
| Date of the actual completion of the international search 07 July 2005 (07.07.2005) | | Date of mailing of the international search report 14 OCT 2005 |
| Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450 Alexandria, Virginia 22313-1450 Facsimile No. (703) 305-3230 | | Authorized officer  Kimberly Chong Telephone No. 571-272-0564 |

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US04/15439

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)This International Searching Authority found multiple inventions in this international application, as follows:
Please See Continuation Sheet

1. As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1-11
- Remark on Protest The additional search fees were accompanied by the applicant's protest.
 No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORTInternational application No.
PCT/US04/15439**BOX III. OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING**

Group 1, claims 1-11, drawn to a method of silencing a target gene in a eukaryotic cell by RNA interference comprising administering a siRNA.

Group 2, claims 12-29, drawn to a method of silencing a plurality of different target genes in a eukaryotic cell by RNA interference comprising administering a siRNA.

Group 3, claims 30-39, 142-157, drawn to a method of designing a siRNA for silencing a first gene but not a second gene in a eukaryotic cell by RNA interference.

Group 4, claims 40-62, drawn to a method of determining an effect of an siRNA on an eukaryotic cell.

Group 5, claims 63-72, drawn to a method of identifying one or more gene in an eukaryotic cell which are silenced directly by a plurality of different siRNA.

Group 6, claims 73-91, drawn to a method of identifying one or more candidate genes in a eukaryotic cell.

Group 7, claims 92-119, drawn to a method of identifying one or more genes that are regulated by a first protein but not a second protein in an eukaryotic cell.

Group 8, claims 120-128, drawn to a method of identifying genes that are differentially silenced by a first and a second siRNA molecule.

Group 9, claims 129-135, drawn to a method of treating a disease or undesirable condition in a mammal comprising administering a siRNA.

Group 10, claims 136-141, drawn to a method of selecting one or more siRNA for silencing a gene in a eukaryotic cell.

Group 11, claims 158-172, drawn to a method of evaluating relative gene silencing activity of the first and second strands of a siRNA.

According to PCT Rule 13.2, unity of invention exists only when the shared same or corresponding technical feature is a contribution over the prior art. The inventions listed as Groups 1-11 do not relate to a single general inventive concept because they lack the same or corresponding special technical feature.

The technical feature of Group 1 is a small interfering RNA molecule targeted to a eukaryotic cell, wherein the siRNA comprises a sense strand and an antisense strand that is identical to a sequence of a transcript of said target gene.

Caplen et al. (August 2001, PNAS Vol. 98: 9742-9747) discloses a small interfering RNA molecule targeted to a eukaryotic cell, wherein the siRNA comprises a sense strand and an antisense strand that is identical to a sequence of a transcript of said target gene (see entire

INTERNATIONAL SEARCH REPORTInternational application No.
PCT/US04/15439

document, especially page 9744). The special technical feature of Group 1 is shown by Caplen et al. to lack novelty or inventive step and does not make a contribution over the prior art.

Therefore, the special technical feature of group 1 does not constitute a special technical feature as defined by PCT Rule 13.2, as it does not define a contribution of the prior art.

Thus, no special technical feature is shared between groups 1-12 and unity of invention is lacking.

フロントページの続き

| | | | |
|----------------------------------|--|-----------------|------------|
| (51) Int.Cl. | | F I | テーマコード(参考) |
| A 6 1 K 48/00 (2006.01) | | A 6 1 K 48/00 | |
| A 6 1 K 31/7088 (2006.01) | | A 6 1 K 31/7088 | |

(81) 指定国 AP(BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), EP(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, 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, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW

(特許庁注：以下のものは登録商標)

フロッピー

マッキントッシュ

UNIX

Linux

JAVA

- (72) 発明者 ジャクソン, エイミー, エル.
 アメリカ合衆国 9 8 1 2 5 ワシントン州, シアトル, ルーズベルト ウェイ エヌイー 1 2
 0 4 6 - エー
- (72) 発明者 パーツ, スティーヴン, アール.
 アメリカ合衆国 9 8 1 0 7 ワシントン州, シアトル, エヌダブリュ 4 9 ティーエイチ スト
 リート 1 0 9
- (72) 発明者 バーチャード, ジュリア
 アメリカ合衆国 9 8 2 7 0 ワシントン州, メリーズヴィル, 7 4 ティーエイチ ドライブ エ
 ヌイー 8 4 1 2
- (72) 発明者 シェルター, ジャネル, エム.
 アメリカ合衆国 9 8 0 0 8 ワシントン州, ベルビュー, 1 6 6 ティーエイチ アベニュー エ
 スイー 2 0 1 7
- (72) 発明者 リンズリー, ピーター, エス.
 アメリカ合衆国 9 8 1 1 9 ワシントン州, シアトル, 1 1 ティーエイチ アベニュー ダブリ
 ユ 2 5 2 8

F ターム(参考) 4B024 AA01 CA11 DA01 DA02 DA11 EA10 GA11 HA17
 4B063 QA01 QA05 QA08 QA18 QQ07 QQ08 QQ09 QQ52 QQ79 QR08
 QR32 QR56 QR62 QS25 QS34 QX02
 4B065 AA57X AA57Y AA87X AA87Y AB01 AC14 BA01 CA44 CA46 CA60
 4C084 AA13 NA14
 4C086 AA01 AA02 EA16 MA01 NA14