

(19)日本国特許庁(JP)

## (12)特許公報(B2)

(11)特許番号  
特許第7661009号  
(P7661009)

(45)発行日 令和7年4月14日(2025.4.14)

(24)登録日 令和7年4月4日(2025.4.4)

(51)国際特許分類

C 1 2 N	15/54 (2006.01)	C 1 2 N	15/54
A 6 1 K	31/7105(2006.01)	A 6 1 K	31/7105
A 6 1 K	45/00 (2006.01)	A 6 1 K	45/00
A 6 1 K	48/00 (2006.01)	A 6 1 K	48/00
A 6 1 P	9/00 (2006.01)	A 6 1 P	9/00

F I

請求項の数 61 (全108頁) 最終頁に続く

(21)出願番号 特願2019-567349(P2019-567349)  
 (86)(22)出願日 平成30年6月4日(2018.6.4)  
 (65)公表番号 特表2020-527329(P2020-527329  
 A)  
 (43)公表日 令和2年9月10日(2020.9.10)  
 (86)国際出願番号 PCT/US2018/035806  
 (87)国際公開番号 WO2018/226560  
 (87)国際公開日 平成30年12月13日(2018.12.13)  
 審査請求日 令和3年6月3日(2021.6.3)  
 審判番号 不服2023-7632(P2023-7632/J1)  
 審判請求日 令和5年5月10日(2023.5.10)  
 (31)優先権主張番号 62/515,140  
 (32)優先日 平成29年6月5日(2017.6.5)  
 (33)優先権主張国・地域又は機関  
 米国(US)

最終頁に続く

(73)特許権者 597160510  
 リジェネロン・ファーマシューティカル  
 ズ・インコーポレイテッド  
 R E G E N E R O N P H A R M A C E  
 U T I C A L S , I N C .  
 アメリカ合衆国 1 0 5 9 1 - 6 7 0 7 2  
 ューヨーク州タリータウン、オールド・  
 ソー・ミル・リバー・ロード 7 7 7 番  
 (73)特許権者 517322042  
 ユニバーシティ オブ メリーランド ポ  
 ルチモア  
 アメリカ合衆国 メリーランド州 2 1 2  
 0 1 ポルチモア ウエスト レキシントン  
 ストリート 6 2 0 フォース フロア  
 (74)代理人 100105957

最終頁に続く

(54)【発明の名称】 B 4 G A L T 1 バリエント及びその使用

## (57)【特許請求の範囲】

## 【請求項 1】

低密度リポタンパク質コレステロール (LDL) の増加、総コレステロールの増加、フィブリノーゲンの増加、推算糸球体濾過量 (eGFR) の増加、またはアスパラギン酸トランスアミナーゼ (AST) の減少を含む心血管病態の発症に対するヒト対象の感受性を判定するために使用される単離された核酸分子であって、配列番号 1 と、少なくとも 90%、少なくとも 95%、少なくとも 98%、または少なくとも 99% 同一の核酸配列を含み、前記核酸配列が、配列番号 1 の 53575 ~ 53577 位に対応してセリンをコードするコドンを含むことを条件とする、前記単離された核酸分子、またはその相補体。

## 【請求項 2】

前記核酸配列が、配列番号 2 の 53575 ~ 53577 位に対応するヌクレオチドを含む、請求項 1 に記載の単離された核酸分子。

## 【請求項 3】

前記核酸配列が、B 4 G A L T 1 遺伝子のエクソン 1 ~ 6 を含む配列番号 2 の部分と、少なくとも 90%、少なくとも 95%、少なくとも 98%、または少なくとも 99% 同一である、請求項 1 または 2 に記載の単離された核酸分子。

## 【請求項 4】

前記核酸配列が、配列番号 2 を含む、請求項 1 または 2 に記載の単離された核酸分子。

## 【請求項 5】

請求項 1 ~ 4 のいずれか一項に記載の単離された核酸分子を含むベクター。

**【請求項 6】**

請求項 1 ~ 4 のいずれか一項に記載の単離された核酸分子もしくは請求項 5 に記載のベクターと、担体とを含む、組成物。

**【請求項 7】**

請求項 1 ~ 4 のいずれか一項に記載の単離された核酸分子もしくは請求項 5 に記載のベクターを含む、宿主細胞。

**【請求項 8】**

低密度リポタンパク質コレステロール (LDL) の増加、総コレステロールの増加、フィブリノーゲンの増加、推算糸球体濾過量 (eGFR) の増加、またはアスパラギン酸トランスアミナーゼ (AST) の減少を含む心血管病態の発症に対するヒト対象の感受性を判定するために使用される単離された核酸分子であって、配列番号 4 と、少なくとも 90%、少なくとも 95%、少なくとも 98%、または少なくとも 99% 同一の核酸配列を含み、前記核酸配列が、完全長 / 成熟型 B4GALT1 ポリペプチドの 352 位に対応する位置にセリンをコードするコドンを含むことを条件とする、前記単離された核酸分子、またはその相補体。

10

**【請求項 9】**

前記核酸配列が、B4GALT1 遺伝子のエクソン 1 ~ 6 を含む配列番号 4 の部分と、少なくとも 90%、少なくとも 95%、少なくとも 98%、または少なくとも 99% 同一である、請求項 8 に記載の単離された核酸分子。

**【請求項 10】**

20

前記核酸配列が、配列番号 4 を含む、請求項 8 または 9 に記載の単離された核酸分子。

**【請求項 11】**

請求項 8 ~ 10 のいずれか一項に記載の単離された核酸分子を含む、ベクター。

**【請求項 12】**

請求項 8 ~ 10 のいずれか一項に記載の単離された核酸分子もしくは請求項 11 に記載のベクターと、担体とを含む、組成物。

**【請求項 13】**

請求項 8 ~ 10 のいずれか一項に記載の単離された核酸分子もしくは請求項 11 に記載のベクターを含む、宿主細胞。

**【請求項 14】**

30

低密度リポタンパク質コレステロール (LDL) の増加、総コレステロールの増加、フィブリノーゲンの増加、推算糸球体濾過量 (eGFR) の増加、またはアスパラギン酸トランスアミナーゼ (AST) の減少を含む心血管病態の発症に対するヒト対象の感受性を判定するために使用される単離された核酸分子であって、配列番号 8 と、少なくとも 90%、少なくとも 95%、少なくとも 98%、または少なくとも 99% 同一であるポリペプチドをコードする核酸配列を含み、前記ポリペプチドが、352 位にセリンを含むことを条件とする、前記核酸分子、またはその相補体。

**【請求項 15】**

前記核酸配列が、配列番号 8 のポリペプチド配列をコードする、請求項 14 に記載の単離された核酸分子。

40

**【請求項 16】**

請求項 14 または 15 に記載の単離された核酸分子を含む、ベクター。

**【請求項 17】**

請求項 14 または 15 に記載の単離された核酸分子もしくは請求項 16 に記載のベクターと、担体とを含む、組成物。

**【請求項 18】**

請求項 14 または 15 に記載の単離された核酸分子もしくは請求項 16 に記載のベクターを含む、宿主細胞。

**【請求項 19】**

低密度リポタンパク質コレステロール (LDL) の増加、総コレステロールの増加、フィ

50

ブリノーゲンの増加、推算糸球体濾過量（eGFR）の増加、またはアスパラギン酸トランスマニナーゼ（AST）の減少を含む心血管病態の発症に対するヒト対象の感受性を判定するために使用されるcDNAであって、配列番号6と、少なくとも90%、少なくとも95%、少なくとも98%、または少なくとも99%同一の核酸配列を含み、ヒト-1,4-ガラクトシルトランスフェラーゼ1（B4GALT1）タンパク質をコードし、前記核酸配列が、完全長／成熟型B4GALT1ポリペプチドの352位に対応する位置でセリンをコードすることを条件とする、前記cDNA、またはその相補体。

【請求項20】

前記核酸配列が、配列番号6を含む、請求項19に記載のcDNA。

【請求項21】

請求項19または20に記載のcDNAを含む、ベクター。

10

【請求項22】

請求項19または20に記載のcDNAもしくは請求項21に記載のベクターと、担体とを含む、組成物。

【請求項23】

請求項19または20に記載のcDNAもしくは請求項21に記載のベクターを含む、宿主細胞。

【請求項24】

低密度リポタンパク質コレステロール（LDL）の増加、総コレステロールの増加、フィブリノーゲンの増加、推算糸球体濾過量（eGFR）の増加、またはアスパラギン酸トランスマニナーゼ（AST）の減少を含む心血管病態の発症に対するヒト対象の感受性を判定するために使用される単離されたポリペプチドであって、配列番号8を有するB4GALT1バリアントポリペプチドと、少なくとも90%、少なくとも95%、少なくとも98%、または少なくとも99%同一のアミノ酸配列を含み、前記ポリペプチドが、配列番号8の352位に対応するセリンを含むことを条件とする、前記単離されたポリペプチド。

20

【請求項25】

前記B4GALT1バリアントポリペプチドが、配列番号8を含む、請求項24に記載のポリペプチド。

【請求項26】

請求項24または請求項25に記載のポリペプチドと、担体または賦形剤とを含む、組成物。

30

【請求項27】

請求項24または請求項25に記載のポリペプチドを発現する、宿主細胞。

【請求項28】

前記ポリペプチドをコードする核酸分子を含む宿主細胞を培養し、それにより前記細胞に前記ポリペプチドを発現させ、前記発現させたポリペプチドを回収することを含む、請求項24または請求項25に記載のポリペプチドの製造方法。

【請求項29】

ヒト対象におけるB4GALT1バリアント核酸分子の検出方法であって、前記対象から採取した試料をアッセイして、前記試料中の核酸分子が、配列番号8のアミノ酸配列を有する完全長／成熟型B4GALT1ポリペプチドの352位に対応する位置にセリンをコードする核酸配列を含むかどうかを判定することを含み、

40

前記試料中に前記核酸分子が存在することは、前記対象が、低密度リポタンパク質コレステロール（LDL）の増加、総コレステロールの増加、フィブリノーゲンの増加、推算糸球体濾過量（eGFR）の増加、またはアスパラギン酸トランスマニナーゼ（AST）の減少、を発症しにくいことを示す、検出方法。

【請求項30】

前記アッセイが：

前記試料中の核酸分子のB4GALT1ゲノム配列の部分を配列決定することであって、配列決定される前記部分が、配列番号2の53575～53577位に対応する位置を

50

含む、前記配列決定すること、

前記試料中の核酸分子の B 4 G A L T 1 m R N A 配列の部分を配列決定することであつて、配列決定される前記部分が、配列番号 4 の 1 2 4 3 ~ 1 2 4 5 位に対応する位置を含む、前記配列決定すること、または

前記試料中の核酸分子の B 4 G A L T 1 c D N A 配列の部分を配列決定することであつて、配列決定される前記部分が、配列番号 6 の 1 0 5 4 ~ 1 0 5 6 位に対応する位置を含む、前記配列決定すること、を含む、請求項 2 9 に記載の方法。

【請求項 3 1】

前記アッセイが：

a ) 前記試料を： i ) 配列番号 2 の 5 3 5 7 5 ~ 5 3 5 7 7 位に対応する B 4 G A L T 1 ゲノム配列の位置に近接する B 4 G A L T 1 ゲノム配列の部分、 i i ) 配列番号 4 の 1 2 4 3 ~ 1 2 4 5 位に対応する B 4 G A L T 1 m R N A の位置に近接する B 4 G A L T 1 m R N A 配列の部分、または i i i ) 配列番号 6 の 1 0 5 4 ~ 1 0 5 6 位に対応する B 4 G A L T 1 c D N A の位置に近接する B 4 G A L T 1 c D N A 配列の部分にハイブリダイズするプライマーと接触させること、

b ) 少なくとも： i ) 5 3 5 7 5 ~ 5 3 5 7 7 位に対応する B 4 G A L T 1 ゲノム配列の位置、 i i ) 1 2 4 3 ~ 1 2 4 5 位に対応する B 4 G A L T 1 m R N A の位置、または i i i ) 1 0 5 4 ~ 1 0 5 6 位に対応する B 4 G A L T 1 c D N A の位置を介して前記プライマーを伸長すること、及び

c ) 前記プライマーの伸長産物が： 配列番号 8 の 3 5 2 位でセリンをコードするヌクレオチドを i ) B 4 G A L T 1 ゲノム配列の 5 3 5 7 5 ~ 5 3 5 7 7 位に対応する、 i i ) B 4 G A L T 1 m R N A の 1 2 4 3 ~ 1 2 4 5 位に対応する、または i i i ) B 4 G A L T 1 c D N A の 1 0 5 4 ~ 1 0 5 6 位に対応する位置に含むかどうかを判定することを含む、請求項 2 9 に記載の方法。

【請求項 3 2】

前記アッセイが、前記試料を、ストリンジエントな条件下で、 B 4 G A L T 1 バリアントゲノム配列、 m R N A 配列、または c D N A 配列に特異的にハイブリダイズし、対応する野生型 B 4 G A L T 1 配列にはハイブリダイズしないプライマーまたはプローブと接触させること、及びハイブリダイゼーションが生じたかどうかを判定することを含む、請求項 2 9 に記載の方法。

【請求項 3 3】

ヒト対象における配列番号 8 のアミノ酸配列を含む B 4 G A L T 1 A s n 3 5 2 S e r の存在の検出方法であつて、前記ヒト対象から採取した試料に対してアッセイを行い、前記試料中の B 4 G A L T 1 タンパク質が配列番号 8 の 3 5 2 位に対応する位置にセリン残基を含むかどうかを判定することを含み、

前記試料中に前記 B 4 G A L T 1 A s n 3 5 2 S e r が存在することは、前記対象が、低密度リポタンパク質コレステロール ( L D L ) の増加、総コレステロールの増加、フィブリノーゲンの増加、推算糸球体濾過量 ( e G F R ) の増加、またはアスパラギン酸トランスアミナーゼ ( A S T ) の減少、を発症しにくいことを示す、前記検出方法。

【請求項 3 4】

低密度リポタンパク質コレステロール ( L D L ) の増加、総コレステロールの増加、フィブリノーゲンの増加、推算糸球体濾過量 ( e G F R ) の増加、またはアスパラギン酸トランスアミナーゼ ( A S T ) の減少を含む心血管病態の発症に対するヒト対象の感受性をインピトロで判定する方法であつて：

a ) 前記対象から採取した試料をアッセイして、前記試料中の核酸分子が、配列番号 8 のアミノ酸配列を有する完全長 / 成熟型 B 4 G A L T 1 ポリペプチドの 3 5 2 位に対応する位置にセリンをコードする核酸配列を含むかどうかを判定すること、及び

b ) 前記核酸分子が前記完全長 / 成熟型 B 4 G A L T 1 ポリペプチドの 3 5 2 位に対応する位置にセリンをコードする核酸配列を含む場合、 L D L の増加、総コレステロールの増加、フィブリノーゲンの増加、 e G F R の増加、または A S T の減少を発症するリスク

10

20

30

40

50

が低いと前記ヒト対象を分類するか、または前記核酸分子が前記完全長／成熟型 B 4 G A L T 1 ポリペプチドの 352 位に対応する位置にセリンをコードする核酸配列を含まない場合、L D L の増加、総コレステロールの増加、フィブリノーゲンの増加、e G F R の増加、または A S T の減少を発症するリスクが高いと前記ヒト対象を分類することを含む、方法。

【請求項 35】

前記アッセイが：

前記試料中の核酸分子の B 4 G A L T 1 ゲノム配列の部分を配列決定することであって、配列決定される前記部分が、配列番号 2 の 53575 ~ 53577 位に対応する位置を含む、前記配列決定すること、

10

前記試料中の核酸分子の B 4 G A L T 1 m R N A 配列の部分を配列決定することであって、配列決定される前記部分が、配列番号 4 の 1243 ~ 1245 位に対応する位置を含む、前記配列決定すること、または

前記試料中の核酸分子の B 4 G A L T 1 c D N A 配列の部分を配列決定することであって、配列決定される前記部分が、配列番号 6 の 1054 ~ 1056 位に対応する位置を含む、前記配列決定すること、を含む、請求項 34 に記載の方法。

【請求項 36】

前記アッセイが：

a) 前記試料を： i) 配列番号 2 の 53575 ~ 53577 位に対応する B 4 G A L T 1 ゲノム配列の位置に近接する B 4 G A L T 1 ゲノム配列の部分、 ii) 配列番号 4 の 1243 ~ 1245 位に対応する B 4 G A L T 1 m R N A の位置に近接する B 4 G A L T 1 m R N A 配列の部分、または iii) 配列番号 6 の 1054 ~ 1056 位に対応する B 4 G A L T 1 c D N A の位置に近接する B 4 G A L T 1 c D N A 配列の部分にハイブリダイズするプライマーと接触させること、

20

b) 少なくとも： i) 53575 ~ 53577 位に対応する B 4 G A L T 1 ゲノム配列の位置、 ii) 1243 ~ 1245 位に対応する B 4 G A L T 1 m R N A の位置、または iii) 1054 ~ 1056 位に対応する B 4 G A L T 1 c D N A の位置を介して前記プライマーを伸長すること、及び

c) 前記プライマーの伸長産物が：配列番号 8 の 352 位でセリンをコードするヌクレオチドを i) B 4 G A L T 1 ゲノム配列の 53575 ~ 53577 位に対応する、 ii) B 4 G A L T 1 m R N A の 1243 ~ 1245 位に対応する、または iii) B 4 G A L T 1 c D N A の 1054 ~ 1056 位に対応する位置に含むかどうかを判定することを含む、請求項 34 に記載の方法。

30

【請求項 37】

前記アッセイが、前記試料を、ストリンジェントな条件下で、B 4 G A L T 1 バリアントゲノム配列、m R N A 配列、または c D N A 配列に特異的にハイブリダイズし、対応する野生型 B 4 G A L T 1 配列にはハイブリダイズしないプライマーまたはプローブと接触させること、及びハイブリダイゼーションが生じたかどうかを判定することを含む、請求項 34 に記載の方法。

【請求項 38】

低密度リポタンパク質コレステロール (L D L) の増加、総コレステロールの増加、フィブリノーゲンの増加、推算糸球体濾過量 (e G F R) の増加、またはアスパラギン酸トランスアミナーゼ (A S T) の減少を含む心血管病態の発症に対するヒト対象の感受性をインピトロで判定する方法であって：

40

a) 前記ヒト対象から採取した試料に対してアッセイを行い、前記試料中の B 4 G A L T 1 タンパク質が配列番号 8 の 352 位に対応する位置にセリン残基を含むかどうかを判定すること、及び

b) B 4 G A L T 1 ポリペプチドが完全長／成熟型 B 4 G A L T 1 ポリペプチドの配列番号 8 の 352 位に対応する位置にセリンを含む場合、L D L の増加、総コレステロールの増加、フィブリノーゲンの増加、e G F R の増加、または A S T の減少を発症するリス

50

クが低いと前記ヒト対象を分類するか、またはB 4 G A L T 1 ポリペプチドが完全長／成熟型B 4 G A L T 1 ポリペプチドの配列番号8の352位に対応する位置にセリンを含まない場合、LDLの増加、総コレステロールの増加、フィブリノーゲンの増加、eGFRの増加、またはASTの減少を発症するリスクが高いと前記ヒト対象を分類することを含む、方法。

【請求項39】

前記心血管病態が、LDL以外の1つ以上の血清脂質のレベルの上昇をさらに含み、前記方法は、

前記核酸分子が前記完全長／成熟型B 4 G A L T 1 ポリペプチドの配列番号8の352位に対応する位置にセリンをコードする核酸配列を含む場合、前記LDL以外の1つ以上の血清脂質のレベルの上昇を発症するリスクが低いと前記ヒト対象を分類するか、または前記核酸分子が前記完全長／成熟型B 4 G A L T 1 ポリペプチドの配列番号8の352位に対応する位置にセリンをコードする核酸配列を含まない場合、前記LDL以外の1つ以上の血清脂質のレベルの上昇を発症するリスクが高いと前記ヒト対象を分類することを含む、請求項34～37のいずれか一項に記載の方法。

10

【請求項40】

前記LDL以外の1つ以上の血清脂質が、コレステロール、HDL、トリグリセリド、HDLコレステロール、及び非HDLコレステロールの1つ以上を含む、請求項39に記載の方法。

【請求項41】

前記心血管病態が、冠動脈石灰化レベルの上昇をさらに含み、前記方法は、前記核酸分子が前記完全長／成熟型B 4 G A L T 1 ポリペプチドの配列番号8の352位に対応する位置にセリンをコードする核酸配列を含む場合、前記冠動脈石灰化レベルの上昇を発症するリスクが低いと前記ヒト対象を分類するか、または前記核酸分子が前記完全長／成熟型B 4 G A L T 1 ポリペプチドの配列番号8の352位に対応する位置にセリンをコードする核酸配列を含まない場合、前記冠動脈石灰化レベルの上昇を発症するリスクが高いと前記ヒト対象を分類することを含む、請求項34～37のいずれか一項に記載の方法。

20

【請求項42】

前記心血管病態が、心膜脂肪レベルの上昇をさらに含み、前記方法は、前記核酸分子が前記完全長／成熟型B 4 G A L T 1 ポリペプチドの配列番号8の352位に対応する位置にセリンをコードする核酸配列を含む場合、前記心膜脂肪レベルの上昇を発症するリスクが低いと前記ヒト対象を分類するか、または前記核酸分子が前記完全長／成熟型B 4 G A L T 1 ポリペプチドの配列番号8の352位に対応する位置にセリンをコードする核酸配列を含まない場合、前記心膜脂肪レベルの上昇を発症するリスクが高いと前記ヒト対象を分類することを含む、請求項34～37のいずれか一項に記載の方法。

30

【請求項43】

前記心血管病態が、アテローム血栓性病態をさらに含み、前記方法は、前記核酸分子が前記完全長／成熟型B 4 G A L T 1 ポリペプチドの配列番号8の352位に対応する位置にセリンをコードする核酸配列を含む場合、前記アテローム血栓性病態を発症するリスクが低いと前記ヒト対象を分類するか、または前記核酸分子が前記完全長／成熟型B 4 G A L T 1 ポリペプチドの配列番号8の352位に対応する位置にセリンをコードする核酸配列を含まない場合、前記アテローム血栓性病態を発症するリスクが高いと前記ヒト対象を分類することを含む、請求項34～37のいずれか一項に記載の方法。

40

【請求項44】

前記アテローム血栓性病態が、フィブリノーゲン活性の関与から形成される血餅をさらに含む、請求項43に記載の方法。

【請求項45】

前記心血管病態が、フィブリノーゲン活性の関与から形成される血餅をさらに含み、前記方法は、

50

前記核酸分子が前記完全長 / 成熟型 B 4 G A L T 1 ポリペプチドの配列番号 8 の 3 5 2 位に対応する位置にセリンをコードする核酸配列を含む場合、前記フィブリノーゲン活性の関与から形成される血餅を発症するリスクが低いと前記ヒト対象を分類するか、または前記核酸分子が前記完全長 / 成熟型 B 4 G A L T 1 ポリペプチドの配列番号 8 の 3 5 2 位に対応する位置にセリンをコードする核酸配列を含まない場合、前記フィブリノーゲン活性の関与から形成される血餅を発症するリスクが高いと前記ヒト対象を分類することを含む、請求項 3 4 ~ 3 7 のいずれか一項に記載の方法。

【請求項 4 6】

ガイド R N A および C a s 酵素を含む細胞において、前記ガイド R N A は、内在性 B 4 G A L T 1 遺伝子に結合または切断するように前記 C a s 酵素を誘導するのに有効であり、前記ガイド R N A が、配列番号 1 の 5 3 5 7 5 ~ 5 3 5 7 7 位に対応する位置を含むかまたは配列番号 1 の 5 3 5 7 5 ~ 5 3 5 7 7 位に対応する位置から約 1 0 0 0 ヌクレオチド以内にある前記内在性 B 4 G A L T 1 遺伝子内のガイド R N A 認識配列にハイブリダイズする D N A ターゲティングセグメントを含む、細胞。 10

【請求項 4 7】

前記ガイド R N A 認識配列が、配列番号 9 ~ 1 2 から選択される、請求項 4 6 に記載の細胞。

【請求項 4 8】

B 4 G A L T 1 バリアントの保因者ではなく、低密度リポタンパク質コレステロール ( L D L ) の増加、総コレステロールの増加、フィブリノーゲンの増加、推算糸球体濾過量 ( e G F R ) の増加、またはアスパラギン酸トランスアミナーゼ ( A S T ) の減少を含む心血管病態を発症しているか、または発症しやすい対象を治療するための医薬組成物であつて、前記医薬組成物は：

- a ) C a s タンパク質または前記 C a s タンパク質をコードする核酸、
- b ) ガイド R N A または前記ガイド R N A をコードする核酸であつて、前記ガイド R N A が、前記 C a s タンパク質と複合体を形成し、内在性 B 4 G A L T 1 遺伝子内のガイド R N A 認識配列にハイブリダイズし、前記ガイド R N A 認識配列が、配列番号 1 の 5 3 5 7 5 ~ 5 3 5 7 7 位に対応する位置を含むかまたはそれに近接する、前記ガイド R N A または前記ガイド R N A をコードする前記核酸、及び
- c ) 配列番号 1 の 5 3 5 7 5 ~ 5 3 5 7 7 位に対応する位置の 5 ' の標的配列にハイブリダイズする 5 ' 相同性アーム、配列番号 1 の 5 3 5 7 5 ~ 5 3 5 7 7 位に対応する位置の 3 ' の標的配列にハイブリダイズする 3 ' 相同性アーム、ならびに前記 5 ' 相同性アーム及び前記 3 ' 相同性アームに隣接し、配列番号 2 の 5 3 5 7 5 ~ 5 3 5 7 7 位に対応する位置にセリンをコードするヌクレオチド配列を含む核酸インサートを含む外来性ドナー配列、を含み、

前記医薬組成物は、前記対象に導入されるものである、医薬組成物。

【請求項 4 9】

B 4 G A L T 1 バリアントの保因者ではなく、低密度リポタンパク質コレステロール ( L D L ) の増加、総コレステロールの増加、フィブリノーゲンの増加、推算糸球体濾過量 ( e G F R ) の増加、またはアスパラギン酸トランスアミナーゼ ( A S T ) の減少を含む心血管病態を発症しているか、または発症しやすい対象を治療するための医薬組成物であつて、前記医薬組成物は：

- a ) C a s タンパク質または前記 C a s タンパク質をコードする核酸、
- b ) 第 1 のガイド R N A または前記第 1 のガイド R N A をコードする核酸であつて、前記第 1 のガイド R N A が、前記 C a s タンパク質と複合体を形成し、内在性 B 4 G A L T 1 遺伝子内の第 1 のガイド R N A 認識配列にハイブリダイズし、前記第 1 のガイド R N A 認識配列が、内在性 B 4 G A L T 1 遺伝子の開始コドンを含むか、または開始コドンから約 1 , 0 0 0 ヌクレオチド以内にあるか、または配列番号 9 ~ 1 2 から選択される、前記第 1 のガイド R N A または前記第 1 のガイド R N A をコードする前記核酸、及び
- c ) 配列番号 2 の 5 3 5 7 5 ~ 5 3 5 7 7 位に対応する位置にセリンをコードするヌク

10

20

30

40

50

レオチド配列を含む組換え B 4 G A L T 1 遺伝子を含む発現ベクター、を含み、前記医薬組成物は、前記対象に導入されるものである、医薬組成物。

【請求項 5 0】

B 4 G A L T 1 バリアントの保因者ではなく、低密度リポタンパク質コレステロール (L D L ) の増加、総コレステロールの増加、フィブリノーゲンの増加、推算糸球体濾過量 (e G F R ) の増加、またはアスパラギン酸トランスアミナーゼ (A S T ) の減少を含む心血管病態を発症しているか、または発症しやすい対象を治療するための医薬組成物であって、前記医薬組成物は、前記対象の細胞の内在性 B 4 G A L T 1 遺伝子内の配列にハイブリダイズして B 4 G A L T 1 ポリペプチドの発現を低下させるアンチセンス R N A 、s i R N A 、または s h R N A を含む、医薬組成物。

10

【請求項 5 1】

前記心血管病態が、L D L 以外の 1 つ以上の血清脂質のレベルの上昇を含む、請求項 5 0 に記載の医薬組成物。

【請求項 5 2】

前記 L D L 以外の 1 つ以上の血清脂質が、コレステロール、H D L 、トリグリセリド、H D L コレステロール、及び非 H D L コレステロールの 1 つ以上を含む、請求項 5 1 に記載の医薬組成物。

20

【請求項 5 3】

前記心血管病態が、冠動脈石灰化レベルの上昇をさらに含む、請求項 5 0 に記載の医薬組成物。

【請求項 5 4】

前記心血管病態が、心膜脂肪レベルの上昇をさらに含む、請求項 5 0 に記載の医薬組成物。

【請求項 5 5】

前記心血管病態が、アテローム血栓性病態をさらに含む、請求項 5 0 に記載の医薬組成物。

【請求項 5 6】

前記アテローム血栓性病態が、フィブリノーゲン活性の関与から形成される血餅を含む、請求項 5 5 に記載の医薬組成物。

30

【請求項 5 7】

前記心血管病態が、フィブリノーゲン活性の関与から形成される血餅を含む、請求項 5 0 に記載の医薬組成物。

【請求項 5 8】

B 4 G A L T 1 バリアントの保因者ではなく、低密度リポタンパク質コレステロール (L D L ) の増加、総コレステロールの増加、フィブリノーゲンの増加、推算糸球体濾過量 (e G F R ) の増加、またはアスパラギン酸トランスアミナーゼ (A S T ) の減少を含む心血管病態を発症しているか、または発症しやすい対象を治療するための医薬組成物であって、前記医薬組成物は、配列番号 2 の 5 3 5 7 5 ~ 5 3 5 7 7 位に対応する位置にセリンをコードするヌクレオチド配列を含む組換え B 4 G A L T 1 遺伝子を含む発現ベクターを含み、前記医薬組成物は、前記対象に導入されるものである、医薬組成物。

40

【請求項 5 9】

B 4 G A L T 1 バリアントの保因者ではなく、低密度リポタンパク質コレステロール (L D L ) の増加、総コレステロールの増加、フィブリノーゲンの増加、推算糸球体濾過量 (e G F R ) の増加、またはアスパラギン酸トランスアミナーゼ (A S T ) の減少を含む心血管病態を発症しているか、または発症しやすい対象を治療するための医薬組成物であって、前記医薬組成物は、配列番号 8 と、少なくとも 9 0 % 、少なくとも 9 5 % 、少なくとも 9 8 % または少なくとも 9 9 % 同一である B 4 G A L T 1 ポリペプチドであって、前記ポリペプチドが、配列番号 8 の 3 5 2 位に対応するセリンを含むことを条件とする同ポリペプチドをコードする核酸を含む発現ベクターを含み、前記医薬組成物は、前記対象に導入されるものである、医薬組成物。

50

## 【請求項 6 0】

B 4 G A L T 1 バリアントの保因者ではなく、低密度リポタンパク質コレステロール (L D L ) の増加、総コレステロールの増加、フィブリノーゲンの増加、推算糸球体濾過量 (e G F R ) の増加、またはアスパラギン酸トランスアミナーゼ (A S T ) の減少を含む心血管病態を発症しているか、または発症しやすい対象を治療するための医薬組成物であって、前記医薬組成物は、配列番号 8 と、少なくとも 9 0 %、少なくとも 9 5 %、少なくとも 9 8 % または少なくとも 9 9 % 同一である B 4 G A L T 1 ポリペプチドであって、前記ポリペプチドが、配列番号 8 の 3 5 2 位に対応するセリンを含むことを条件とする同ポリペプチドをコードする m R N A を含み、前記医薬組成物は、前記対象に導入されるものである、医薬組成物。

10

## 【請求項 6 1】

B 4 G A L T 1 バリアントの保因者ではなく、低密度リポタンパク質コレステロール (L D L ) の増加、総コレステロールの増加、フィブリノーゲンの増加、推算糸球体濾過量 (e G F R ) の増加、またはアスパラギン酸トランスアミナーゼ (A S T ) の減少を含む心血管病態を発症しているか、または発症しやすい対象を治療するための医薬組成物であって、前記医薬組成物は B 4 G A L T 1 A s n 3 5 2 S e r タンパク質またはその断片を含み、前記 B 4 G A L T 1 ポリペプチドが、配列番号 8 と、少なくとも 9 0 %、少なくとも 9 5 %、少なくとも 9 8 % または少なくとも 9 9 % 同一であり、かつ前記ポリペプチドが配列番号 8 の 3 5 2 位に対応するセリンを含むことを条件とし、前記医薬組成物は、前記対象に導入されるものである、医薬組成物。

20

## 【発明の詳細な説明】

## 【技術分野】

## 【0 0 0 1】

政府補助金への言及

本発明は、国立衛生研究所により授与された H L 1 2 1 0 0 7 に基づく政府の支援によりなされた。政府は本発明に一定の権利を有する。

## 【0 0 0 2】

配列表の参照

本出願には、ファイルサイズ 1 6 1 K B を有する 1 8 9 2 3 8 0 0 2 0 2 S E Q という名称のテキストファイルとして、2 0 1 8 年 6 月 4 日に作成し、電子的に提出した配列表が含まれる。配列表は、参照により本明細書に援用される。

30

## 【0 0 0 3】

本開示は、バリアント型 B 4 G A L T 1 ゲノム、m R N A 、及び c D N A 核酸分子、ならびにポリペプチド、これらの分子の存在の検出方法、内在性 B 4 G A L T 1 ゲノム、m R N A 、及び c D N A 核酸分子、ならびにポリペプチドの調節方法、バリアント型 B 4 G A L T 1 ゲノム、m R N A 、及び c D N A 核酸分子、ならびにポリペプチドの存在または非存在の検出による心血管病態の発症リスクの確認方法、ならびに心血管病態の治療方法を提供する。

## 【背景技術】

## 【0 0 0 4】

特許、公開出願、登録番号、技術論文、及び学術論文を含む様々な刊行物を、本明細書全体において引用する。引用する各刊行物は、あらゆる目的のためにその全体を参照により本明細書に援用する。

40

## 【0 0 0 5】

- 1 , 4 - ガラクトシルトランスフェラーゼ 1 ( B 4 G A L T 1 ) は、様々な複合糖質及び糖構造の生合成に役割を果たす I I 型膜結合糖タンパク質をコードする - 1 , 4 - ガラクトシルトランスフェラーゼ遺伝子ファミリーのメンバーである。 B 4 G A L T 1 によってコードされる酵素は、糖タンパク質の N 結合オリゴ糖部分のプロセシングに重要な役割を果たし、タンパク質結合糖鎖は、多くの場合、糖タンパク質の生物学的機能を調節する。したがって、 B 4 G A L T 1 活性の欠損は、 N 結合型オリゴ糖を含むすべての糖

50

タンパク質の構造を変化させる可能性がある。ロング型の B 4 G A L T 1 酵素はトランスゴルジ網に局在しており、N 結合オリゴ糖の高マンノース型から複合型への生合成処理の過程で、ガラクトシル残基を N - アセチルグルコサミン残基へ転移させる。ガラクトシル残基の付加はシアル酸付加の前提条件であるため、B 4 G A L T 1 の欠損は間接的な影響を及ぼしてシアル酸残基の付加をブロックし、したがって血漿糖タンパク質の半減期を変化させ得る。グリコシリ化の欠損により、LDL 受容体を含む様々な糖タンパク質の細胞内輸送が損なわれることが報告されている。さらに、N 結合型オリゴ糖の構造異常は、タンパク質の折りたたみを変化させる可能性があり、それが糖タンパク質の機能及びその分泌を変化させ得る。タンパク質の大部分は N 結合型グリコシリ化を有しており、例えば、細胞表面受容体（例えば、LDL 受容体及びインスリン受容体）及び様々な循環血漿タンパク質（例えば、アポリポタンパク質 B 及びフィブリノーゲン）が挙げられる。B 4 G A L T 1 遺伝子のタンパク質短縮型変異のホモ接合性に起因する遺伝性疾患の患者の報告が存在する。そのような患者の一人は、a) 重度の神経発達異常（水頭症を含む）、b) ミオパチー、及び c) 血液凝固異常を特徴とする重度の表現型を有していた。予測通り、循環トランスフェリンに由来するオリゴ糖は、ガラクトース及びシアル酸残基を欠いていた。同じ遺伝的欠損を有する 2 名の追加の患者は、凝固障害、肝障害、及び異形性の形質を特徴とする、より軽い表現型を示した。

#### 【 0 0 0 6 】

米国及びその他の西欧諸国において、心血管疾患は主要な死因である。脳卒中及び心筋梗塞などのアテローム血栓性心血管疾患の主なリスク因子として、血中コレステロールの増加及び血栓傾向が挙げられる。脂質代謝及び凝固に関する多くのタンパク質はグリコシリ化されており、したがって B 4 G A L T 1 による調節を受ける。心血管病態の発症と進行の根底にある遺伝的因子を知ることにより、リスク層別化を向上させ、新規治療戦略の基盤を提供することができるだろう。

#### 【発明の概要】

#### 【課題を解決するための手段】

#### 【 0 0 0 7 】

本開示は、B 4 G A L T 1 バリアント型ゲノム配列（r s 5 5 1 5 6 4 6 8 3 と呼ばれる SNP を含む）と少なくとも約 90 % 同一の核酸配列を含み、また、その核酸配列が完全長 / 成熟型 B 4 G A L T 1 ポリペプチドの 352 位に対応する位置にセリンをコードするスクレオチドを含むことを条件とする核酸分子を提供する。

#### 【 0 0 0 8 】

本開示はまた、B 4 G A L T 1 バリアント型 mRNA 配列（r s 5 5 1 5 6 4 6 8 3 と呼ばれる SNP を含む）と少なくとも約 90 % 同一の核酸配列を含み、また、その核酸配列が完全長 / 成熟型 B 4 G A L T 1 ポリペプチドの 352 位に対応する位置でセリンをコードすることを条件とする核酸分子を提供する。

#### 【 0 0 0 9 】

本開示はまた、B 4 G A L T 1 バリアント型 cDNA 配列（r s 5 5 1 5 6 4 6 8 3 と呼ばれる SNP を含む）と少なくとも約 90 % 同一の核酸配列を含み、また、その核酸配列が、完全長 / 成熟型 B 4 G A L T 1 ポリペプチドの 352 位に対応する位置でセリンをコードすることを条件とする、B 4 G A L T 1 ポリペプチドをコードする cDNA 分子を提供する。

#### 【 0 0 1 0 】

本開示はまた、これらの核酸分子のいずれか 1 つ以上を含むベクターまたは外来性ドナー配列を提供する。

本開示はまた、B 4 G A L T 1 ポリペプチドと少なくとも約 90 % 同一のアミノ酸配列を含み、完全長 / 成熟型 B 4 G A L T 1 ポリペプチドの 352 位に対応する位置にセリンを含む単離されたポリペプチドを提供する。

#### 【 0 0 1 1 】

本開示はまた、宿主細胞において活性な異種プロモーターに作動可能に連結したこれら

の核酸分子のいずれか1つ以上を含む宿主細胞を提供する。

本開示はまた、B 4 G A L T 1 ポリペプチドをコードする核酸分子を含む宿主細胞を培養し、単離されたポリペプチドを回収することによる、B 4 G A L T 1 ポリペプチドの產生方法を提供し、その場合、核酸分子は宿主細胞において活性な異種プロモーターに作動可能に連結しており、それによって核酸分子は発現する。

#### 【 0 0 1 2 】

本開示はまた、これらの核酸分子またはポリペプチド、及びそれらの安定性を高めるための担体を含む組成物を提供する。

本開示はまた、ヒト対象由来の生体試料に対してアッセイを行い、生体試料中の核酸分子が、完全長／成熟型B 4 G A L T 1 ポリペプチドの352位に対応する位置にセリンを含むバリアント型B 4 G A L T 1 ポリペプチドをコードする核酸配列を含むかどうかを判定することを含む、ヒト対象におけるB 4 G A L T 1 バリアント型核酸分子(r s 5 5 1 5 6 4 6 8 3と呼ばれるSNPを含む)の存在または非存在の検出方法を提供する。

10

#### 【 0 0 1 3 】

本開示はまた、ヒト対象由来の生体試料に対してアッセイを行い、バリアント型B 4 G A L T 1 ポリペプチドの存在を判定することを含む、ヒト対象の完全長／成熟型B 4 G A L T 1 ポリペプチドの352位に対応する位置にセリンを含むバリアント型B 4 G A L T 1 ポリペプチドの存在の検出方法を提供する。

#### 【 0 0 1 4 】

本開示はまた：a)ヒト対象由来の生体試料に対してアッセイを行い、生体試料中の核酸分子が、完全長／成熟型B 4 G A L T 1 ポリペプチドの352位に対応する位置にセリンを含むバリアント型B 4 G A L T 1 ポリペプチドをコードする核酸配列を含むかどうかを判定し、b)完全長／成熟型B 4 G A L T 1 ポリペプチドの352位に対応する位置にセリンを含むバリアント型B 4 G A L T 1 ポリペプチドをコードする核酸配列を含む核酸分子が生体試料中に検出される場合、ヒト対象を、心血管病態を発症するリスクが低いと分類し、あるいは、完全長／成熟型B 4 G A L T 1 ポリペプチドの352位に対応する位置にセリンを含むバリアント型B 4 G A L T 1 ポリペプチドをコードする核酸配列を含む核酸分子が生体試料中に検出されない場合、ヒト対象を、心血管病態を発症するリスクが高いと分類することを含む、心血管病態の発症に対するヒト対象の感受性の判定方法を提供する。

20

#### 【 0 0 1 5 】

本開示はまた：a)ヒト対象由来の生体試料に対してアッセイを行い、生体試料中のB 4 G A L T 1 ポリペプチドが、352位に対応する位置にセリンを含むかどうかを判定し、b)完全長／成熟型B 4 G A L T 1 ポリペプチドの352位に対応する位置にセリンを含むB 4 G A L T 1 ポリペプチドが生体試料中に検出される場合、ヒト対象を、心血管病態を発症するリスクが低いと分類し、あるいは、完全長／成熟型B 4 G A L T 1 ポリペプチドの352位に対応する位置にセリンを含むB 4 G A L T 1 ポリペプチドが生体試料中に検出されない場合、ヒト対象を、心血管病態を発症するリスクが高いと判定することを含む、心血管病態の発症に対するヒト対象の感受性の判定方法を提供する。

30

#### 【 0 0 1 6 】

本開示はまた、内在性B 4 G A L T 1 遺伝子に結合するか、または切断するようにCas 酵素を誘導するのに有効なガイドRNA分子を提供し、ガイドRNAは、野生型B 4 G A L T 1 遺伝子の53575～53577の位置に対応する位置を含むか、またはそれに近接する(例えば、以下で説明するような特定の数のヌクレオチド以内で)内在性B 4 G A L T 1 遺伝子内のガイドRNA認識配列にハイブリダイズするDNAターゲティングセグメントを含む。

40

#### 【 0 0 1 7 】

本開示はまた、細胞のゲノムを：a)Casタンパク質、及びb)Casタンパク質と複合体を形成し、内在性B 4 G A L T 1 遺伝子内のガイドRNA認識配列にハイブリダイズするガイドRNAと接触させることを含む、細胞の内在性B 4 G A L T 1 遺伝子の改変

50

方法を提供し、その場合、ガイドRNA認識配列は、野生型B4GALT1遺伝子の53575～53577位に対応する位置を含むか、またはそれに近接し（例えば、以下で説明するような特定の数のヌクレオチド以内で）、Casタンパク質は内在性B4GALT1遺伝子を切断する。

【0018】

本開示はまた、細胞のゲノムを：a) Casタンパク質、及びb) Casタンパク質と複合体を形成し、内在性B4GALT1遺伝子内の第1のガイドRNA認識配列にハイブリダイズする第1のガイドRNAと接触させることを含む、細胞の内在性B4GALT1遺伝子の改変方法を提供し、その場合、第1のガイドRNA認識配列は、B4GALT1遺伝子の開始コドンを含むか、開始コドンから約1,000ヌクレオチド以内に存在し、Casタンパク質は、内在性B4GALT1遺伝子の発現を切断するか、または変化させる。

【0019】

本開示はまた、発現ベクターを細胞に導入することを含む、細胞の改変方法を提供し、発現ベクターは、完全長／成熟型B4GALT1ポリペプチドの352位に対応する位置にセリンを含むB4GALT1ポリペプチドをコードするヌクレオチド配列を含む組換えB4GALT1遺伝子を含む。

【0020】

本開示はまた、発現ベクターを細胞に導入することを含む、細胞の改変方法を提供し、発現ベクターは、完全長／成熟型B4GALT1ポリペプチドの352位に対応する位置にセリンを含むB4GALT1ポリペプチドと少なくとも約90%同一であるポリペプチドをコードする核酸分子を含み、そのポリペプチドもまた、完全長／成熟型B4GALT1ポリペプチドの352位に対応する位置にセリンを含む。

【0021】

本開示はまた、ポリペプチドまたはその断片を細胞に導入することを含む細胞の改変方法を提供し、ポリペプチドは、完全長／成熟型B4GALT1ポリペプチドの352位に対応する位置にセリンを含むB4GALT1ポリペプチドと少なくとも90%同一であり、そのポリペプチドもまた、完全長／成熟型B4GALT1ポリペプチドの352位に対応する位置にセリンを含む。

【0022】

本開示はまた、B4GALT1バリアント型核酸分子またはポリペプチド（rs551564683と呼ばれるSNPを含む）の保因者ではなく、心血管病態を発症しているか、または発症しやすい対象の治療方法を提供し、本方法は、対象に：a) Casタンパク質またはCasタンパク質をコードする核酸；b) ガイドRNAまたはガイドRNAをコードする核酸を導入することを含み、ガイドRNAは、Casタンパク質と複合体を形成し、内在性B4GALT1遺伝子内のガイドRNA認識配列にハイブリダイズし、ガイドRNA認識配列は、野生型B4GALT1遺伝子の53575～53577位に対応する位置を含むか、またはそれに近接し、本方法はまた、c) 野生型B4GALT1遺伝子の53575～53577位に対応する位置の5'の標的配列にハイブリダイズする5'相同性アーム、野生型B4GALT1遺伝子の53575～53577位に対応する位置の3'の標的配列にハイブリダイズする3'相同性アーム、ならびに5'相同性アーム及び3'相同性アームに隣接する、完全長／成熟型B4GALT1ポリペプチドの352位に対応する位置にセリンを含むB4GALT1ポリペプチドをコードするヌクレオチド配列を含む核酸インサートを含む、外来性ドナー配列を導入することを含み、Casタンパク質は、対象の細胞の内在性B4GALT1遺伝子を切断し、外来性ドナー配列は、細胞の内在性B4GALT1遺伝子と組み換わり、外来性ドナー配列の内在性B4GALT1遺伝子との組換え時に、野生型B4GALT1遺伝子の53575～53577位に対応するヌクレオチドにセリンが挿入される。

【0023】

本開示はまた、B4GALT1バリアント型核酸分子またはポリペプチド（rs551

10

20

30

40

50

5 6 4 6 8 3 と呼ばれる S N P を含む) の保因者ではなく、心血管病態を発症しているか、または発症しやすい対象の治療方法を提供し、本方法は、対象に： a ) C a s タンパク質または C a s タンパク質をコードする核酸、 b ) 第 1 のガイド R N A または第 1 のガイド R N A をコードする核酸を導入することを含み、第 1 のガイド R N A は、 C a s タンパク質と複合体を形成し、内在性 B 4 G A L T 1 遺伝子内の第 1 のガイド R N A 認識配列にハイブリダイズし、第 1 のガイド R N A 認識配列は、内在性 B 4 G A L T 1 遺伝子の開始コドンを含むか、開始コドンから約 1,000 ヌクレオチド以内に存在し、本方法はまた、 c ) 完全長 / 成熟型 B 4 G A L T 1 ポリペプチドの 352 位に対応する位置にセリンを含む B 4 G A L T 1 ポリペプチドをコードするヌクレオチド配列を含む組換え B 4 G A L T 1 遺伝子を含む発現ベクターを導入することを含み、 C a s タンパク質は、対象の細胞の内在性 B 4 G A L T 1 遺伝子を切断するかまたは発現を変化させ、発現ベクターは、対象の細胞内で組換え B 4 G A L T 1 遺伝子を発現する。

【 0 0 2 4 】

本開示はまた、 B 4 G A L T 1 バリアント型核酸分子またはポリペプチド ( r s 5 5 1 5 6 4 6 8 3 と呼ばれる S N P を含む) の保因者ではなく、心血管病態を発症しているか、または発症しやすい対象の治療方法であって、対象の細胞の内在性 B 4 G A L T 1 遺伝子内の配列にハイブリダイズして B 4 G A L T 1 ポリペプチドの発現を減少させるアンチセンス D N A 、 R N A 、 s i R N A 、または s h R N A を対象に導入することを含む、方法を提供する。

【 0 0 2 5 】

本開示はまた、発現ベクターを対象に導入することを含む、 B 4 G A L T 1 バリアント型核酸分子またはポリペプチド ( r s 5 5 1 5 6 4 6 8 3 と呼ばれる S N P を含む) の保因者ではなく、心血管病態を発症しているか、または発症しやすい対象の治療方法を提供し、発現ベクターは、完全長 / 成熟型 B 4 G A L T 1 ポリペプチドの 352 位に対応する位置にセリンを含む B 4 G A L T 1 ポリペプチドをコードするヌクレオチド配列を含む組換え B 4 G A L T 1 遺伝子を含み、発現ベクターは、対象の細胞内で組換え B 4 G A L T 1 遺伝子を発現する。

【 0 0 2 6 】

本開示はまた、発現ベクターを対象に導入することを含む、 B 4 G A L T 1 バリアント型核酸分子またはポリペプチド ( r s 5 5 1 5 6 4 6 8 3 と呼ばれる S N P を含む) の保因者ではなく、心血管病態を発症しているか、または発症しやすい対象の治療方法を提供し、発現ベクターは、完全長 / 成熟型 B 4 G A L T 1 ポリペプチドの 352 位に対応する位置にセリンを含む B 4 G A L T 1 ポリペプチドをコードする核酸分子を含み、発現ベクターは、対象の細胞内で B 4 G A L T 1 ポリペプチドをコードする核酸を発現する。

【 0 0 2 7 】

本開示はまた、 m R N A を対象に導入することを含む、 B 4 G A L T 1 バリアント型核酸分子またはポリペプチド ( r s 5 5 1 5 6 4 6 8 3 と呼ばれる S N P を含む) の保因者ではなく、心血管病態を発症しているか、または発症しやすい対象の治療方法を提供し、 m R N A は、完全長 / 成熟型 B 4 G A L T 1 ポリペプチドの 352 位に対応する位置にセリンを含む B 4 G A L T 1 ポリペプチドをコードし、 m R N A は、対象の細胞内で B 4 G A L T 1 ポリペプチドを発現する。

【 0 0 2 8 】

本開示はまた、完全長 / 成熟型 B 4 G A L T 1 ポリペプチドの 352 位に対応する位置にセリンを含む B 4 G A L T 1 ポリペプチドまたはその断片を対象に導入することを含む、 B 4 G A L T 1 バリアント型核酸分子またはポリペプチド ( r s 5 5 1 5 6 4 6 8 3 と呼ばれる S N P を含む) の保因者ではなく、心血管病態を発症しているか、または発症しやすい対象の治療方法を提供する。

【 0 0 2 9 】

本明細書に記載または例示する方法のいずれかにおいて、心血管病態は、アテローム性動脈硬化症のリスクを増加させる 1 つ以上の血清脂質のレベルを有し得る。血清脂質には

10

20

30

40

50

、コレステロール、LDL、HDL、トリグリセリド、HDLコレステロール、及び非HDLコレステロール、またはその任意の亜分画（例えば、HDL2、HDL2a、HDL2b、HDL2c、HDL3、HDL3a、HDL3b、HDL3c、HDL3d、LDL1、LDL2、LDL3、リポタンパク質A、Lpa1、Lpa1、Lpa3、Lpa4、またはLpa5）の1つ以上が含まれる。心血管病態には、冠動脈石灰化レベルの上昇が含まれ得る。心血管病態には、心膜脂肪レベルの上昇が含まれ得る。心血管病態には、アテローム血栓性病態が含まれ得る。アテローム血栓性病態には、フィブリノーゲンレベルの上昇が含まれ得る。アテローム血栓性病態には、フィブリノーゲン媒介性血餅が含まれ得る。心血管病態には、フィブリノーゲンレベルの上昇が含まれ得る。心血管病態には、フィブリノーゲン媒介性血餅が含まれ得る。心血管病態には、フィブリノーゲン活性の関与から形成される血餅が含まれ得る。フィブリノーゲン媒介性血餅またはフィブリノーゲン活性の関与から形成される血餅は、体内的どの静脈または動脈にも存在し得る。

【図面の簡単な説明】

【0030】

【図1】バリアント型B4GALT1とLDLの代表的なゲノム全体での関連付けの結果を示す。

【図2】バリアント型B4GALT1とLDLの代表的なTOPMed WGS関連付けの結果を示す。

【図3】B4GALT1に関連する上位SNPの代表的なハプロタイプ構造の結果を示す。

【図4】エクソーム配列決定によって同定したアーミッシュにおけるLDLとバリアント型B4GALT1遺伝子の相関性を示す。

【図5-1】アーミッシュでは、バリアント型B4GALT1遺伝子の出現頻度が1000倍超濃い頻度であることを示す。

【図5-2】アーミッシュでは、バリアント型B4GALT1遺伝子の出現頻度が1000倍超濃い頻度であることを示す。

【図6-1】B4GALT1 Asn352Serと血清脂質の減少との相関性を示す。

【図6-2】B4GALT1 Asn352Serと血清脂質の減少との相関性を示す。

【図7-1】B4GALT1 Asn352Serと血清脂質の減少及びASTの増加との高い相関性を示す。

【図7-2】B4GALT1 Asn352Serと血清脂質の減少及びASTの増加との高い相関性を示す。

【図8】B4GALT1 Asn352Serとすべての脂質亜分画との相関性を示す。

【図9】B4GALT1 Asn352Serとフィブリノーゲンレベルの低下との相関性を示す。

【図10】示された濃度のアンチセンスモルホリノオリゴヌクレオチドを注射したゼブラフィッシュ幼虫の受精後5日でのb4galta1転写産物の減少を示す。

【図11】示された濃度のアンチセンスモルホリノオリゴヌクレオチドを注射したゼブラフィッシュ幼虫の受精5日後でのアンチセンスモルホリノオリゴヌクレオチドのオフターゲット効果の診断マーカーを示す。

【図12】実験ごとの受精後5日間のゼブラフィッシュ幼虫100匹のホモジネートの平均LDL濃度を示す。

【図13】ゼブラフィッシュにおける50pgのヒトB4GALT1 mRNAの共発現によるLDL-c表現型のレスキューを示す。

【図14】標的遺伝子型判定を使用したB4GALT1 N352SとLDLの遺伝的関連付けの結果を示す。

【図15】Flag-352AsnまたはFlag-352Serの細胞内局在の共焦点顕微鏡画像を示す。

【図16】トランスゴルジ網マーカーTGN46に関連する内在性B4GALT1、Flag-352Asn、及びFlag-352Ser細胞内局在の共焦点顕微鏡画像を示す。

【図17-1】（パネルA及びB）は、B4GALT1タンパク質の定常状態レベルに対

10

20

30

40

50

する 352Ser の効果を示す。(パネル A) 352Asn または 352Ser F1a g タグタンパク質融合体のいずれかを遊離 EGFP とともに発現する COS7 細胞。(パネル B) RT-qPCR 分析によって測定した B4GALT1 遺伝子の mRNA 発現レベル。

【図 17-2】(パネル A 及び B) は、B4GALT1 タンパク質の定常状態レベルに対する 352Ser の効果を示す。(パネル A) 352Asn または 352Ser F1a g タグタンパク質融合体のいずれかを遊離 EGFP とともに発現する COS7 細胞。(パネル B) RT-qPCR 分析によって測定した B4GALT1 遺伝子の mRNA 発現レベル。

【図 18】(パネル A、B、及び C) は、352Ser 変異が活性に及ぼす影響を示す。(パネル A 及び B) COS7 細胞で発現し、B4GALT1 または F1ag のウエスタンプロットで分析した 352Asn または 352Ser F1ag タグタンパク質融合体を発現する COS7 細胞。(パネル C) 免疫沈降物の B4GALT1 活性。

【図 19】B4GALT1 N352S 遺伝子型群ごとのトリシアロ/ジオリゴ比を示す。

【図 20-1】B4GALT1 N352S のマイナー (SS) 及びメジャー (NN) ホモ接合体の一致ペア由来糖タンパク質の N-グリカン分析の代表的な HILIC-FLR-MS スペクトルを示す。

【図 20-2】B4GALT1 N352S のマイナー (SS) 及びメジャー (NN) ホモ接合体の一致ペア由来糖タンパク質の N-グリカン分析の代表的な HILIC-FLR-MS スペクトルを示す。

#### 【発明を実施するための形態】

##### 【0031】

本明細書に記載するように、配列決定研究により、完全長 / 成熟型 B4GALT1 ポリペプチドの 352 位に対応する位置にアスパラギンの代わりにセリンを含む B4GALT1 のバリエントが、オールド・オーダー・アーミッシュ (OOA) の約 11% ~ 12% の個体に存在し (代替対立遺伝子出現頻度 = 6%)、また、一般的な集団においては非常にまれであることが同定された。この変異は、398 アミノ酸長のヒトタンパク質の 352 位において (N352S)、またはショート型アイソフォームの 311 位において、アスパラギンをセリンへ変化させる。バリエント型 B4GALT1 は、低密度リボタンパク質コレステロール (LDL)、総コレステロール、ならびにフィブリノーゲン及び eGFR のレベルが低いこと、アスパラギン酸トランスアミナーゼ (AST) (ただし、アラニントランスアミナーゼ (ALT) ではない) のレベル、ならびにクレアチニナーゼ及びクレアチニンの血清レベルが高いこと、筋肉組織 (ただし、肝臓または赤血球ではない) での発現、ならびに好塩基球の減少と相関性があることが観察されている。N352S バリエントは、1 つ以上の心血管病態に対して保護作用があると考えられている。さらに、そのバリエント状態を含む B4GALT1 は、患者の心血管病態の発症リスクを診断するために使用し得ると考えられる。

##### 【0032】

所与のアミノ酸配列またはポリヌクレオチド配列の番号付けに関して使用する場合の語句「に対応する」とは、所与のアミノ酸配列またはポリヌクレオチド配列を参照配列と比較する場合の指定する参照配列の残基の番号付けを指す (本明細書中の参照配列は、(野生型 / 完全長) B4GALT1 のポリヌクレオチド (gDNA 配列、mRNA 配列、cDNA 配列) またはポリペプチドである)。言い換えれば、所与のポリマーの残基数または残基位置は、所与のアミノ酸配列またはポリヌクレオチド配列内の残基の実際の数値位置ではなく、参照配列に対して指定される。例えば、2 つの配列間の残基の一致性を最適化するためにギャップを導入することにより、特定のアミノ酸配列を参照配列にアライメントすることができる。これらの場合において、ギャップは存在するものの、特定のアミノ酸配列またはポリヌクレオチド配列における残基の番号付けは、それをアライメントした参照配列に対して行う。

##### 【0033】

10

20

30

40

50

本明細書中で使用する場合、単数形の冠詞「a」、「an」、及び「the」は、文脈からそうでないことが明確に示されない限り、複数の参照を含む。

本明細書中で使用する場合、及び文脈からそうでないことが明白でない限り、「約」は、述べられた値の測定の標準誤差（例えば、SEM）の範囲内の値を包含する。

#### 【0034】

本明細書中で使用する場合、「及び／または」とは、1つ以上の関連する列挙項目のありとあらゆる可能な組み合わせ、及び代替（「または」）で解釈する場合は組み合わせの欠如を指し、また、それらを包含する。

#### 【0035】

本明細書中で使用する場合、用語「含む（comprising）」または「含む（including）」とは、列挙する要素の1つ以上が、具体的に列挙していない他の要素を含み得ることを意味する。例えば、タンパク質を「含む（comprising）」または「含む（including）」組成物は、タンパク質を単独で、または他の成分と組み合わせて含み得る。移行句「本質的に～からなる」とは、請求項の範囲が、その請求項において列挙する特定の要素、及び請求する対象の基本的及び新規の特徴（複数可）に実質的に影響しない要素を包含すると解釈されることを意味する。したがって、本開示の請求項において使用する場合、用語「本質的に～からなる」は、「含む」と同義であると解釈されることを意図していない。

10

#### 【0036】

本明細書中で使用する場合、「任意の」または「場合により」とは、続いて記載する事象または状況が発生する場合と発生しない場合があり、その記載が、事象または状況が発生する場合と発生しない場合を含むことを意味する。

20

#### 【0037】

本明細書中で使用する場合、「または」とは、特定のリストの任意の1つのメンバーを指し、そのリストのメンバーの任意の組み合わせも含む。

値の範囲の指定には、範囲内のまたは範囲を定義するすべての整数（2つの境界値を含む）、及び範囲内の整数によって定義されるすべての部分範囲が含まれる。

#### 【0038】

明確にするために別個の実施形態で記載している本開示の特定の特徴は、組み合わせて単一の実施形態として提供することもできることを理解されたい。逆に、簡潔にするために単一の実施形態で記載している本開示の様々な特徴は、別個にまたは任意の適切な部分組み合わせとして提供することもできる。

30

#### 【0039】

本開示は、単離されたB4GALT1ゲノム及びmRNAバリアント、B4GALT1 cDNAバリアント、またはその任意の相補体、ならびに単離されたB4GALT1ポリペプチドバリアントを提供する。これらのバリアントは、血清脂質レベルの上昇、フィブリノーゲンレベルの上昇、冠動脈石灰化、冠動脈疾患（CAD）、及びアスパラギン酸アミノトランスフェラーゼ（AST）レベルの増加（ただし、アラニントランスアミナーゼ（ALT）は増加しない）を含むがこれらに限定されない様々な心血管病態の発症リスクの低下に関連すると考えられる。理論に拘泥することを望むものではないが、これらのB4GALT1バリアントは、実験的に観察されたASTレベルの増加とALTが増加しなかったことによって証明されるように、肝臓や赤血球ではなく筋肉組織での発現に関連すると考えられる。B4GALT1ゲノムバリアント及びmRNAバリアント、B4GALT1 cDNAバリアント、及び単離されたB4GALT1ポリペプチドバリアントを含む組成物も本明細書において提供する。B4GALT1ゲノムバリアント及びmRNAバリアント及びB4GALT1 cDNAバリアントにハイブリダイズする核酸分子も本明細書において提供する。本開示はまた、B4GALT1ゲノムバリアント及びmRNAバリアント、B4GALT1 cDNAバリアント、及びB4GALT1ポリペプチドバリアントを含むベクター及び細胞も提供する。

40

#### 【0040】

50

本開示はまた、生体試料中のゲノムバリアント及び／またはmRNAバリアント、B4GALT1 cDNAバリアント、またはその相補体、及び／またはB4GALT1ポリペプチドバリアントの存在及び／またはレベルの検出方法を提供する。また、心血管病態の発症に対する対象の感受性の判定方法、及び心血管病態を発症しているか、または心血管病態のリスクを有する対象の診断方法も提供する。また、ヌクレアーゼ剤、外来性ドナーパリエント、転写活性化因子、転写抑制因子、及び組換えB4GALT1遺伝子またはB4GALT1ポリペプチドをコードする核酸を発現するための発現ベクターの任意の組み合わせの使用による細胞の改変方法も提供する。また、心血管病態を発症しているか、または心血管病態を発症するリスクを有する対象を治療するための治療及び予防方法も提供する。

## 【0041】

10

野生型ヒトゲノムB4GALT1核酸は、長さがおよそ56.7kbであり、6つのエクソンを有し、ヒトゲノムの染色体9に位置する。例示的な野生型ヒトゲノムB4GALT1配列には、NCBI登録番号NG\_008919.1(配列番号1)が割り当てられている。ヒトゲノムB4GALT1のバリアントは、配列番号2に示しており、一塩基多型(SNP)を含んでいる(53576位のAからGへ、本明細書中ではバリアント型B4GALT1と呼ばれる)。このバリアントのSNPにより、完全長／成熟型B4GALT1ポリペプチドの352位に対応する位置に、野生型B4GALT1ポリペプチドによってコードされるアスパラギンではなく、コードされるB4GALT1バリアントポリペプチドのセリンが生じる。バリアント型ヒトゲノムB4GALT1核酸は、例えば、野生型ヒトゲノムB4GALT1の53575～53577位に対応する位置における3つの塩基「aat」とは対照的に、野生型ヒトゲノムB4GALT1の53575～53577位に対応する位置におけるセリンをコードする3つの塩基(例えば「agt」)を含む(それぞれ、配列番号2と配列番号1とを比較)。いくつかの実施形態では、単離された核酸分子は配列番号2を含む。いくつかの実施形態では、単離された核酸分子は配列番号2からなる。いくつかの実施形態では、単離された核酸分子は、本明細書に開示する任意のゲノムB4GALT1核酸分子の相補体である。

## 【0042】

20

いくつかの実施形態では、単離された核酸分子は、配列番号2と、少なくとも約70%、少なくとも約75%、少なくとも約80%、少なくとも約85%、少なくとも約90%、少なくとも約95%、少なくとも約96%、少なくとも約97%、少なくとも約98%、少なくとも約99%、または100%同一である核酸配列を含むかまたはそれからなる。いくつかの実施形態では、そのような核酸配列はまた、配列番号2の53575～53577位に対応するヌクレオチドを含む。いくつかの実施形態では、単離された核酸分子は、B4GALT1遺伝子のエクソン1～6を含む配列番号2の部分と、少なくとも約70%、少なくとも約75%、少なくとも約80%、少なくとも約85%、少なくとも約90%、少なくとも約95%、少なくとも約96%、少なくとも約97%、少なくとも約98%、少なくとも約99%、または100%同一である核酸配列を含むかまたはそれからなる。いくつかの実施形態では、そのような核酸配列は、配列番号2の53575～53577位に対応するヌクレオチドも含む。いくつかの実施形態では、単離された核酸分子は、エクソン5を含む配列番号2の部分と、少なくとも約70%、少なくとも約75%、少なくとも約80%、少なくとも約85%、少なくとも約90%、少なくとも約95%、少なくとも約96%、少なくとも約97%、少なくとも約98%、少なくとも約99%、または100%同一である核酸配列を含むかまたはそれからなる。いくつかの実施形態では、そのような核酸配列はまた、配列番号2の53575～53577位に対応するヌクレオチドを含む。いくつかの実施形態では、単離された核酸分子は、配列番号2と少なくとも約90%同一の核酸配列を含み、ただしその核酸配列は、配列番号2の53575～53577位に対応するヌクレオチドを含む。

## 【0043】

40

核酸内の核酸配列の特定のストレッチ間の相補性の割合は、BLASTプログラム(基本的なローカルアライメント検索ツール)及びPowerBLASTプログラム(Alt

50

schul et al., J. Mol. Biol., 1990, 215, 403-410; Zhang and Madden, Genome Res., 1997, 7, 649-656)を使用して、またはSmith and Watermanのアルゴリズム(Adv. Appl. Math., 1981, 2, 482-489)を使用するGapプログラム(Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, Madison Wis.)を使用してデフォルトの設定を用いて、ルーチン的に決定することができる。

#### 【0044】

いくつかの実施形態では、単離された核酸分子は、ゲノム配列全体よりも短い配列を含む。いくつかの実施形態では、単離された核酸分子は、配列番号2の、少なくとも約15、少なくとも約20、少なくとも約25、少なくとも約30、少なくとも約35、少なくとも約40、少なくとも約45、少なくとも約50、少なくとも約60、少なくとも約70、少なくとも約80、少なくとも約90、少なくとも約100、少なくとも約200、少なくとも約300、少なくとも約400、少なくとも約500、少なくとも約600、少なくとも約700、少なくとも約800、少なくとも約900、少なくとも約1000、少なくとも約2000、少なくとも約3000、少なくとも約4000、少なくとも約5000、少なくとも約6000、少なくとも約7000、少なくとも約8000、少なくとも約9000、少なくとも約10000、少なくとも約11000、少なくとも約12000、少なくとも約13000、少なくとも約14000、少なくとも約15000、少なくとも約16000、少なくとも約17000、少なくとも約18000、少なくとも約19000、または少なくとも約20000の連続ヌクレオチドを含むかまたはそれからなる。いくつかの実施形態では、そのような単離された核酸分子は、配列番号2の53575～53577位に対応するヌクレオチドも含む。いくつかの実施形態では、単離された核酸分子は、配列番号2の、少なくとも約15、少なくとも約20、少なくとも約25、少なくとも約30、少なくとも約35、少なくとも約40、少なくとも約45、少なくとも約50、少なくとも約60、少なくとも約70、少なくとも約80、少なくとも約90、少なくとも約100、少なくとも約200、少なくとも約300、少なくとも約400、少なくとも約500、少なくとも約600、少なくとも約700、少なくとも約800、少なくとも約900、または少なくとも約1000の連続ヌクレオチドを含むかまたはそれからなる。いくつかの実施形態では、そのような単離された核酸分子は、配列番号2の53575～53577位に対応するヌクレオチドも含む。いくつかの実施形態では、単離された核酸分子は、配列番号2のエクソン5の、少なくとも約15、少なくとも約20、少なくとも約25、少なくとも約30、少なくとも約35、少なくとも約40、少なくとも約45、少なくとも約50、少なくとも約60、少なくとも約70、少なくとも約80、少なくとも約90、少なくとも約100、少なくとも約200、少なくとも約300、少なくとも約400、少なくとも約500、少なくとも約600、少なくとも約700、少なくとも約800、少なくとも約900、または少なくとも約1000の連続ヌクレオチドを含むかまたはそれからなる。いくつかの実施形態では、そのような単離された核酸分子は、配列番号2の53575～53577位に対応するヌクレオチドも含む。

#### 【0045】

例えば、いくつかの実施形態では、単離された核酸分子は、配列番号2の少なくとも15の連続ヌクレオチドを含み、連続ヌクレオチドは、配列番号2のヌクレオチド53575～53577を含む。いくつかのそのような実施形態では、単離された核酸分子は、配列番号2の、少なくとも20、少なくとも25または少なくとも30の連続ヌクレオチドを含む。いくつかの実施形態では、単離された核酸分子は、配列番号2の15～50の連続ヌクレオチドを含み、連続ヌクレオチドは、配列番号2のヌクレオチド53575～53577を含む。いくつかのそのような実施形態では、単離された核酸分子は、配列番号2の、少なくとも20、少なくとも25または少なくとも30の連続ヌクレオチドを含む。

10

20

30

40

50

## 【0046】

いくつかの実施形態では、本開示は、配列番号2の部分と少なくとも90%同一である核酸配列を含む単離された核酸を提供し、配列番号2の部分は、配列番号2のヌクレオチド53575～53577を含み、配列番号2の部分は、少なくとも15ヌクレオチド長である。いくつかのそのような実施形態では、配列番号2の部分は、少なくとも20ヌクレオチド長、少なくとも25ヌクレオチド長、または少なくとも30ヌクレオチド長である。いくつかの実施形態では、本開示は、配列番号2の部分と少なくとも90%同一である核酸配列を含む単離された核酸を提供し、配列番号2の部分は、配列番号2のヌクレオチド53575～53577を含み、配列番号2の部分は、15～50ヌクレオチド長である。いくつかのそのような実施形態では、配列番号2の部分は、少なくとも20ヌクレオチド長、少なくとも25ヌクレオチド長、または少なくとも30ヌクレオチド長である。

10

## 【0047】

いくつかの実施形態では、本開示は、配列番号2の部分と少なくとも95%同一である核酸配列を含む単離された核酸を提供し、配列番号2の部分は、配列番号2のヌクレオチド53575～53577を含み、配列番号2の部分は、少なくとも15ヌクレオチド長である。いくつかのそのような実施形態では、配列番号2の部分は、少なくとも20ヌクレオチド長、少なくとも25ヌクレオチド長、または少なくとも30ヌクレオチド長である。いくつかの実施形態では、本開示は、配列番号2の部分と少なくとも95%同一である核酸配列を含む単離された核酸を提供し、配列番号2の部分は、配列番号2のヌクレオチド53575～53577を含み、配列番号2の部分は、15～50ヌクレオチド長である。いくつかのそのような実施形態では、配列番号2の部分は、少なくとも20ヌクレオチド長、少なくとも25ヌクレオチド長、または少なくとも30ヌクレオチド長である。

20

## 【0048】

そのような単離された核酸分子は、例えば、バリアント型B4GALT1 mRNA及びタンパク質を発現するために、または外来性ドナー配列として使用することができる。集団内の遺伝子配列は、SNPなどの多型により変化し得ることが理解される。本明細書において提供する例は、例示的な配列に過ぎず、他の配列も可能である。

## 【0049】

いくつかの実施形態では、単離された核酸分子は、バリアント型B4GALT1ミニ遺伝子を含み、この遺伝子では、対応する野生型B4GALT1遺伝子に対して、配列番号2の1つ以上の非必須セグメントが欠失している。いくつかの実施形態では、欠失した非必須セグメントには、1つ以上のインtron配列が含まれる。いくつかの実施形態では、B4GALT1ミニ遺伝子は、例えば、バリアント型B4GALT1(配列番号2)由來のエクソン1～6のいずれか1つ以上に対応するエクソン、またはそのようなエクソンの任意の組み合わせを含むことができる。いくつかの実施形態では、ミニ遺伝子は、配列番号2のエクソン5を含むかまたはそれからなる。いくつかの実施形態では、B4GALT1ミニ遺伝子は、エクソン1～6のいずれか1つ以上を含む配列番号2の部分、またはそのようなエクソンの任意の組み合わせと、少なくとも約70%、少なくとも約75%、少なくとも約80%、少なくとも約85%、少なくとも約90%、少なくとも約95%、少なくとも約96%、少なくとも約97%、少なくとも約98%、少なくとも約99%、または100%同一である。いくつかの実施形態では、B4GALT1ミニ遺伝子は、エクソン1～6のいずれか1つ以上を含む配列番号2の部分、またはそのようなエクソンの任意の組み合わせと、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、または100%同一であり、配列番号2の53575～53577位に対応するヌクレオチドを含む。いくつかの実施形態では、B4GALT1ミニ遺伝子は、エクソン5を含む配列番号2の部分と、少なくとも約70%、少なくとも約75%、少なくとも約80%、少なくとも約85%、少なくとも約90%、少なくとも約95%、少なくとも約96%、少なくとも約97%、少なくとも約98%、少なくとも約99%、または100%同一である。

30

40

50

## 【0050】

本開示は、バリアント型 B 4 G A L T 1 ゲノム配列またはバリアント型 B 4 G A L T 1 ミニ遺伝子にハイブリダイズする単離された核酸分子も提供する。いくつかの実施形態では、そのような単離された核酸分子は、少なくとも約 15、少なくとも約 20、少なくとも約 25、少なくとも約 30、少なくとも約 35、少なくとも約 40、少なくとも約 45、少なくとも約 50、少なくとも約 60、少なくとも約 70、少なくとも約 80、少なくとも約 90、少なくとも約 100、少なくとも約 200、少なくとも約 300、少なくとも約 400、少なくとも約 500、少なくとも約 600、少なくとも約 700、少なくとも約 800、少なくとも約 900、少なくとも約 1000、少なくとも約 2000、少なくとも約 3000、少なくとも約 4000、少なくとも約 5000、少なくとも約 6000、少なくとも約 7000、少なくとも約 8000、少なくとも約 9000、少なくとも約 10000、少なくとも約 11000、少なくとも約 12000、少なくとも約 13000、少なくとも約 14000、少なくとも約 15000、少なくとも約 16000、少なくとも約 17000、少なくとも約 18000、少なくとも約 19000、または少なくとも約 20000 ヌクレオチドを含むかそれからなる。いくつかの実施形態では、そのような単離された核酸分子はまた、配列番号 2 の 53575 ~ 53577 位にハイブリダイズする。いくつかの実施形態では、単離された核酸分子は、配列番号 2 の 53575 ~ 53577 位の、約 1000、約 500、約 400、約 300、約 200、約 100、約 50、約 45、約 40、約 35、約 30、約 25、約 20、約 15、約 10、または約 5 ヌクレオチドを含むかまたはそれ以内であるセグメントにおいて、バリアント型 B 4 G A L T 1 ゲノムまたはミニ遺伝子の部分にハイブリダイズする。いくつかの実施形態では、単離された核酸分子は、バリアント型 B 4 G A L T 1 ゲノム D N A またはミニ遺伝子と、少なくとも約 70%、少なくとも約 75%、少なくとも約 80%、少なくとも約 85%、少なくとも約 90%、少なくとも約 95%、少なくとも約 96%、少なくとも約 97%、少なくとも約 98%、少なくとも約 99%、または 100% 同一である核酸分子の少なくとも約 15 の連続ヌクレオチドにハイブリダイズする。いくつかの実施形態では、そのような単離された核酸分子はまた、配列番号 2 の 53575 ~ 53577 位にハイブリダイズする。いくつかの実施形態では、単離された核酸分子は、約 15 ~ 約 100 ヌクレオチド、または約 15 ~ 約 35 ヌクレオチドを含むか、またはそれからなる。

## 【0051】

例えば、いくつかの実施形態では、本開示は、少なくとも 15 ヌクレオチドを含む単離された核酸分子を提供し、単離された核酸分子は、配列番号 2 の配列を含む核酸にハイブリダイズし、単離された核酸分子は、配列番号 2 の部分にハイブリダイズし、配列番号 2 の部分は、配列番号 2 のヌクレオチド 53575 ~ 53577 を含む。いくつかのそのような実施形態では、単離された核酸分子は、少なくとも 20、少なくとも 25、または少なくとも 30 のヌクレオチドを含む。いくつかの実施形態では、本開示は、15 ~ 50 ヌクレオチドを含む単離された核酸分子を提供し、単離された核酸分子は、配列番号 2 の配列を含む核酸にハイブリダイズし、単離された核酸分子は、配列番号 2 の部分にハイブリダイズし、配列番号 2 の部分は、配列番号 2 のヌクレオチド 53575 ~ 53577 を含む。いくつかのそのような実施形態では、単離された核酸分子は、少なくとも 20、少なくとも 25、または少なくとも 30 のヌクレオチドを含む。

## 【0052】

いくつかの実施形態では、単離された核酸分子は、核酸の少なくとも 15 の連続ヌクレオチドにハイブリダイズし、連続ヌクレオチドは、配列番号 2 の部分と少なくとも 90% 同一であり、連続ヌクレオチドは、配列番号 2 の 53757 ~ 53577 位に対応する位置で、配列番号 2 のヌクレオチド 53575 ~ 53577 を含む。いくつかのそのような実施形態では、連続ヌクレオチドは、少なくとも 20、少なくとも 25、または少なくとも 30 ヌクレオチド長である。いくつかの実施形態では、単離された核酸分子は、核酸の少なくとも 15 の連続ヌクレオチドにハイブリダイズし、連続ヌクレオチドは、配列番号 2 の部分と少なくとも 95% 同一であり、連続ヌクレオチドは、配列番号 2 の 53757

10

20

30

40

50

~ 5 3 5 7 7 位に対応する位置で、配列番号 2 のヌクレオチド 5 3 5 7 5 ~ 5 3 5 7 7 を含む。いくつかのそのような実施形態では、連続ヌクレオチドは、少なくとも 2 0 、少なくとも 2 5 、または少なくとも 3 0 ヌクレオチド長である。いくつかの実施形態では、単離された核酸分子は、核酸の少なくとも 1 5 の連続ヌクレオチドとハイブリダイズし、連続ヌクレオチドは、配列番号 2 の部分と少なくとも 1 0 0 % 同一であり、連続ヌクレオチドは、配列番号 2 の 5 3 7 5 7 ~ 5 3 5 7 7 位に対応する位置で、配列番号 2 のヌクレオチド 5 3 5 7 5 ~ 5 3 5 7 7 を含む。いくつかのそのような実施形態では、連続ヌクレオチドは、少なくとも 2 0 、少なくとも 2 5 、または少なくとも 3 0 ヌクレオチド長である。

#### 【 0 0 5 3 】

いくつかの実施形態では、単離された核酸分子は、核酸の 1 5 ~ 5 0 の連続ヌクレオチドにハイブリダイズし、連続ヌクレオチドは、配列番号 2 の部分と少なくとも 9 0 % 同一であり、連続ヌクレオチドは、配列番号 2 の 5 3 7 5 7 ~ 5 3 5 7 7 位に対応する位置で、配列番号 2 のヌクレオチド 5 3 5 7 5 ~ 5 3 5 7 7 を含む。いくつかのそのような実施形態では、連続ヌクレオチドは、少なくとも 2 0 、少なくとも 2 5 、または少なくとも 3 0 ヌクレオチド長である。いくつかの実施形態では、単離された核酸分子は、核酸の 1 5 ~ 5 0 の連続ヌクレオチドにハイブリダイズし、連続ヌクレオチドは、配列番号 2 の部分と少なくとも 9 5 % 同一であり、連続ヌクレオチドは、配列番号 2 の 5 3 7 5 7 ~ 5 3 5 7 7 位に対応する位置で、配列番号 2 のヌクレオチド 5 3 5 7 5 ~ 5 3 5 7 7 を含む。いくつかのそのような実施形態では、連続ヌクレオチドは、少なくとも 2 0 、少なくとも 2 5 、または少なくとも 3 0 ヌクレオチド長である。いくつかの実施形態では、単離された核酸分子は、核酸の 1 5 ~ 5 0 の連続ヌクレオチドにハイブリダイズし、連続ヌクレオチドは、配列番号 2 の部分と少なくとも 1 0 0 % 同一であり、連続ヌクレオチドは、配列番号 2 の 5 3 7 5 7 ~ 5 3 5 7 7 位に対応する位置で、配列番号 2 のヌクレオチド 5 3 5 7 5 ~ 5 3 5 7 7 を含む。いくつかのそのような実施形態では、連続ヌクレオチドは、少なくとも 2 0 、少なくとも 2 5 、または少なくとも 3 0 ヌクレオチド長である。

#### 【 0 0 5 4 】

そのような単離された核酸分子は、例えば、ガイド R N A 、プライマー、プローブ、または外来性ドナー配列として使用することができる。

代表的な野生型 B 4 G A L T 1 ゲノム配列を、配列番号 1 に記載する。代表的なバリアント型 B 4 G A L T 1 ゲノム配列バリアントを、配列番号 2 に記載する。

#### 【 0 0 5 5 】

本開示はまた、 B 4 G A L T 1 m R N A のバリアントを含む単離された核酸分子を提供する。例示的な野生型ヒト B 4 G A L T 1 m R N A には、 N C B I 登録番号 N M \_ 0 0 1 4 9 7 ( 配列番号 3 ) が割り当てられており、これは 4 2 1 4 ヌクレオチド塩基からなる。ヒト B 4 G A L T 1 m R N A のバリアントは、配列番号 4 に示されており、 S N P ( 1 2 4 4 位の A から G へ、本明細書中ではバリアント型 B 4 G A L T 1 と呼ばれる ) を含んでおり、その結果、コードされた B 4 G A L T 1 バリアントポリペプチドの 3 5 2 位に対応する位置にセリンが生じる。バリアント型ヒト B 4 G A L T 1 m R N A は、例えれば、野生型ヒト B 4 G A L T 1 m R N A の 1 2 4 3 ~ 1 2 4 5 位に対応する位置に、野生型ヒト B 4 G A L T 1 m R N A の 1 2 4 3 ~ 1 2 4 5 位の 3 塩基「 a a u 」とは対照的に、セリンをコードする 3 塩基「 a g u 」を含む ( それぞれ配列番号 4 と配列番号 3 とを比較 ) 。いくつかの実施形態では、単離された核酸分子は配列番号 4 を含む。いくつかの実施形態では、単離された核酸分子は配列番号 4 からなる。

#### 【 0 0 5 6 】

いくつかの実施形態では、単離された核酸分子は、配列番号 4 と、少なくとも約 7 0 % 、少なくとも約 7 5 % 、少なくとも約 8 0 % 、少なくとも約 8 5 % 、少なくとも約 9 0 % 、少なくとも約 9 5 % 、少なくとも約 9 6 % 、少なくとも約 9 7 % 、少なくとも約 9 8 % 、少なくとも約 9 9 % 、または 1 0 0 % 同一である核酸配列を含むかまたはそれからなる。いくつかの実施形態では、そのような核酸配列はまた、配列番号 4 の 1 2 4 3 ~ 1 2 4 5 位に対応するヌクレオチドを含む。いくつかの実施形態では、単離された核酸分子は、

10

20

30

40

50

エクソン1～6を含む配列番号4の部分と、少なくとも約70%、少なくとも約75%、少なくとも約80%、少なくとも約85%、少なくとも約90%、少なくとも約95%、少なくとも約96%、少なくとも約97%、少なくとも約98%、少なくとも約99%、または100%同一であるヌクレオチド配列を含むかまたはそれからなる。いくつかの実施形態では、そのような核酸配列はまた、配列番号4の1243～1245位に対応するヌクレオチドを含む。いくつかの実施形態では、単離された核酸分子は、本明細書に開示する任意のB4GALT1 mRNA分子の相補体である。

#### 【0057】

いくつかの実施形態では、単離された核酸分子は、mRNA配列全体よりも短い配列を含む。いくつかの実施形態では、単離された核酸分子は、配列番号4の、少なくとも約15、少なくとも約20、少なくとも約25、少なくとも約30、少なくとも約35、少なくとも約40、少なくとも約45、少なくとも約50、少なくとも約60、少なくとも約70、少なくとも約80、少なくとも約90、少なくとも約100、少なくとも約200、少なくとも約300、少なくとも約400、少なくとも約500、少なくとも約600、少なくとも約700、少なくとも約800、少なくとも約900、少なくとも約1000、少なくとも約2000、少なくとも約3000、または少なくとも約4000の連続ヌクレオチドを含むかまたはそれからなる。いくつかの実施形態では、そのような単離された核酸分子はまた、配列番号4の1243～1245位に対応するヌクレオチドを含む。いくつかの実施形態では、単離された核酸分子は、配列番号4の、少なくとも約15、少なくとも約20、少なくとも約25、少なくとも約30、少なくとも約35、少なくとも約40、少なくとも約45、少なくとも約50、少なくとも約60、少なくとも約70、少なくとも約80、少なくとも約90、少なくとも約100、少なくとも約200、少なくとも約300、少なくとも約400、少なくとも約500、少なくとも約600、少なくとも約700、少なくとも約800、少なくとも約900、または少なくとも約1000の連続ヌクレオチドを含むかまたはそれからなる。いくつかの実施形態では、そのような単離された核酸分子はまた、配列番号4の1243～1245位に対応するヌクレオチドを含む。いくつかの実施形態では、単離された核酸分子は、配列番号4のエクソン1～6の、少なくとも約15、少なくとも約20、少なくとも約25、少なくとも約30、少なくとも約35、少なくとも約40、少なくとも約45、少なくとも約50、少なくとも約60、少なくとも約70、少なくとも約80、少なくとも約90、少なくとも約100、少なくとも約200、少なくとも約300、少なくとも約400、少なくとも約500、少なくとも約600、少なくとも約700、少なくとも約800、少なくとも約900、または少なくとも約1000の連続ヌクレオチドを含むかまたはそれからなる。いくつかの実施形態では、そのような単離された核酸分子はまた、配列番号4の1243～1245位に対応するヌクレオチドを含む。

#### 【0058】

いくつかの実施形態では、本開示は、配列番号4の部分と少なくとも90%同一である核酸配列を含む単離された核酸分子を提供し、配列番号4の部分は、配列番号4のヌクレオチド1243～1245を含み、配列番号4の部分は、配列番号4の少なくとも15ヌクレオチドを含む。いくつかのそのような実施形態では、配列番号4の部分は、配列番号4の、少なくとも20、少なくとも25または少なくとも30ヌクレオチドである。いくつかの実施形態では、本開示は、配列番号4の部分と少なくとも95%同一である核酸配列を含む単離された核酸分子を提供し、配列番号4の部分は、配列番号4のヌクレオチド1243～1245を含み、配列番号4の部分は、配列番号4の少なくとも15ヌクレオチドを含む。いくつかのそのような実施形態では、配列番号4の部分は、配列番号4の、少なくとも20、少なくとも25または少なくとも30ヌクレオチドである。いくつかの実施形態では、本開示は、配列番号4の部分と100%同一である核酸配列を含む単離された核酸分子を提供し、配列番号4の部分は、配列番号4のヌクレオチド1243～1245を含み、配列番号4の部分は、配列番号4の少なくとも15ヌクレオチドを含む。いくつかのそのような実施形態では、配列番号4の部分は、配列番号4の、少なくとも20

10

20

30

40

50

、少なくとも 25 または少なくとも 30 ヌクレオチドである。いくつかの実施形態では、本開示は、配列番号 4 の部分と少なくとも 90 % 同一である核酸配列を含む単離された核酸分子を提供し、配列番号 4 の部分は、配列番号 4 のヌクレオチド 1243 ~ 1245 を含み、配列番号 4 の部分は、配列番号 4 の 15 ~ 50 ヌクレオチドを含む。いくつかのそのような実施形態では、配列番号 4 の部分は、配列番号 4 の、少なくとも 20 、少なくとも 25 または少なくとも 30 ヌクレオチドである。いくつかの実施形態では、本開示は、配列番号 4 の部分と少なくとも 95 % 同一である核酸配列を含む単離された核酸分子を提供し、配列番号 4 の部分は、配列番号 4 のヌクレオチド 1243 ~ 1245 を含み、配列番号 4 の部分は、配列番号 4 の 15 ~ 50 ヌクレオチドを含む。いくつかのそのような実施形態では、配列番号 4 の部分は、配列番号 4 の、少なくとも 20 、少なくとも 25 または少なくとも 30 ヌクレオチドである。いくつかの実施形態では、本開示は、配列番号 4 の部分と 100 % 同一である核酸配列を含む単離された核酸分子を提供し、配列番号 4 の部分は、配列番号 4 のヌクレオチド 1243 ~ 1245 を含み、配列番号 4 の部分は、配列番号 4 の 15 ~ 50 ヌクレオチドを含む。いくつかのそのような実施形態では、配列番号 4 の部分は、配列番号 4 の、少なくとも 20 、少なくとも 25 または少なくとも 30 ヌクレオチドである。

#### 【 0059 】

そのような単離された核酸分子は、例えば、B4GALT1 バリアントポリペプチドを発現するために、または外来性ドナー配列として使用することができる。集団内の遺伝子配列は、SNP などの多型により変化し得ることが理解される。本明細書において提供する例は、例示的な配列に過ぎず、他の配列も可能である。

#### 【 0060 】

いくつかの実施形態では、単離された核酸分子は、バリアント型 Asn352Ser B4GALT1 ポリペプチド ( 配列番号 8 ) と、少なくとも約 75 % 、少なくとも約 80 % 、少なくとも約 85 % 、少なくとも約 90 % 、少なくとも約 91 % 、少なくとも約 92 % 、少なくとも約 93 % 、少なくとも約 94 % 、少なくとも約 95 % 、少なくとも約 96 % 、少なくとも約 97 % 、少なくとも約 98 % 、少なくとも約 99 % 、または 100 % 同一であるポリペプチドをコードする核酸配列を含むかまたはそれからなり、ただし、そのポリペプチドは、352 位に対応する位置にセリンを含む。いくつかの実施形態では、単離された核酸分子は、配列番号 8 と少なくとも約 90 % 同一のポリペプチドをコードする核酸配列を含むかまたはそれからなり、ただし、そのポリペプチドは、352 位に対応する位置にセリンを含む。いくつかの実施形態では、単離された核酸分子は、配列番号 8 と少なくとも約 95 % 同一のポリペプチドをコードする核酸配列を含むかまたはそれからなり、ただし、そのポリペプチドは、352 位に対応する位置にセリンを含む。

#### 【 0061 】

例えば、いくつかの実施形態では、単離された核酸分子は、少なくとも 10 アミノ酸長のアミノ酸配列を有するポリペプチドをコードする核酸配列を含み、そのアミノ酸配列は、配列番号 8 のアミノ酸配列の部分と 90 % 同一であり、その部分は、配列番号 8 の 352 位に対応する位置にセリンを含む。いくつかのそのような実施形態では、核酸配列は、少なくとも 15 アミノ酸、少なくとも 20 アミノ酸または少なくとも 25 アミノ酸長のアミノ酸配列を有するポリペプチドをコードする。いくつかの実施形態では、単離された核酸分子は、少なくとも 10 アミノ酸長のアミノ酸配列を有するポリペプチドをコードする核酸配列を含み、アミノ酸配列は、配列番号 8 のアミノ酸配列の部分と 95 % 同一であり、その部分は、配列番号 8 の 352 位に対応する位置にセリンを含む。いくつかのそのような実施形態では、核酸配列は、少なくとも 15 アミノ酸、少なくとも 20 アミノ酸または少なくとも 25 アミノ酸長のアミノ酸配列を有するポリペプチドをコードする。いくつかの実施形態では、単離された核酸分子は、10 ~ 50 アミノ酸長のアミノ酸配列を有するポリペプチドをコードする核酸配列を含み、アミノ酸配列は、配列番号 8 のアミノ酸配列の部分と 90 % 同一であり、その部分は、配列番号 8 の 352 位に対応する位置にセリンを含む。いくつかのそのような実施形態では、核酸配列は、少なくとも 15 アミノ酸、

少なくとも 20 アミノ酸または少なくとも 25 アミノ酸長のアミノ酸配列を有するポリペプチドをコードする。いくつかの実施形態では、単離された核酸分子は、10～50 アミノ酸長のアミノ酸配列を有するポリペプチドをコードする核酸配列を含み、アミノ酸配列は、配列番号 8 のアミノ酸配列の部分と 95% 同一であり、その部分は、配列番号 8 の 352 位に対応する位置にセリンを含む。いくつかのそのような実施形態では、核酸配列は、少なくとも 15 アミノ酸、少なくとも 20 アミノ酸または少なくとも 25 アミノ酸長のアミノ酸配列を有するポリペプチドをコードする。いくつかの実施形態では、単離された核酸分子は、配列番号 8 と同一のポリペプチドをコードする核酸配列を含むかまたはそれからなる。

【0062】

本開示はまた、バリアント型 B4GALT1 mRNA 配列にハイブリダイズする単離された核酸分子を提供する。いくつかの実施形態では、そのような単離された核酸分子は、少なくとも約 15、少なくとも約 20、少なくとも約 25、少なくとも約 30、少なくとも約 35、少なくとも約 40、少なくとも約 45、少なくとも約 50、少なくとも約 60、少なくとも約 70、少なくとも約 80、少なくとも約 90、少なくとも約 100、少なくとも約 200、少なくとも約 300、少なくとも約 400、少なくとも約 500、少なくとも約 600、少なくとも約 700、少なくとも約 800、少なくとも約 900、少なくとも約 1000、少なくとも約 2000、少なくとも約 3000、または少なくとも約 4000 ヌクレオチドを含むかまたはそれからなる。いくつかの実施形態では、そのような単離された核酸分子はまた、配列番号 4 の 1243～1245 位にハイブリダイズする。いくつかの実施形態では、単離された核酸分子は、配列番号 4 の 1243～1245 位の、約 1000、約 500、約 400、約 300、約 200、約 100、約 50、約 45、約 40、約 35、約 30、約 25、約 20、約 15、約 10、または約 5 ヌクレオチドを含むかまたはそれ以内のセグメントにおいてバリアント型 B4GALT1 mRNA の部分にハイブリダイズする。

【0063】

いくつかの実施形態では、単離された核酸分子は、少なくとも 15 ヌクレオチドを含むかまたはそれからなり、配列番号 4 の 1243～1245 位の 5 ヌクレオチドを含むかまたはそれ以内のセグメントにおいてバリアント型 B4GALT1 mRNA ( 例えば、配列番号 4 ) の部分にハイブリダイズする。いくつかのそのような実施形態では、単離された核酸分子は、少なくとも 20、少なくとも 25 または少なくとも 30 ヌクレオチドを含む。いくつかの実施形態では、単離された核酸分子は、少なくとも 15 ヌクレオチドを含むかまたはそれからなり、配列番号 4 の 1243～1245 位の 5 ヌクレオチドを含むかまたはそれ以内のセグメントにおいてバリアント型 B4GALT1 mRNA ( 例えば、配列番号 4 ) の部分にハイブリダイズし、配列番号 4 の 1243～1245 位にハイブリダイズする。いくつかのそのような実施形態では、単離された核酸分子は、少なくとも 20、少なくとも 25 または少なくとも 30 ヌクレオチドを含む。いくつかの実施形態では、単離された核酸分子は、15～50 ヌクレオチドを含み、配列番号 4 の 1243～1245 位を含むセグメントにおいてバリアント型 B4GALT1 mRNA ( 例えば、配列番号 4 ) の部分にハイブリダイズし、配列番号 4 の 1243～1245 位にハイブリダイズする。いくつかのそのような実施形態では、単離された核酸分子は、少なくとも 20、少なくとも 25 または少なくとも 30 ヌクレオチドを含む。

【0064】

いくつかの実施形態では、単離された核酸分子は、バリアント型 B4GALT1 mRNA ( 例えば、配列番号 4 ) と、少なくとも約 70%、少なくとも約 75%、少なくとも約 80%、少なくとも約 85%、少なくとも約 90%、少なくとも約 95%、少なくとも約 96%、少なくとも約 97%、少なくとも約 98%、少なくとも約 99%、または 100% 同一である核酸分子の少なくとも約 15 の連続ヌクレオチドにハイブリダイズする。いくつかの実施形態では、単離された核酸分子はまた、配列番号 4 の 1243～1245 位にハイブリダイズする。いくつかの実施形態では、単離された核酸分子は、約 15～約

10

20

30

40

50

100 ヌクレオチド、または約 15 ~ 約 35 ヌクレオチドを含むか、またはそれからなる。

【0065】

いくつかの実施形態では、単離された核酸分子は、少なくとも 15 ヌクレオチドを含むかまたはそれからなり、配列番号 4 の 1243 ~ 1245 位の 5 ヌクレオチドを含むかまたはそれ以内のセグメントにおいてバリアント型 B4GALT1 mRNA の部分にハイブリダイズし、その場合、バリアント型 B4GALT1 mRNA は、バリアント型 B4GALT1 mRNA ( 例えれば、配列番号 4 など ) と少なくとも 90 % 同一である。いくつかのそのような実施形態では、単離された核酸分子は、少なくとも 20 、少なくとも 25 または少なくとも 30 ヌクレオチドを含む。いくつかの実施形態では、単離された核酸分子は、少なくとも 15 ヌクレオチドを含むかまたはそれからなり、配列番号 4 の 1243 ~ 1245 位の 5 ヌクレオチドを含むかまたはそれ以内のセグメントにおいてバリアント型 B4GALT1 mRNA の部分にハイブリダイズし、その場合、バリアント型 B4GALT1 mRNA は、バリアント型 B4GALT1 mRNA ( 例えれば、配列番号 4 など ) と少なくとも 95 % 同一である。いくつかのそのような実施形態では、単離された核酸分子は、少なくとも 20 、少なくとも 25 または少なくとも 30 ヌクレオチドを含む。いくつかの実施形態では、単離された核酸分子は、少なくとも 15 ヌクレオチドを含むかまたはそれからなり、配列番号 4 の 1243 ~ 1245 位の 5 ヌクレオチドを含むかまたはそれ以内のセグメントにおいてバリアント型 B4GALT1 mRNA の部分にハイブリダイズし、配列番号 4 の 1243 ~ 1245 位にハイブリダイズし、その場合、バリアント型 B4GALT1 mRNA は、バリアント型 B4GALT1 mRNA ( 例えれば、配列番号 4 など ) と少なくとも 90 % 同一である。いくつかのそのような実施形態では、単離された核酸分子は、少なくとも 20 、少なくとも 25 または少なくとも 30 ヌクレオチドを含む。いくつかの実施形態では、単離された核酸分子は、少なくとも 15 ヌクレオチドを含むかまたはそれからなり、配列番号 4 の 1243 ~ 1245 位の 5 ヌクレオチドを含むかまたはそれ以内のセグメントにおいてバリアント型 B4GALT1 mRNA の部分にハイブリダイズし、配列番号 4 の 1243 ~ 1245 位にハイブリダイズし、その場合、バリアント型 B4GALT1 mRNA は、バリアント型 B4GALT1 mRNA ( 例えれば、配列番号 4 など ) と少なくとも 95 % 同一である。いくつかのそのような実施形態では、単離された核酸分子は、少なくとも 20 、少なくとも 25 または少なくとも 30 ヌクレオチドを含む。いくつかの実施形態では、単離された核酸分子は、15 ~ 100 ヌクレオチド、または 15 ~ 35 ヌクレオチドを含むか、またはそれからなる。

【0066】

そのような単離された核酸分子は、例えれば、ガイド RNA 、プライマー、プローブ、または外来性ドナー配列として使用することができる。

代表的な野生型 B4GALT1 mRNA 配列を、配列番号 3 に記載する。代表的なバリアント型 B4GALT1 mRNA 配列を、配列番号 4 に記載する。

【0067】

本開示はまた、 B4GALT1 バリアントポリペプチドの全部または一部をコードする B4GALT1 cDNA のバリアントを含む核酸分子を提供する。例示的な野生型ヒト B4GALT1 cDNA ( 例えれば、 DNA として記述された mRNA のコード領域 ) は、 1197 ヌクレオチド塩基からなる ( 配列番号 5 ) 。ヒト B4GALT1 cDNA のバリアントは配列番号 6 に示されており、このバリアントは SNP ( 1055 位の A から G へ、本明細書中ではバリアント型 B4GALT1 と呼ばれる ) を含んでおり、その結果、コードされた B4GALT1 バリアントポリペプチドの 352 位に対応する位置にセリンが生じる。バリアント型ヒト B4GALT1 cDNA は、例えれば、完全長 / 成熟型の野生型ヒト B4GALT1 cDNA の 1054 ~ 1056 位における 3 つの塩基「 a a t 」とは対照的に、野生型ヒト B4GALT1 cDNA の 1054 ~ 1056 位に対応する位置にセリンをコードする「 a g t 」を含む ( それぞれ、配列番号 6 と配列番号 5 とを比較 ) 。いくつかの実施形態では、核酸分子は配列番号 6 を含む。いくつかの実施形態

10

20

30

40

50

では、核酸分子は配列番号 6 からなる。いくつかの実施形態では、c DNA 分子は単離される。

【0068】

いくつかの実施形態では、c DNA 分子は、配列番号 6 と、少なくとも約 70%、少なくとも約 75%、少なくとも約 80%、少なくとも約 85%、少なくとも約 90%、少なくとも約 95%、少なくとも約 96%、少なくとも約 97%、少なくとも約 98%、少なくとも約 99%、または 100% 同一である核酸配列を含むかまたはそれからなる。いくつかの実施形態では、c DNA 分子はまた、配列番号 6 の 1054 ~ 1056 位に対応するヌクレオチドを含む。いくつかの実施形態では、単離された核酸分子は、本明細書に開示する任意の B4GALT1 c DNA 分子の相補体である。

10

【0069】

いくつかの実施形態では、c DNA 分子は、c DNA 配列全体よりも短い配列を含む。いくつかの実施形態では、c DNA 分子は、配列番号 6 の、少なくとも約 15、少なくとも約 20、少なくとも約 25、少なくとも約 30、少なくとも約 35、少なくとも約 40、少なくとも約 45、少なくとも約 50、少なくとも約 60、少なくとも約 70、少なくとも約 80、少なくとも約 90、少なくとも約 100、少なくとも約 200、少なくとも約 300、少なくとも約 400、少なくとも約 500、少なくとも約 600、少なくとも約 700、少なくとも約 800、少なくとも約 900、少なくとも約 1000、または少なくとも約 1100 の連続ヌクレオチドを含むかまたはそれからなる。いくつかの実施形態では、そのような c DNA 分子はまた、配列番号 6 の 1054 ~ 1056 位に対応するヌクレオチドを含む。いくつかの実施形態では、c DNA 分子は、配列番号 6 の、少なくとも約 15、少なくとも約 20、少なくとも約 25、少なくとも約 30、少なくとも約 35、少なくとも約 40、少なくとも約 45、少なくとも約 50、少なくとも約 60、少なくとも約 70、少なくとも約 80、少なくとも約 90、少なくとも約 100、少なくとも約 200、少なくとも約 300、少なくとも約 400、または少なくとも約 500 の連続ヌクレオチドを含むかまたはそれからなる。いくつかの実施形態では、そのような c DNA 分子はまた、配列番号 6 の 1054 ~ 1056 位に対応するヌクレオチドを含む。

20

【0070】

例えば、いくつかの実施形態では、c DNA 分子は、配列番号 6 の少なくとも 15 の連続ヌクレオチドを含み、連続ヌクレオチドは、配列番号 6 のヌクレオチド 1054 ~ 1056 を含む。いくつかのそのような実施形態では、単離された核酸分子は、配列番号 6 の、少なくとも 20、少なくとも 25 または少なくとも 30 の連続ヌクレオチドを含む。いくつかの実施形態では、c DNA 分子は、配列番号 6 の 15 ~ 50 の連続ヌクレオチドを含み、連続ヌクレオチドは、配列番号 6 のヌクレオチド 1054 ~ 1056 を含む。いくつかのそのような実施形態では、単離された核酸分子は、配列番号 6 の、少なくとも 20、少なくとも 25 または少なくとも 30 の連続ヌクレオチドを含む。いくつかの実施形態では、本開示は、配列番号 6 の部分と少なくとも 90% 同一である核酸配列を含む c DNA 分子を提供し、配列番号 6 の部分は、配列番号 6 のヌクレオチド 1054 ~ 1056 を含み、配列番号 6 の部分は、配列番号 6 の少なくとも 15 の連続ヌクレオチドを含む。いくつかのそのような実施形態では、配列番号 6 の部分は、配列番号 6 の、少なくとも 20、少なくとも 25 または少なくとも 30 の連続ヌクレオチドである。いくつかの実施形態では、本開示は、配列番号 6 の部分と少なくとも 95% 同一である核酸配列を含む c DNA 分子を提供し、配列番号 6 の部分は、配列番号 6 のヌクレオチド 1054 ~ 1056 を含み、配列番号 6 の部分は、配列番号 6 の少なくとも 15 の連続ヌクレオチドを含む。いくつかのそのような実施形態では、配列番号 6 の部分は、配列番号 6 の、少なくとも 20、少なくとも 25 または少なくとも 30 の連続ヌクレオチドである。いくつかの実施形態では、本開示は、配列番号 6 の部分と少なくとも 90% 同一である核酸配列を含む c DNA 分子を提供し、配列番号 6 の部分は、配列番号 6 のヌクレオチド 1054 ~ 1056 を含み、配列番号 6 の部分は、配列番号 6 の 15 ~ 50 の連続ヌクレオチドを含む。いくつかのそのような実施形態では、配列番号 6 の部分は、配列番号 6 の、少なくとも 20、少

30

40

50

なくとも 25 または少なくとも 30 の連続ヌクレオチドである。いくつかの実施形態では、本開示は、配列番号 6 の部分と少なくとも 95 % 同一である核酸配列を含む cDNA 分子を提供し、配列番号 6 の部分は、配列番号 6 のヌクレオチド 1054 ~ 1056 を含み、配列番号 6 の部分は、配列番号 6 の 15 ~ 50 の連続ヌクレオチドを含む。いくつかのそのような実施形態では、配列番号 6 の部分は、配列番号 6 の、少なくとも 20、少なくとも 25 または少なくとも 30 の連続ヌクレオチドである。いくつかの実施形態では、本開示は、配列番号 6 のヌクレオチド 1054 ~ 1056 に対応する位置に配列番号 6 のヌクレオチド 1054 ~ 1056 を含む cDNA 分子を提供し、cDNA 分子は、配列番号 6 の部分と少なくとも 90 % 同一の核酸配列を含み、配列番号 6 の部分は、配列番号 6 のヌクレオチド 1054 ~ 1056 を含み、配列番号 6 の部分は、配列番号 6 の少なくとも 15 の連続ヌクレオチドを含む。いくつかのそのような実施形態では、配列番号 6 の部分は、配列番号 6 の、少なくとも 20、少なくとも 25 または少なくとも 30 の連続ヌクレオチドである。いくつかの実施形態では、本開示は、配列番号 6 のヌクレオチド 1054 ~ 1056 に対応する位置に配列番号 6 のヌクレオチド 1054 ~ 1056 を含む cDNA 分子を提供し、cDNA 分子は、配列番号 6 の部分と少なくとも 95 % 同一の核酸配列を含み、配列番号 6 の部分は、配列番号 6 のヌクレオチド 1054 ~ 1056 を含み、配列番号 6 の部分は、配列番号 6 の少なくとも 15 の連続ヌクレオチドを含む。いくつかのそのような実施形態では、配列番号 6 の部分は、配列番号 6 の、少なくとも 20、少なくとも 25 または少なくとも 30 の連続ヌクレオチドである。いくつかの実施形態では、本開示は、配列番号 6 のヌクレオチド 1054 ~ 1056 に対応する位置に配列番号 6 のヌクレオチド 1054 ~ 1056 を含む cDNA 分子を提供し、cDNA 分子は、配列番号 6 の部分と少なくとも 90 % 同一の核酸配列を含み、配列番号 6 の部分は、配列番号 6 のヌクレオチド 1054 ~ 1056 を含み、配列番号 6 の部分は、配列番号 6 の 15 ~ 50 の連続ヌクレオチドを含む。いくつかのそのような実施形態では、配列番号 6 の部分は、配列番号 6 の、少なくとも 20、少なくとも 25 または少なくとも 30 の連続ヌクレオチドである。 10

いくつかの実施形態では、本開示は、配列番号 6 のヌクレオチド 1054 ~ 1056 に対応する位置に配列番号 6 のヌクレオチド 1054 ~ 1056 を含む cDNA 分子を提供し、cDNA 分子は、配列番号 6 の部分と少なくとも 95 % 同一の核酸配列を含み、配列番号 6 の部分は、配列番号 6 のヌクレオチド 1054 ~ 1056 を含み、配列番号 6 の部分は、配列番号 6 の 15 ~ 50 の連続ヌクレオチドを含む。いくつかのそのような実施形態では、配列番号 6 の部分は、配列番号 6 の、少なくとも 20、少なくとも 25 または少なくとも 30 の連続ヌクレオチドである。 20

いくつかの実施形態では、本開示は、配列番号 6 のヌクレオチド 1054 ~ 1056 に対応する位置に配列番号 6 のヌクレオチド 1054 ~ 1056 を含む cDNA 分子を提供し、cDNA 分子は、配列番号 6 の部分と少なくとも 90 % 同一の核酸配列を含み、配列番号 6 の部分は、配列番号 6 のヌクレオチド 1054 ~ 1056 を含み、配列番号 6 の部分は、配列番号 6 の 15 ~ 50 の連続ヌクレオチドを含む。いくつかのそのような実施形態では、配列番号 6 の部分は、配列番号 6 の、少なくとも 20、少なくとも 25 または少なくとも 30 の連続ヌクレオチドである。 30

いくつかの実施形態では、本開示は、配列番号 6 のヌクレオチド 1054 ~ 1056 に対応する位置に配列番号 6 のヌクレオチド 1054 ~ 1056 を含む cDNA 分子を提供し、cDNA 分子は、配列番号 6 の部分と少なくとも 95 % 同一の核酸配列を含み、配列番号 6 の部分は、配列番号 6 のヌクレオチド 1054 ~ 1056 を含み、配列番号 6 の部分は、配列番号 6 の 15 ~ 50 の連続ヌクレオチドを含む。いくつかのそのような実施形態では、配列番号 6 の部分は、配列番号 6 の、少なくとも 20、少なくとも 25 または少なくとも 30 の連続ヌクレオチドである。

### 【0071】

そのような cDNA 分子は、例えば、B4GALT1 バリアントタンパク質を発現するために、または外来性ドナー配列として使用することができる。集団内の遺伝子配列は、SNP などの多型により変化し得ることが理解される。本明細書において提供する例は、例示的な配列に過ぎず、他の配列も可能である。

### 【0072】

いくつかの実施形態では、cDNA 分子は、バリアント型 Asn352Ser B4GALT1 ポリペプチド（配列番号 8）と、少なくとも約 75 %、少なくとも約 80 %、少なくとも約 85 %、少なくとも約 90 %、少なくとも約 91 %、少なくとも約 92 %、少なくとも約 93 %、少なくとも約 94 %、少なくとも約 95 %、少なくとも約 96 %、少なくとも約 97 %、少なくとも約 98 %、少なくとも約 99 %、または 100 % 同一のポリペプチドをコードする核酸配列を含むかまたはそれからなり、そのポリペプチドは 352 位に対応する位置にセリンを含むことを条件とする。いくつかの実施形態では、cDNA 分子は、配列番号 8 と少なくとも約 90 % 同一のポリペプチドをコードする核酸配列を含むかまたはそれからなり、そのポリペプチドは、352 位に対応する位置にセリンを含むことを条件とする。いくつかの実施形態では、cDNA 分子は、配列番号 8 と、少なくとも約 95 % 同一のポリペプチドをコードする核酸配列を含むかまたはそれからなり、そのポリペプチドは、352 位に対応する位置にセリンを含むことを条件とする。いくつかの 40

10

20

30

40

50

実施形態では、cDNA分子は、配列番号8と同一のポリペプチドをコードする核酸配列を含むかまたはそれからなる。

【 0 0 7 3 】

本開示はまた、バリアント型 B 4 G A L T 1 c D N A 配列にハイブリダイズする単離された核酸分子を提供する。いくつかの実施形態では、そのような単離された核酸分子は、少なくとも約 15、少なくとも約 20、少なくとも約 25、少なくとも約 30、少なくとも約 35、少なくとも約 40、少なくとも約 45、少なくとも約 50、少なくとも約 60、少なくとも約 70、少なくとも約 80、少なくとも約 90、少なくとも約 100、少なくとも約 200、少なくとも約 300、少なくとも約 400、少なくとも約 500、少なくとも約 600、少なくとも約 700、少なくとも約 800、少なくとも約 900、少なくとも約 1000、または少なくとも約 1100 ヌクレオチドを含むかまたはそれからなる。いくつかの実施形態では、そのような単離された核酸分子はまた、配列番号 6 の 1054 ~ 1056 位にハイブリダイズする。いくつかの実施形態では、そのような単離された核酸分子は、配列番号 6 の 1054 ~ 1056 位の約 600、約 500、約 400、約 300、約 200、約 100、約 50、約 45、約 40、約 35、約 30、約 25、約 20、約 15、約 10、または約 5 ヌクレオチドを含むかまたはそれ以内のセグメントにおいてバリアント型 B 4 G A L T 1 c D N A の部分にハイブリダイズする。いくつかの実施形態では、単離された核酸分子は、バリアント型 B 4 G A L T 1 c D N A ( 例えば、配列番号 6 ) と、少なくとも約 70%、少なくとも約 75%、少なくとも約 80%、少なくとも約 85%、少なくとも約 90%、少なくとも約 95%、少なくとも約 96%、少なくとも約 97%、少なくとも約 98%、少なくとも約 99%、または 100% 同一である c D N A 分子の少なくとも約 15 の連続ヌクレオチドにハイブリダイズする。いくつかの実施形態では、単離された核酸分子はまた、配列番号 6 の 1054 ~ 1056 位にハイブリダイズする。いくつかの実施形態では、単離された核酸分子は、約 15 ~ 約 100 ヌクレオチド、または約 15 ~ 約 35 ヌクレオチドを含むか、またはそれからなる。

〔 0 0 7 4 〕

いくつかの実施形態では、単離された核酸分子は、少なくとも 15 ヌクレオチドを含むかまたはそれらからなり、配列番号 6 の 1054 ~ 1056 位の 5 ヌクレオチドを含むかまたはそれ以内のセグメントにおいてバリアント型 B4GALT1 cDNA の部分にハイブリダイズし、その場合、バリアント型 B4GALT1 cDNA は、バリアント型 B4GALT1 cDNA ( 例えば、配列番号 6 など ) と少なくとも 90 % 同一である。いくつかの実施形態では、単離された核酸分子は、少なくとも 15 ヌクレオチドを含むかまたはそれらからなり、配列番号 6 の 1054 ~ 1056 位の 5 ヌクレオチドを含むかまたはそれ以内のセグメントにおいてバリアント型 B4GALT1 cDNA の部分にハイブリダイズし、その場合、バリアント型 B4GALT1 cDNA は、バリアント型 B4GALT1 cDNA ( 例えば、配列番号 6 など ) と少なくとも 95 % 同一である。いくつかの実施形態では、単離された核酸分子は、少なくとも 15 ヌクレオチドを含むかまたはそれらからなり、配列番号 6 の 1054 ~ 1056 位の 5 ヌクレオチドを含むかまたはそれ以内のセグメントにおいてバリアント型 B4GALT1 cDNA の部分にハイブリダイズし、その場合、バリアント型 B4GALT1 cDNA は、バリアント型 B4GALT1 cDNA ( 例えば、配列番号 6 など ) と 100 % 同一である。いくつかの実施形態では、単離された核酸分子は、少なくとも 15 ヌクレオチドを含むかまたはそれらからなり、配列番号 6 の 1054 ~ 1056 位の 5 ヌクレオチドを含むかまたはそれ以内のセグメントにおいてバリアント型 B4GALT1 cDNA の部分にハイブリダイズし、配列番号 6 の 1054 ~ 1056 位にハイブリダイズし、その場合、バリアント型 B4GALT1 cDNA は、バリアント型 B4GALT1 cDNA ( 例えば、配列番号 6 など ) と少なくとも 90 % 同一である。いくつかの実施形態では、単離された核酸分子は、少なくとも 15 ヌクレオチドを含むかまたはそれらからなり、配列番号 6 の 1054 ~ 1056 位の 5 ヌクレオチドを含むかまたはそれ以内のセグメントにおいてバリアント型 B4GALT1 cDNA の部分にハイブリダイズし、配列番号 6 の 1054 ~ 1056 位にハイブリダイズし、その場合、バリアント型 B4GALT1 cDNA は、バリアント型 B4GALT1 cDNA ( 例えば、配列番号 6 など ) と少なくとも 90 % 同一である。

プリダイズし、その場合、バリアント型 B 4 G A L T 1 c D N A は、バリアント型 B 4 G A L T 1 c D N A ( 例えは、配列番号 6 など ) と少なくとも 95 % 同一である。いくつかの実施形態では、単離された核酸分子は、少なくとも 15 ヌクレオチドを含むかまたはそれからなり、配列番号 6 の 1054 ~ 1056 位の 5 ヌクレオチドを含むかまたはそれ以内のセグメントにおいてバリアント型 B 4 G A L T 1 c D N A の部分にハイブリダイズし、配列番号 6 の 1054 ~ 1056 位にハイブリダイズし、その場合、バリアント型 B 4 G A L T 1 c D N A は、バリアント型 B 4 G A L T 1 c D N A ( 例えは、配列番号 6 など ) と 100 % 同一である。いくつかの実施形態では、単離された核酸分子は、15 ~ 100 ヌクレオチド、または 15 ~ 35 ヌクレオチドを含むか、またはそれからなる。

10

#### 【 0075 】

そのような単離された核酸分子は、例えは、ガイド R N A 、プライマー、プローブ、外来性ドナー配列、アンチセンス R N A 、 s i R N A 、または s h R N A として使用することができる。

#### 【 0076 】

代表的な野生型 B 4 G A L T 1 c D N A 配列を、配列番号 5 に記載する。代表的なバリアント型 B 4 G A L T 1 c D N A 配列を、配列番号 6 に記載する。

本明細書に開示する核酸分子には、天然の B 4 G A L T 1 遺伝子または m R N A 転写産物の核酸配列を含ませることができ、または非天然の配列を含ませることができる。いくつかの実施形態では、天然の配列は、同義的変異またはコードする B 4 G A L T 1 ポリペプチドに影響を及ぼさない変異のために、非天然の配列と異なり得る。例えは、配列は、同義的変異またはコードする B 4 G A L T 1 ポリペプチドに影響を及ぼさない変異を除いて同一であり得る。同義的変異または置換とは、生成されるアミノ酸配列を改変しないような、タンパク質をコードする遺伝子のエクソンにおけるあるヌクレオチドから別のヌクレオチドへの置換である。これは、遺伝暗号の縮重により可能となっており、いくつかのアミノ酸は複数の 3 塩基対コドンによってコードされる。同義的置換は、例えは、コドン最適化のプロセスで使用される。本明細書に開示する核酸分子は、コドン最適化することができる。

20

#### 【 0077 】

本明細書ではまた、開示する核酸分子と相互作用することができる機能性ポリヌクレオチドを提供する。機能性ポリヌクレオチドは、標的分子の結合または特異的反応の触媒など、特定の機能を有する核酸分子である。機能性ポリヌクレオチドの例として、アンチセンス分子、アプタマー、リボザイム、三重鎖形成分子、及び外部ガイド配列が挙げられるが、これらに限定されない。機能性ポリヌクレオチドは、標的分子が有する特異的活性のエフェクター、阻害剤、モジュレーター、及び刺激剤として作用することができるか、または機能性ポリヌクレオチドは、他の分子とは独立した新規活性を有し得る。

30

#### 【 0078 】

アンチセンス分子は、標準的または非標準的な塩基対形成のいずれかを介して標的核酸分子と相互作用するように設計される。アンチセンス分子と標的分子の相互作用は、例えは、 R N a s e - H を介した R N A - D N A ハイブリッド分解による標的分子の破壊を促進するように設計される。あるいは、アンチセンス分子は、転写または複製などの、通常であれば標的分子で起こるであろう処理機能を妨げるよう設計される。アンチセンス分子は、標的分子の配列に基づいて設計することができる。標的分子の最もアクセスしやすい領域を特定することによってアンチセンス効率を最適化するための多くの方法が存在する。例示的な方法として、 i n v i t r o 選択実験及び D M S 及び D E P C を使用した D N A 修飾試験が挙げられるが、これらに限定されない。アンチセンス分子は、一般的に、解離定数 ( k d ) 約 10<sup>-6</sup> 以下、約 10<sup>-8</sup> 以下、約 10<sup>-10</sup> 以下、または約 10<sup>-1</sup> 以下で標的分子に結合する。アンチセンス分子の設計及び使用を支援する方法及び技術の代表的なサンプルは、以下の非限定的な米国特許のリスト : 5,135,917, 5,294,533, 5,627,158, 5,641,754, 5,691,317, 5,

40

50

780, 607, 5, 786, 138, 5, 849, 903, 5, 856, 103, 5,  
 919, 772, 5, 955, 590, 5, 990, 088, 5, 994, 320, 5,  
 998, 602, 6, 005, 095, 6, 007, 995, 6, 013, 522, 6,  
 017, 898, 6, 018, 042, 6, 025, 198, 6, 033, 910, 6,  
 040, 296, 6, 046, 004, 6, 046, 319、及び6, 057, 437に  
 見出すことができる。アンチセンス分子の例として、アンチセンスRNA、低分子干渉RNA(silRNA)、及びショートヘアピンRNA(shRNA)が挙げられるが、これらに限定されない。

#### 【0079】

本明細書に開示する単離された核酸分子には、RNA、DNA、またはRNAとDNAの両方を含ませることができる。単離された核酸分子を、ベクターまたは異種標識などの異種核酸配列に連結または融合することもできる。例えば、本明細書に開示する単離された核酸分子を、単離された核酸分子及び異種核酸配列を含むベクターまたは外来性ドナー配列中に存在させることができる。単離された核酸分子を、蛍光標識などの異種標識に連結または融合することもできる。標識の他の例は、本明細書の他の場所に開示する。

10

#### 【0080】

標識は、直接検出可能（例えば、フルオロフォア）または間接的に検出可能（例えば、ハプテン、酵素、またはフルオロフォア消光剤）であり得る。そのような標識は、分光学的、光化学的、生化学的、免疫化学的、または化学的手段により検出可能であり得る。そのような標識として、例えば、放射線計数装置で測定することができる放射性標識、分光度計で視覚的に観察または測定することができる色素、染料、または他の色素原、スピニラベルアナライザーで測定することができるスピニン標識、及び適切な分子付加物の励起によって出力シグナルが生成され、色素によって吸収される光で励起することにより視覚化することができるか、または標準的な蛍光計または画像化システムで測定することができる蛍光標識（例えば、フルオロフォア）が挙げられる。標識はまた、例えば、化学発光物質とすることもでき、その場合、出力シグナルは、シグナル化合物、金属含有物質、または酵素の化学修飾によって生成され、酵素の場合、無色の基質からの着色生成物の形成など、酵素依存性の二次シグナル生成が生じる。用語「標識」とはまた、複合体分子に選択的に結合させることができる「タグ」またはハプテンをも指し、それにより、基質を続けて添加する場合、複合体分子は検出可能なシグナルを生成するために使用される。例えば、ビオチンをタグとして使用し、次いでアビジンまたはストレプトアビジンの西洋ワサビペルオキシダーゼ(HRP)コンジュゲートを使用してタグに結合させ、次いで比色基質（例えば、テトラメチルベンジン(TMB)）または蛍光基質を使用してHRPの存在を検出することができる。精製を促進するためのタグとして使用することができる例示的な標識として、myc、HA、FLAGもしくは3×FLAG、6×Hisもしくはポリヒスチジン、グルタチオン-S-トランスフェラーゼ(GST)、マルトース結合タンパク質、エピトープタグ、または免疫グロブリンのFc部分が挙げられるが、これらに限定されない。多数の標識が知られており、例えば、粒子、フルオロフォア、ハプテン、酵素ならびにそれらの比色、蛍光及び化学発光基質ならびに他の標識が挙げられる。

20

#### 【0081】

開示する核酸分子は、例えば、ヌクレオチド、またはヌクレオチド類似体もしくはヌクレオチド置換体などの非天然または修飾ヌクレオチドから構成され得る。そのようなヌクレオチドには、修飾塩基、糖、またはリン酸基が含まれるか、またはその構造に非天然部分が組み込まれたヌクレオチドが含まれる。非天然ヌクレオチドの例として、ジデオキシヌクレオチド、ビオチン化、アミノ化、脱アミノ化、アルキル化、ベンジル化、及びフルオロフォア標識ヌクレオチドが挙げられるが、これらに限定されない。

30

#### 【0082】

本明細書に開示する核酸分子にはまた、1つ以上のヌクレオチド類似体または置換を含ませることができる。ヌクレオチド類似体は、塩基、糖、またはリン酸部分のいずれかへの修飾を含むヌクレオチドである。塩基部分への修飾として、A、C、G、及びT/Uの

40

50

天然及び合成修飾、ならびに例えればプソイドウリジン、ウラシル-5-イル、ヒポキサンチン-9-イル(Ⅰ)、及び2-アミノアデニン-9-イルなどの異なるプリンまたはピリミジン塩基が挙げられるが、これらに限定されない。修飾塩基として、5-メチルシトシン(5-me-C)、5-ヒドロキシメチルシトシン、キサンチン、ヒポキサンチン、2-アミノアデニン、6-メチルならびにアデニン及びグアニンの他のアルキル誘導体、2-プロピルならびにアデニン及びグアニンの他のアルキル誘導体、2-チオウラシル、2-チオチミン及び2-チオシトシン、5-ハロウラシル及びシトシン、5-プロピニルウラシル及びシトシン、6-アゾウラシル、シトシン及びチミン、5-ウラシル(プソイドウラシル)、4-チオウラシル、8-ハロ、8-アミノ、8-チオール、8-チオアルキル、8-ヒドロキシル及び他の8-置換アデニン及びグアニン、5-ハロ、特に5-ブロモ、5-トリフルオロメチルならびに他の5-置換ウラシル及びシトシン、7-メチルグアニン及び7-メチルアデニン、8-アザグアニン及び8-アザアデニン、7-デアザグアニン及び7-デアザアデニン、及び3-デアザグアニン及び3-デアザアデニンが挙げられるが、これらに限定されない。例えば、5-置換ピリミジン、6-アザピリミジン、及び2-アミノプロピルアデニン、5-プロピニルウラシル、5-プロピニルシトシン、ならびに5-メチルシトシンを含むがこれらに限定されないN-2、N-6及びO-6置換プリンなどの特定のヌクレオチド類似体は、二本鎖形成の安定性を高めることができる。多くの場合、2'-O-メトキシエチルなどの糖修飾と塩基修飾を組み合わせて、二本鎖安定性の増加などのユニークな特性を実現することができる。

#### 【0083】

ヌクレオチド類似体は、糖部分の修飾を含み得る。糖部分への修飾には、リボース及びデオキシリボースの天然修飾、ならびに合成修飾が含まれるが、これらに限定されない。糖の修飾には、2'位での以下の修飾: O H ; F ; O - 、 S - 、もしくは N - アルキル ; O - 、 S - 、もしくは N - アルケニル ; O - 、 S - 、もしくは N - アルキニル ; または O - アルキル - O - アルキルが含まれるが、これらに限定されず、アルキル、アルケニル、及びアルキニルは、置換または非置換の C<sub>1</sub>-C<sub>10</sub> アルキルまたは C<sub>2</sub>-C<sub>10</sub> アルケニル、及び C<sub>2</sub>-C<sub>10</sub> アルキニルであり得る。例示的な2'糖修飾として、-O[(CH<sub>2</sub>)<sub>n</sub>O]<sub>m</sub>CH<sub>3</sub>、-O(CH<sub>2</sub>)<sub>n</sub>OCH<sub>3</sub>、-O(CH<sub>2</sub>)<sub>n</sub>NH<sub>2</sub>、-O(CH<sub>2</sub>)<sub>n</sub>CH<sub>3</sub>、-O(CH<sub>2</sub>)<sub>n</sub>-ONH<sub>2</sub>、及び-O(CH<sub>2</sub>)<sub>n</sub>ON[(CH<sub>2</sub>)<sub>n</sub>CH<sub>3</sub>]<sub>2</sub>(式中、n及びmは1~約10である)も挙げられるが、これらに限定されない。

#### 【0084】

2'位の他の修飾として、C<sub>1</sub>-C<sub>10</sub> アルキル、置換低級アルキル、アルカリール、アラルキル、O-アルカリールまたはO-アラルキル、SH、SCH<sub>3</sub>、OCN、C1、Br、CN、CF<sub>3</sub>、OCF<sub>3</sub>、SOCH<sub>3</sub>、SO<sub>2</sub>CH<sub>3</sub>、ONO<sub>2</sub>、NO<sub>2</sub>、N<sub>3</sub>、NH<sub>2</sub>、ヘテロシクロアルキル、ヘテロシクロアルカリール、アミノアルキルアミノ、ポリアルキルアミノ、置換シリル、RNA切断基、レポーター基、インターラーカー、オリゴヌクレオチドの薬物動態特性を向上させるための基、またはオリゴヌクレオチドの薬力学的特性を向上させるための基、及び同様の特性を有する他の置換基が挙げられるが、これらに限定されない。同様の修飾を、糖における他の位置、特に3'末端ヌクレオチドまたは2'-5'結合オリゴヌクレオチドの糖の3'位、及び5'末端ヌクレオチドの5'位でも行ってよい。修飾糖には、CH<sub>2</sub>やSなど、架橋環の酸素に修飾を含むものも含まれる。ヌクレオチド糖類似体には、ペントフラノシリル糖の代わりにシクロブチル部分などの糖模倣物を含ませることもできる。

#### 【0085】

ヌクレオチド類似体は、リン酸部分で修飾することもできる。修飾リン酸部分には、2つのヌクレオチド間の結合に、ホスホロチオエート、キラルホスホロチオエート、ホスホロジチオエート、ホスホトリエステル、アミノアルキルホスホトリエステル、メチルならびに3'-アルキレンホスホネート及びキラルホスホネートなどの他のアルキルホスホネート、ホスフィネート、3'-アミノホスホルアミデート及びアミノアルキルホスホルアミデートなどのホスホルアミデート、チオノホスホルアミデート、チオノアルキルホスホネート

10

20

30

50

ト、チオノアルキルホスホトリエステル、ならびにボラノホスフェートが含まれるように修飾することができるものが含まれるが、これらに限定されない。2つのヌクレオチド間のこれらのリン酸結合または修飾リン酸結合は、3' - 5' 結合または2' - 5' 結合を介するものとすることができ、結合には、3' - 5' から5' - 3' または2' - 5' から5' - 2' などの逆極性が含まれ得る。様々な塩、混合塩、及び遊離酸の形態も含まれる。

#### 【 0 0 8 6 】

ヌクレオチド置換体には、ヌクレオチドと同様の機能特性を有するがリン酸部分を含まない分子、例えばペプチド核酸（PNA）が含まれる。ヌクレオチド置換体には、ワトソン - クリックまたはフーグスティーン様式で核酸を認識するが、リン酸部分以外の部分を介して一緒に連結する分子が含まれる。ヌクレオチド置換体は、適切な標的核酸と相互作用する場合に二重らせん型構造に適合することができる。

10

#### 【 0 0 8 7 】

ヌクレオチド置換体には、リン酸部分または糖部分が置換されているヌクレオチドまたはヌクレオチド類似体も含まれる。いくつかの実施形態では、ヌクレオチド置換体は、標準リン原子を含まない場合がある。リン酸の置換体は、例えば、短鎖アルキルもしくはシクロアルキルヌクレオシド間結合、混合ヘテロ原子及びアルキルもしくはシクロアルキルヌクレオシド間結合、または1つ以上の短鎖ヘテロ原子ヌクレオシド間結合もしくは複素環ヌクレオシド間結合であり得る。これらには、モルホリノ結合（ヌクレオシドの糖部分から一部形成される）；シロキサン骨格；硫化物、スルホキシド及びスルホン骨格；ホルムアセチル及びチオホルムアセチル骨格；メチレンホルムアセチル及びチオホルムアセチル骨格；アルケン含有骨格；スルファメート骨格；メチレンイミノ及びメチレンヒドラジノ骨格；スルホン酸及びスルホンアミド骨格；アミド骨格；ならびに混合N、O、S、及びC H<sub>2</sub>構成要素部分を有する他の成分を有するヌクレオチド置換体が含まれる。

20

#### 【 0 0 8 8 】

ヌクレオチド置換体では、ヌクレオチドの糖及びリン酸部分の両方が、例えば、アミド型結合（アミノエチルグリシン）（PNA）によって置換され得ることも理解されたい。

また、他のタイプの分子（コンジュゲート）をヌクレオチドまたはヌクレオチド類似体に連結して、例えば、細胞取り込みを増強することも可能である。コンジュゲートは、ヌクレオチドまたはヌクレオチド類似体に化学的に結合させることができる。そのようなコンジュゲートとして、例えば、コレステロール部分などの脂質部分、コール酸、ヘキシリ-S-トリチルチオールなどのチオエーテル、チオコレステロール、ドデカンジオールまたはウンデシル残基などの脂肪族鎖、ジ-ヘキサデシル-r a c-グリセロールまたはトリエチルアンモニウム1,2-ジ-O-ヘキサデシル-r a c-グリセロ-3-H-ホスホネートなどのリン脂質、ポリアミンまたはポリエチレングリコール鎖、アダマンタン酢酸、パルミチル部分、またはオクタデシルアミンもしくはヘキシルアミノ-カルボニル-オキシコレステロール部分が挙げられる。

30

#### 【 0 0 8 9 】

本開示はまた、本明細書に開示する核酸分子のいずれか1つ以上を含むベクターを提供する。いくつかの実施形態では、ベクターは、本明細書に開示する核酸分子のいずれか1つ以上及び異種核酸を含む。ベクターは、核酸分子を輸送できるウイルスベクターまたは非ウイルスベクターであり得る。いくつかの実施形態では、ベクターは、プラスミドまたはコスミド（例えば、追加のDNAセグメントをライゲーションすることができる環状二本鎖DNA）である。いくつかの実施形態では、ベクターはウイルスベクターであり、追加のDNAセグメントをウイルスゲノムにライゲーションすることができる。いくつかの実施形態では、ベクターは、それを導入する宿主細胞内で自律的に複製することができる（例えば、細菌の複製起点を有する細菌ベクター及びエピソーム哺乳類ベクター）。いくつかの実施形態では、ベクター（例えば、非エピソーム哺乳類ベクター）は、宿主細胞への導入時に宿主細胞のゲノム内に組み込むことができ、それによって宿主ゲノムとともに複製される。さらに、特定のベクターは、それらに作動可能に連結している遺伝子の発現を誘導することができる。そのようなベクターは、本明細書中では「組換え発現ベクター

40

50

」または「発現ベクター」と呼ばれる。そのようなベクターは、ターゲティングベクター（すなわち、外来性ドナー配列）でもあり得る。

【0090】

いくつかの実施形態では、開示する遺伝子バリエントをコードする核酸分子を発現ベクターに挿入し、それによって、転写及び翻訳制御配列などの発現制御配列に遺伝子を作動可能に連結することによって、本明細書に開示する様々な遺伝子バリエントによってコードされるタンパク質を発現させる。発現ベクターとして、プラスミド、コスミド、レトロウイルス、アデノウイルス、アデノ随伴ウイルス（A A V）、カリフラワーモザイクウイルス及びタバコモザイクウイルスなどの植物ウイルス、酵母人工染色体（Y A C）、エプスタイン・バー（E B V）由来エピソームなどが挙げられるが、これらに限定されない。いくつかの実施形態では、ベクター内の転写及び翻訳制御配列が遺伝子バリエントの転写及び翻訳を調節する意図された機能を果たすように、開示する遺伝子バリエントを含む核酸分子をベクターにライゲーションすることができる。発現ベクター及び発現制御配列は、使用する発現宿主細胞と適合するように選択する。開示する遺伝子バリエントを含む核酸分子をベクターにライゲーションすることができる。開示する遺伝子バリエントを含む核酸配列は、バリエントの遺伝情報とは別個のベクターまたは同じ発現ベクターに挿入することができる。開示する遺伝子バリエントを含む核酸配列は、標準的な方法（例えば、開示する遺伝子バリエントとベクターを含む核酸上の相補的制限酵素部位のライゲーション、または制限酵素部位が存在しない場合は平滑末端ライゲーション）によって発現ベクターに挿入することができる。

10

【0091】

開示する遺伝子バリエントを含む核酸配列に加えて、組換え発現ベクターには、宿主細胞内での遺伝子バリエントの発現を制御する調節配列を保有させることができる。調節配列の選択を含む発現ベクターの設計は、形質転換する宿主細胞の選択、所望のタンパク質の発現レベルなどの要因に依存する。哺乳類宿主細胞発現に望ましい調節配列として、例えば、哺乳類細胞において高レベルのタンパク質発現を誘導するウイルス因子、例えば、レトロウイルスLTR、サイトメガロウイルス（CMV）（CMVプロモーター／エンハンサーなど）、サルウイルス40（SV40）（SV40プロモーター／エンハンサーなど）、アデノウイルス（例えば、アデノウイルス主要後期プロモーター（AdMLP））、ポリオーマに由来するプロモーター及び／またはエンハンサー、ならびに強力な哺乳類プロモーター、例えば、ネイティブ免疫グロブリンプロモーター及びアクチングロモーターなどが挙げられる。細菌細胞または真菌細胞（例えば、酵母細胞）におけるポリペプチドの発現方法もまた、周知である。

20

【0092】

プロモーターは、例えば、構成的に活性なプロモーター、条件的プロモーター、誘導性プロモーター、時間的に制限されたプロモーター（例えば、発生的に調節されたプロモーター）、または空間的に制限されたプロモーター（例えば、細胞特異的または組織特異的プロモーター）であり得る。プロモーターの例は、例えばWO2013/176772に見出すことができる。

30

【0093】

誘導性プロモーターの例として、例えば、化学的に調節されるプロモーター及び物理的に調節されるプロモーターが挙げられる。化学的に調節されるプロモーターとして、例えば、アルコール調節プロモーター（例えば、アルコールデヒドロゲナーゼ（alcA）遺伝子プロモーター）、テトラサイクリン調節プロモーター（例えば、テトラサイクリン応答性プロモーター、テトラサイクリンオペレーター配列（tetO）、tet-t-Onプロモーター、またはtet-offプロモーター）、ステロイド調節プロモーター（例えば、ラットグルココルチコイド受容体、エストロゲン受容体のプロモーター、またはエクジソン受容体のプロモーター）、または金属調節プロモーター（例えば、金属タンパク質プロモーター）が挙げられる。物理的に調節されるプロモーターとして、例えば、温度調節プロモーター（例えば、熱ショックプロモーター）及び光調節プロモーター（例えば、光誘導性プロモーターまたは光抑制性プロモーター）が挙げられる。

40

50

## 【0094】

組織特異的プロモーターは、例えば、ニューロン特異的プロモーター、グリア特異的プロモーター、筋肉細胞特異的プロモーター、心臓細胞特異的プロモーター、腎細胞特異的プロモーター、骨細胞特異的プロモーター、内皮細胞特異的プロモーター、または免疫細胞特異的プロモーター（例えば、B細胞プロモーターまたはT細胞プロモーター）であり得る。

## 【0095】

発生的に調節されるプロモーターとして、例えば、発生の胚段階の間のみ、または成体細胞においてのみ活性なプロモーターが挙げられる。

開示する遺伝子バリエント及び調節配列を含む核酸配列に加えて、追加の配列、例えば、宿主細胞内でのベクターの複製を調節する配列（例えば、複製起点）及び選択マーカー遺伝子を、組換え発現ベクターに保有させることができる。選択マーカー遺伝子は、ベクターを導入した宿主細胞の選択を促進することができる（例えば、米国特許第4,399,216号、第4,634,665号、及び第5,179,017号参照）。例えば、選択マーカー遺伝子は、ベクターを導入した宿主細胞に、G418、ハイグロマイシン、メトトレキサートなどの薬剤に対する耐性を付与することができる。例示的な選択マーカー遺伝子として、ジヒドロ葉酸還元酵素（DHFR）遺伝子（メトトレキサート選択／増殖を伴うdhfr宿主細胞での使用）、neo遺伝子（G418選択用）、及びグルタミン酸合成酵素（GS）遺伝子が挙げられるが、これらに限定されない。

## 【0096】

本開示はまた、バリエント型B4GALT1ポリペプチド（Asn352Ser）を含む単離されたポリペプチドを提供する。例示的な野生型ヒトB4GALT1ポリペプチドにはUniProt登録番号P15291（配列番号7）が割り当てられており、これは398アミノ酸からなる。ヒトバリエント型B4GALT1ポリペプチドは、完全長／成熟型B4GALT1ポリペプチド（配列番号8）の352位に対応する位置にセリンを含み、野生型ヒトB4GALT1の同じ位置にあるアスパラギンとは対照的である（それぞれ、配列番号8と配列番号7とを比較）。いくつかの実施形態では、単離されたポリペプチドは配列番号8を含む。いくつかの実施形態では、単離されたポリペプチドは配列番号8からなる。

## 【0097】

いくつかの実施形態では、単離されたポリペプチドは、配列番号8と、少なくとも約70%、少なくとも約75%、少なくとも約80%、少なくとも約85%、少なくとも約90%、少なくとも約95%、少なくとも約96%、少なくとも約97%、少なくとも約98%、少なくとも約99%、または100%同一であるアミノ酸配列を含むかまたはそれからなる。いくつかの実施形態では、単離されたポリペプチドは、配列番号8の352位に対応する位置にセリンを含む。いくつかの実施形態では、単離されたポリペプチドは、配列番号8と、少なくとも約90%、少なくとも約95%、少なくとも約96%、少なくとも約97%、少なくとも約98%、少なくとも約99%、または100%同一であるアミノ酸配列を含むかまたはそれからなる。いくつかの実施形態では、単離されたポリペプチドは、配列番号8の352位に対応する位置にセリンを含む。いくつかの実施形態では、単離されたポリペプチドは、配列番号8の352位に対応する位置にセリンを含む。いくつかの実施形態では、単離されたポリペプチドは、配列番号8と少なくとも約90%同一のアミノ酸配列を含むかまたはそれからなる。いくつかの実施形態では、単離されたポリペプチドは、配列番号8と少なくとも約90%同一のアミノ酸配列を含むかまたはそれからなり、配列番号8の352位に対応する位置にセリンを含む。いくつかの実施形態では、単離されたポリペプチドは、配列番号8と少なくとも約90%同一のアミノ酸配列を含むかまたはそれからなり、ただし、その単離されたポリペプチドは、配列番号8の352位に対応する位置にセリンを含む。

## 【0098】

いくつかの実施形態では、単離されたポリペプチドは、配列番号8の352位に対応する位置にセリンを含む。いくつかの実施形態では、単離されたポリペプチドは、配列番号

10

20

30

40

50





代表的な野生型 B 4 G A L T 1 ポリペプチド配列を、配列番号 7 に記載する。代表的な B 4 G A L T 1 バリアント型ポリペプチド配列を、配列番号 8 に記載する。

本明細書に開示する単離されたポリペプチドは、天然の B 4 G A L T 1 ポリペプチドのアミノ酸配列を含み得るか、または非天然の配列を含み得る。いくつかの実施形態では、天然の配列は、保存的アミノ酸置換のために、非天然の配列と異なり得る。例えば、その配列は、保存的アミノ酸置換を除いて同一であり得る。

#### 【 0 1 0 3 】

いくつかの実施形態では、本明細書に開示する単離されたポリペプチドを、異種ポリペプチドもしくは異種分子または標識に連結または融合させ、その多くの例を本明細書の他の場所に開示する。例えば、タンパク質を異種ポリペプチドに融合させて、安定性を向上または低下させることができる。融合ドメインまたは異種ポリペプチドは、N 末端、C 末端、またはポリペプチド内部に配置することができる。融合パートナーは、例えば、T ヘルパーエピトープの提供を支援することができ（免疫学的融合パートナー）、または天然の組換えポリペプチドよりも高い収率でのタンパク質の発現を支援することができる（発現エンハンサー）。特定の融合パートナーは、免疫学的及び発現増強融合パートナーの両方である。ポリペプチドの溶解度を高めるか、または所望の細胞内区画へのポリペプチドの標的化を促進する他の融合パートナーを選択してもよい。いくつかの融合パートナーは、ポリペプチドの精製を促進する親和性タグを有する。

#### 【 0 1 0 4 】

いくつかの実施形態では、融合タンパク質を、異種分子に直接融合させるか、ペプチドリンカーなどのリンカーを介して異種分子に連結させる。適切なペプチドリンカー配列は、例えば、以下の因子に基づいて選択してもよい：1) 柔軟で拡張的な立体構造をとり得る能力、2) 第 1 及び第 2 のポリペプチド上の機能的エピトープと相互作用することができる二次構造をとることに対する抵抗性、及び 3) ポリペプチドの機能的エピトープと反応する可能性のある疎水性または荷電残基の欠如。例えば、ペプチドリンカー配列は、G 1 y、A s n 及び S e r 残基を含み得る。リンカー配列には、T h r 及び A l a などの他の中性に近いアミノ酸もまた、使用してもよい。リンカーとして有用に使用し得るアミノ酸配列として、例えば、M a r a t e a e t a l . , G e n e , 1 9 8 5 , 4 0 , 3 9 - 4 6 、M u r p h y e t a l . , P r o c . N a t l . A c a d . S c i . U S A , 1 9 8 6 , 8 3 , 8 2 5 8 - 8 2 6 2 、ならびに米国特許第 4 , 9 3 5 , 2 3 3 号及び第 4 , 7 5 1 , 1 8 0 号に開示されているアミノ酸配列が挙げられる。リンカー配列は、一般的に、例えば、長さが 1 ~ 約 5 0 アミノ酸であり得る。第 1 及び第 2 のポリペプチドが、機能的ドメインを分離し、立体干渉を防ぐことができる非必須の N 末端アミノ酸領域を有する場合、リンカー配列は通常必要ない。

#### 【 0 1 0 5 】

いくつかの実施形態では、ポリペプチドを、細胞透過性ドメインに作動可能に連結する。例えば、細胞透過性ドメインは、H I V - 1 T A T タンパク質、ヒト B 型肝炎ウイルス由来 T L M 細胞透過性モチーフ、M P G 、P e p - 1 、V P 2 2 、単純ヘルペスウイルス由来細胞透過性ペプチド、またはポリアルギニンペプチド配列由来であり得る。例えば、W O 2 0 1 4 / 0 8 9 2 9 0 参照。細胞透過性ドメインは、N 末端、C 末端、またはタンパク質内のどこにでも配置することができる。

#### 【 0 1 0 6 】

いくつかの実施形態では、ポリペプチドを、追跡または精製を容易にするために、蛍光タンパク質、精製タグ、またはエピトープタグなどの異種ポリペプチドに作動可能に連結する。蛍光タンパク質の例として、緑色蛍光タンパク質（例えば、G F P 、G F P - 2 、t a g G F P 、t u r b o G F P 、e G F P 、E m e r a l d 、A z a m i G r e e n 、単量体 A z a m i G r e e n 、C o p G F P 、A c e G F P 、Z s G r e e n 1 ）、黄色蛍光タンパク質（例えば、Y F P 、e Y F P 、C i t r i n e 、V e n u s 、Y P e t 、P h i Y F P 、Z s Y e l l o w 1 ）、青色蛍光タンパク質（e B F P 、e B F P 2 、A z u r i t e 、m K a l a m a l 、G F P u v 、S a p p h i r e 、T - S a p p h

10

20

30

40

50

ire)、シアン蛍光タンパク質(eCFP、Cerulean、CyPet、AmCyan1、Midoriishi-Cyan)、赤色蛍光タンパク質(mKate、mKate2、mPlum、DsRed単量体、mCherry、mRFP1、DsRed-Express、DsRed2、DsRed-Monomer、HcRed-Tandem、HcRed1、AsRed2、eqFP611、mRaspberry、mStrawberry、Jred)、橙色蛍光タンパク質(mOrange、mKO、Kusabira-Orange、単量体Kusabira-Orange、mTangerine、tdTomato)、及び任意の他の適切な蛍光タンパク質が挙げられるがこれらに限定されない。タグの例として、グルタチオン-S-トランスフェラーゼ(GST)、キチン結合タンパク質(CBP)、マルトース結合タンパク質、チオレドキシン(TRX)、ポリ(NANP)、タンデムアフィニティー精製(TAP)タグ、myc、AcV5、AU1、AU5、E、ECS、E2、FLAG、血球凝集素(HA)、nus、Softag1、Softag3、Strep、SBP、Glu-Glu、HSV、KT3、S、S1、T7、V5、VSV-G、ヒスチジン(His)、ビオチンカルボキシルキャリアタンパク質(BCCP)、及びカルモジュリンが挙げられるがこれらに限定されない。いくつかの実施形態では、異種分子は、免疫グロブリンFcドメイン、ペプチドタグ、形質導入ドメイン、ポリ(エチレングリコール)、ポリシアル酸、またはグリコール酸である。

#### 【0107】

いくつかの実施形態では、単離されたポリペプチドは、非天然もしくは修飾アミノ酸またはペプチド類似体を含む。例えば、多数のD-アミノ酸または天然アミノ酸とは異なる機能的置換基を有するアミノ酸が存在する。天然ペプチドとは逆の立体異性体、及びペプチド類似体の立体異性体を開示する。これらのアミノ酸は、tRNA分子に最適なアミノ酸をチャージし、例えばアンバーコドンを利用する遺伝子構築物を作出する方法でペプチド鎖にアミノ酸類似体を挿入することにより、ポリペプチド鎖に容易に組み込むことができる。

#### 【0108】

いくつかの実施形態では、単離されたポリペプチドは、ペプチド模倣物であり、ペプチドに類似するように生成され得るが、天然のペプチド結合を介して接続しない。例えば、アミノ酸またはアミノ酸類似体の結合として、-CH<sub>2</sub>NH-、-CH<sub>2</sub>S-、-CH<sub>2</sub>-、-CH=CH- (シス及びトランス)、-COCH<sub>2</sub>-、-CH(OH)CH<sub>2</sub>-、及び-CHH<sub>2</sub>SO-が挙げられるが、これらに限定されない。ペプチド類似体には、結合原子間に、b-アラニン、ガミノ酪酸などの複数の原子を含ませることができる。アミノ酸類似体及びペプチド類似体は、多くの場合、増強されたまたは望ましい特性、例えば、より経済的な生産、より高い化学的安定性、より増強された薬理学的特性(半減期、吸収、効力、有効性など)、特異性の変化(例えば、広範な生物活性)、抗原性の低下、及び他の望ましい特性を有する。

#### 【0109】

いくつかの実施形態では、単離されたポリペプチドはD-アミノ酸を含み、D-アミノ酸はペプチダーゼによって認識されないことから、これらを使用してより安定なペプチドを生成することができる。コンセンサス配列の1つ以上のアミノ酸を、同じタイプのD-アミノ酸(例えば、L-リジンの代わりにD-リジン)で体系的に置換して、より安定したペプチドを生成することができる。システイン残基を使用して、2つ以上のペプチドを環化または結合することができる。これは、ペプチドを特定の立体構造に拘束するのに有益であり得る(例えば、Rizzo and Gierasch, Ann. Rev. Biochem., 1992, 61, 387参照)。

#### 【0110】

本開示はまた、本明細書に開示するポリペプチドのいずれかをコードする核酸分子を提供する。これには、特定のポリペプチド配列に関連するすべての縮重配列(すなわち、1つの特定のポリペプチド配列をコードする配列を有するすべての核酸、及びタンパク質配列の開示するバリエント及び誘導体をコードする縮重核酸を含むすべての核酸)が含まれ

10

20

30

40

50

る。したがって、各々の特定の核酸配列は本明細書に記載し得ないが、各配列はすべて、開示するポリペプチド配列を通じて本明細書に実際に開示及び記載されるものとする。

【0111】

本開示はまた、核酸分子のいずれか1つ以上及び／または本明細書に開示するポリペプチドのいずれか1つ以上を含む組成物を提供する。いくつかの実施形態では、組成物は担体を含む。いくつかの実施形態では、担体は、核酸分子及び／またはポリペプチドの安定性を高める（例えば、所定の保存条件（例えば、-20、4、または周囲温度）下での期間を延長し、その場合、分解産物は、出発核酸またはタンパク質の重量で0.5%未満など、閾値未満で維持され；またはin vivoでの安定性を高める）。担体の例として、ポリ（乳酸）（PLA）ミクロスフェア、ポリ（D、L-乳酸-コグリコール酸）（PLGA）ミクロスフェア、リポソーム、ミセル、逆ミセル、脂質コクリエート、及び脂質微小管が挙げられるが、これらに限定されない。

【0112】

本開示はまた、本明細書に開示するB4GALT1ポリペプチドまたはその断片のいずれかの产生方法を提供する。そのようなB4GALT1ポリペプチドまたはその断片は、任意の適切な方法により产生することができる。例えば、B4GALT1ポリペプチドまたはその断片は、そのようなB4GALT1ポリペプチドまたはその断片をコードする核酸分子（例えば、組換え発現ベクター）を含む宿主細胞から产生することができる。そのような方法は、B4GALT1ポリペプチドまたはその断片を产生するのに十分な条件下でB4GALT1ポリペプチドまたはその断片をコードする核酸分子（例えば、組換え発現ベクター）を含む宿主細胞を培養し、それによりB4GALT1ポリペプチドまたはその断片を产生することを含み得る。核酸を、宿主細胞において活性なプロモーターに作動可能に連結することができ、培養は、核酸が発現する条件下で実施することができる。そのような方法は、発現するB4GALT1ポリペプチドまたはその断片を回収することをさらに含み得る。回収は、B4GALT1ポリペプチドまたはその断片を精製することをさらに含み得る。

【0113】

タンパク質発現に適した系の例として、例えば：細菌細胞発現系（例えば、Escherichia coli、Lactococcus lactis）、酵母細胞発現系（例えば、Saccharomyces cerevisiae、Pichia pastoris）、昆虫細胞発現系（例えば、バキュロウイルスを介したタンパク質発現）、及び哺乳類細胞発現系などの宿主細胞が挙げられる。

【0114】

B4GALT1ポリペプチドまたはその断片をコードする核酸分子の例を、本明細書の他の場所でより詳細に開示する。いくつかの実施形態では、核酸分子を、宿主細胞での発現のためにコドン最適化する。いくつかの実施形態では、核酸分子を、宿主細胞で活性なプロモーターに作動可能に連結する。プロモーターは、異種プロモーター（すなわち、天然のB4GALT1プロモーターではないプロモーター）とすることができます。Escherichia coliに適したプロモーターの例として、アラビノース、lac、tac、及びT7プロモーターが挙げられるが、これらに限定されない。Lactococcus lactisに適したプロモーターの例として、P170及びナイシンプロモーターが挙げられるが、これらに限定されない。Saccharomyces cerevisiaeに適したプロモーターの例として、アルコールデヒドロゲナーゼ（ADH1）またはエノラーゼ（ENO）プロモーターなどの構成的プロモーター、またはPHO、CUP1、GAL1、及びG10などの誘導性プロモーターが挙げられるが、これらに限定されない。Pichia pastorisに適したプロモーターの例として、アルコールオキシダーゼI（AOX I）プロモーター、グリセルアルデヒド3リン酸デヒドロゲナーゼ（GAP）プロモーター、及びグルタチオン依存性ホルムアルデヒドデヒドロゲナーゼ（FLDI）プロモーターが挙げられるが、これらに限定されない。バキュロウイルスを介した系に適したプロモーターの例は、後期ウイルスの強力なポリヘドリンプロモー

10

20

30

40

50

ターである。

【0115】

いくつかの実施形態では、核酸分子は、タンパク質精製を促進するために、B 4 G A L T 1 ポリペプチドまたはその断片とともにインフレームでタグをコードする。タグの例を、本明細書の他の場所で開示する。そのようなタグは、例えば、他のすべてのタンパク質（例えば、宿主細胞タンパク質）からタグ付きタンパク質を単離できるように、パートナーリガンドに結合する（例えば、樹脂に固定化する）ことができる。アフィニティーコロマトグラフィー、高速液体クロマトグラフィー（H P L C）、及びサイズ排除クロマトグラフィー（S E C）は、発現タンパク質の純度を向上させるために使用することができる方法の例である。

10

【0116】

他の方法を使用して、B 4 G A L T 1 ポリペプチドまたはその断片を産生することもできる。例えば、タンパク質化学技術により 2 つ以上のペプチドまたはポリペプチドを連結することができる。例えば、F m o c ( 9 - フルオレニルメチルオキシカルボニル ) または B o c ( t e r t - ブチルオキシカルボノイル ) のいずれかの化学物質を使用して、ペプチドまたはポリペプチドを化学合成することができる。そのようなペプチドまたはポリペプチドは、標準的な化学反応によって合成することができる。例えば、あるペプチドまたはポリペプチドを合成し、それはその合成樹脂から切断され得ない一方で、ペプチドまたはタンパク質の他の断片は、合成され、その後樹脂から切断され、それにより他の断片上では機能的にロックされた末端基を露出させることができる。ペプチド縮合反応により、これらの 2 つの断片を、それぞれカルボキシル末端及びアミノ末端でペプチド結合を介して共有結合することができる。あるいは、ペプチドまたはポリペプチドを、本明細書に記載するように *i n v i v o* で独立して合成することができる。一旦単離した後、これらの独立したペプチドまたはポリペプチドを同様のペプチド縮合反応を介して連結して、ペプチドまたはその断片を形成してもよい。

20

【0117】

いくつかの実施形態では、クローニングしたか、または合成のペプチドセグメントの酵素的ライゲーションにより、比較的短いペプチド断片を連結して、より大きなペプチド断片、ポリペプチド、または全タンパク質ドメインを生成することができる（A b r a h m s e n et al. , B i o c h e m i s t r y , 1 9 9 1 , 3 0 , 4 1 5 1 ）。あるいは、合成ペプチドのネイティブな化学ライゲーションを利用して、短いペプチド断片から大きなペプチドまたはポリペプチドを合成的に構築することができる。この方法は、2 工程の化学反応からなる（D a w s o n et al. , S c i e n c e , 1 9 9 4 , 2 6 6 , 7 7 6 - 7 7 9 参照）。第 1 の工程は、保護されていない合成ペプチド - チオエステルと、アミノ末端 C y s 残基を含む別の保護されていないペプチドセグメントとの化学選択的反応で、第 1 の共有結合生成物としてチオエステル結合中間体を生成する。反応条件を変更することなく、この中間体は自発的かつ迅速な分子内反応を起こし、ライゲーション部位で天然のペプチド結合を形成することができる。

30

【0118】

いくつかの実施形態では、化学ライゲーションの結果としてペプチドセグメント間に形成される結合が非天然（非ペプチド）結合である場合、保護されていないペプチドセグメントを化学的に連結することができる（S c h n o l z e r et al. , S c i e n c e , 1 9 9 2 , 2 5 6 , 2 2 1 参照）。

40

【0119】

本開示はまた、本明細書に開示する任意の 1 つ以上の核酸分子及び / または任意の 1 つ以上のポリペプチドを有する細胞（例えば、組換え宿主細胞）を提供する。細胞は、*i n v i t r o*、*ex v i v o*、または *i n v i v o* であり得る。核酸分子をプロモーター及び他の調節配列に連結し、これを発現させてコードするタンパク質を産生することができる。

【0120】

50

いくつかの実施形態では、細胞は、全能性細胞または多能性細胞（例えば、げっ歯類 E S 細胞、マウス E S 細胞、またはラット E S 細胞などの胚性幹（E S）細胞）である。全能性細胞には、任意の細胞型を生じさせることができる未分化細胞が含まれ、多能性細胞には、複数の分化細胞型に発達する能力を有する未分化細胞が含まれる。そのような多能性及び／または全能性細胞は、例えば、E S 細胞または人工多能性幹（i P S）細胞などの E S 様細胞であり得る。E S 細胞には、胚への導入時に発生中の胚の任意の組織に寄与し得る胚由来の全能性または多能性細胞が含まれる。E S 細胞は胚盤胞の内部細胞塊に由来し、3つの脊椎動物の胚葉（内胚葉、外胚葉、及び中胚葉）のいずれかの細胞に分化することができる。

#### 【0121】

いくつかの実施形態では、細胞は、初代体細胞、または初代体細胞ではない細胞である。体細胞には、配偶子、生殖細胞、配偶子細胞、または未分化幹細胞ではない任意の細胞が含まれ得る。いくつかの実施形態では、細胞は初代細胞とすることもできる。初代細胞には、生物、器官、または組織から直接分離した細胞または細胞の培養物が含まれる。初代細胞には、形質転換も不死化もしていない細胞が含まれる。初代細胞には、組織培養で以前に継代されていないか、または組織培養で以前に継代されているが組織培養で無期限に継代することができない生物、器官、または組織から得られた細胞が含まれる。そのような細胞は、従来技術により単離することができ、例えば、体細胞、造血細胞、内皮細胞、上皮細胞、線維芽細胞、間葉細胞、ケラチノサイト、メラニン細胞、単球、単核細胞、脂肪細胞、前脂肪細胞、ニューロン、グリア細胞、肝細胞、骨格筋芽細胞、及び平滑筋細胞が挙げられる。例えば、初代細胞は、結合組織、筋肉組織、神経系組織、または上皮組織に由来し得る。

#### 【0122】

いくつかの実施形態では、細胞は通常、無期限に増殖できないが、変異または変化により、正常な細胞老化が回避され、それどころか分裂し続けることができる。そのような変異または変化は、自然に発生し得るか、または意図的に誘発され得る。不死化細胞の例として、チャイニーズハムスター卵巣（C H O）細胞、ヒト胎児腎細胞（例えば、H E K 2 9 3 細胞）、及びマウス胚線維芽細胞（例えば、3 T 3 細胞）が挙げられるが、これらに限定されない。多数のタイプの不死化細胞が周知である。不死化細胞または初代細胞には、組換え遺伝子または組換えタンパク質の培養または発現に通常使用される細胞が含まれる。いくつかの実施形態では、細胞は、肝細胞（例えば、ヒト肝細胞）などの分化細胞である。

#### 【0123】

細胞は、任意の供給源に由来し得る。例えば、細胞は、真核細胞、動物細胞、植物細胞、または真菌（例えば、酵母）細胞とすることができます。そのような細胞は、魚類細胞または鳥類細胞であり得るか、またはそのような細胞は、哺乳類細胞、例えば、ヒト細胞、非ヒト哺乳類細胞、げっ歯類細胞、マウス細胞またはラット細胞であり得る。哺乳類として、ヒト、非ヒト霊長類、サル、類人猿、ネコ、イヌ、ウマ、ウシ、シカ、バイソン、ヒツジ、げっ歯類（例えば、マウス、ラット、ハムスター、モルモット）、家畜（例えば、ウシ、去勢ウシなどのウシ種、ヒツジ、ヤギなどのヒツジ種、ならびにブタ及びイノシシなどのブタ種）が挙げられるが、これらに限定されない。鳥類として、ニワトリ、シチメンチョウ、ダチョウ、ガチョウ、アヒルなどが挙げられるが、これらに限定されない。家畜動物及び農業動物も含まれる。用語「非ヒト動物」には、ヒトは含まれない。

#### 【0124】

本開示はまた、ヒト対象由来の生体試料中の B 4 G A L T 1 バリアント遺伝子、m R N A、c D N A、及び／またはポリペプチドの存在の検出方法を提供する。集団内の遺伝子配列、及びそのような遺伝子によってコードされるm R N A 及びタンパク質は、一塩基多型などの多型により変化し得ることを理解されたい。本明細書中で提供する B 4 G A L T 1 遺伝子、m R N A、c D N A、及びポリペプチドの配列は例示的な配列に過ぎない。B 4 G A L T 1 遺伝子、m R N A、c D N A、及びポリペプチドの他の配列も可能である。

10

20

30

40

50

〔 0 1 2 5 〕

生体試料は、対象由来の任意の細胞、組織、または生体液に由来し得る。試料は、臨床的に適切な任意の組織、例えば、骨髄試料、腫瘍生検、細針吸引液、または体液試料、例えば、血液、血漿、血清、リンパ液、腹水、囊胞液、もしくは尿を含み得る。いくつかの場合では、試料には口腔スワブが含まれる。本明細書に開示する方法において使用する試料は、アッセイ形式、検出方法の性質、及び試料として使用する組織、細胞、または抽出物に基づいて様々に異なる。生体試料は、用いるアッセイに応じて異なる方法で処理することができる。例えば、バリアント型 B 4 G A L T 1 核酸分子を検出する場合、ゲノム D N A の試料を単離または濃縮するように設計した予備処理を用いることができる。この目的のために、様々な公知の技術を使用してもよい。B 4 G A L T 1 m R N A のレベルを検出する場合、様々な手法を用いて、生体試料を m R N A で富化することができる。m R N A の存在もしくはレベル、または特定のバリアント型ゲノム D N A 遺伝子座の存在を検出するための様々な方法を使用することができる。

10

【 0 1 2 6 】

いくつかの実施形態では、本開示は、生体試料中の核酸の少なくとも部分を配列決定して、その核酸が、配列番号2の53757～53577位に対応する位置の配列番号2のヌクレオチド53757～53577を含むかどうかを判定することを含む、バリエント型B4GALT1核酸分子の存在または非存在の検出方法を提供する。

20

【 0 1 2 7 】

いくつかの実施形態では、本開示は、生体試料中の核酸の少なくとも部分を配列決定して、その核酸が、配列番号4の1243～1245位に対応する位置の配列番号4のヌクレオチド1243～1245を含むかどうかを判定することを含む、バリエント型B4GA L T1核酸分子の存在または非存在の検出方法を提供する。

【 0 1 2 8 】

いくつかの実施形態では、本開示は、生体試料中の核酸の少なくとも部分を配列決定して、その核酸が、配列番号6の1054～1056位に対応する位置の配列番号6のヌクレオチド1054～1056を含むかどうかを判定することを含む、バリアント型B4GALT1核酸分子の存在または非存在の検出方法を提供する。

30

〔 0 1 2 9 〕

いくつかの実施形態では、ヒト対象におけるバリアント型 B 4 G A L T 1 核酸分子（例えば、遺伝子、mRNA、またはcDNA）の存在または非存在の検出方法は：生体試料中の核酸分子が配列番号 8 の 352 位のセリンをコードする核酸配列を含むかどうかを判定するアッセイを、ヒト対象由来の生体試料に対して実施することを含む。いくつかの実施形態では、生体試料には、細胞または細胞溶解物が含まれる。そのような方法は、例えば、B 4 G A L T 1 遺伝子、mRNA、またはcDNA を含む対象由来の生体試料を取得し、配列番号 2（遺伝子）の 53757 ~ 53577 位、配列番号 4（mRNA）の 1243 ~ 1245 位、または配列番号 6（cDNA）の 1054 ~ 1056 位に対応する B 4 G A L T 1 遺伝子、mRNA、またはcDNA の位置が、バリアント型 B 4 G A L T 1 ポリペプチドの 352 位に対応する位置にアスパラギンの代わりにセリンをコードすることを判定するアッセイを生体試料に対して行うことを含み得る。そのようなアッセイは、例えば、特定の B 4 G A L T 1 核酸分子のこれらの位置の同一性を決定することを含み得る。

40

【 0 1 3 0 】

いくつかの実施形態では、アッセイは：ヒト対象由来の生体試料中の核酸分子の B 4 G A L T 1 ゲノム配列の部分を配列決定することを含み、配列決定する部分には、配列番号 2 の 5 3 5 7 5 ~ 5 3 5 7 7 位に対応する位置が含まれ、ヒト対象由来の生体試料中の核酸分子の B 4 G A L T 1 m R N A 配列の部分を配列決定することを含み、配列決定する部分には、配列番号 4 の 1 2 4 3 ~ 1 2 4 5 位に対応する位置が含まれ、または、ヒト対象由来の生体試料中の核酸分子の B 4 G A L T 1 c D N A 配列の部分を配列決定することを含み、配列決定する部分には、配列番号 6 の 1 0 5 4 ~ 1 0 5 6 位に対応する位置が

50

含まれる。

【0131】

いくつかの実施形態では、アッセイは：a) 生体試料を：i) 配列番号2の53575～53577位に対応するB4GALT1ゲノム配列の位置に近接するB4GALT1ゲノム配列の部分、ii) 配列番号4の1243～1245位に対応するB4GALT1mRNAの位置に近接するB4GALT1mRNA配列の部分、またはiii) 配列番号6の1054～1056位に対応するB4GALT1cDNAの位置に近接するB4GALT1cDNA配列の部分にハイブリダイズするプライマーと接触させること、b) 少なくとも：i) 53575～53577位に対応するB4GALT1ゲノム配列の位置、ii) 1243～1245位に対応するB4GALT1mRNAの位置、またはiii) 1054～1056位に対応するB4GALT1cDNAの位置を介してプライマーを伸長すること、及びc) プライマーの伸長産物が：配列番号8の352位のセリンをコードするヌクレオチドをi) B4GALT1ゲノム配列の53575～53577位に対応する、ii) B4GALT1mRNAの1243～1245位に対応する、またはiii) B4GALT1cDNAの1054～1056位に対応する位置に含むかどうかを判定することを含む。いくつかの実施形態では、B4GALT1ゲノムDNAのみを分析する。いくつかの実施形態では、B4GALT1mRNAのみを分析する。いくつかの実施形態では、B4GALT1cDNAのみを分析する。

10

【0132】

いくつかの実施形態では、アッセイは、ストリンジエントな条件下でバリアント型B4GALT1ゲノム配列、mRNA配列、またはcDNA配列に特異的にハイブリダイズし、対応する野生型B4GALT1配列にはハイブリダイズしないプライマーまたはプローブに生体試料を接触させること、及びハイブリダイゼーションが生じたかどうかを判定することを含む。

20

【0133】

いくつかの実施形態では、上記のアッセイは、RNA配列決定(RNA-Seq)を含む。いくつかの実施形態では、アッセイはまた、逆転写ポリメラーゼ連鎖反応(RT-PCR)を含む。

【0134】

いくつかの実施形態では、方法は、標的核酸配列に結合し、バリアント型B4GALT1遺伝子、mRNA、またはcDNAを含むポリヌクレオチドを特異的に検出及び/または同定するのに十分なヌクレオチド長のプローブ及びプライマーを利用する。ハイブリダイゼーション条件または反応条件は、この結果を達成するために作業者が決定することができる。この長さは、選択する検出方法において使用するのに十分な任意の長さであってよい。一般的に、例えば、約8、約11、約14、約16、約18、約20、約22、約24、約26、約28、約30、約40、約50、約75、約100、約200、約300、約400、約500、約600、または約700ヌクレオチド以上、または約11～約20、約20～約30、約30～約40、約40～約50、約50～約100、約100～約200、約200～約300、約300～約400、約400～約500、約500～約600、約600～約700、または約700～約800、またはそれ以上の長さのヌクレオチドを使用する。そのようなプローブ及びプライマーは、高ストリンジエンシーのハイブリダイゼーション条件下で標的配列に特異的にハイブリダイズすることができる。プローブ及びプライマーは、標的配列に対して連続ヌクレオチドの完全な核酸配列同一性を有していてもよいが、プローブは標的核酸配列とは異なっており、また、標的核酸配列を特異的に検出及び/または同定する能力を保持しており、従来の方法により設計され得る。したがって、プローブ及びプライマーは、標的核酸分子と、約80%、約85%、約90%、約91%、約92%、約93%、約94%、約95%、約96%、約97%、約98%、約99%、または100%の配列同一性または相補性を共有することができる。

30

【0135】

40

50

いくつかの実施形態では、特異的プライマーを使用してバリアント型 B 4 G A L T 1 遺伝子座及び / または B 4 G A L T 1 バリアント型 m R N A または c D N A を增幅してアンプリコンを生成することができ、それは、生体試料中のバリアント型 B 4 G A L T 1 遺伝子座を同定するために、もしくは特定の B 4 G A L T 1 m R N A または c D N A のレベルを測定するために、特異的プローブとして使用することができるか、またはそれ自体を検出することができる。B 4 G A L T 1 バリアント型遺伝子座を使用して、配列番号 2 の 5 3 5 7 5 ~ 5 3 5 7 7 位に対応する位置を含むゲノム核酸配列を示すことができる。プローブが核酸分子に結合可能な条件下でプローブが生体試料中の核酸分子とハイブリダイズする場合、この結合を検出することができ、生体試料中のバリアント型 B 4 G A L T 1 遺伝子座の存在またはバリアント型 B 4 G A L T 1 m R N A または c D N A の存在またはレベルを示すことができる。結合プローブのそのような同定は既に説明した。特異的プローブは、バリアント型 B 4 G A L T 1 遺伝子の特定の領域と、少なくとも約 8 0 %、約 8 0 % ~ 約 8 5 %、約 8 5 % ~ 約 9 0 %、約 9 0 % ~ 約 9 5 %、及び約 9 5 % ~ 約 1 0 0 % 同一の（または相補的な）配列を含んでいてもよい。特異的プローブは、バリアント型 B 4 G A L T 1 m R N A の特定の領域と、少なくとも約 8 0 %、約 8 0 % ~ 約 8 5 %、約 8 5 % ~ 約 9 0 %、約 9 0 % ~ 約 9 5 %、及び約 9 5 % ~ 約 1 0 0 % 同一の（または相補的な）配列を含んでいてもよい。特異的プローブは、バリアント型 B 4 G A L T 1 c D N A の特定の領域と、少なくとも約 8 0 %、約 8 0 % ~ 約 8 5 %、約 8 5 % ~ 約 9 0 %、約 9 0 % ~ 約 9 5 %、及び約 9 5 % ~ 約 1 0 0 % 同一の（または相補的な）配列を含んでいてもよい。

【 0 1 3 6 】

いくつかの実施形態では、生体試料の核酸相補体がバリアント型 B 4 G A L T 1 遺伝子座（配列番号 2）の 5 3 5 7 5 ~ 5 3 5 7 7 位にセリンをコードするヌクレオチドを含むかどうかを判定するために、5 3 5 7 5 ~ 5 3 5 7 7 位に隣接する 5' 隣接配列に由来する第 1 のプライマーと 5 3 5 7 5 ~ 5 3 5 7 7 位に隣接する 3' 隣接配列に由来する第 2 のプライマーを含むプライマーペアを使用して、バリアント型 B 4 G A L T 1 遺伝子座（配列番号：2）の 5 3 5 7 5 ~ 5 3 5 7 7 位における SNP の存在を診断するアンプリコンを生成する核酸増幅法に、生体試料を供してもよい。いくつかの実施形態では、アンプリコンの長さは、プライマー対と 1 ヌクレオチド塩基対とを合わせた長さから DNA 増幅プロトコールにより生成可能なアンプリコンの任意の長さまでの範囲であってよい。この距離は、1 ヌクレオチド塩基対から増幅反応の限界、すなわち約 2 万ヌクレオチド塩基対までの範囲とすることができます。場合により、プライマー対は、5 3 5 7 5 ~ 5 3 5 7 7 位及び 5 3 5 7 5 ~ 5 3 5 7 7 位の両側の少なくとも 1、2、3、4、5、6、7、8、9、10、またはそれ以上のヌクレオチドを含む領域に隣接する。同様のアンプリコンを、mRNA 及び / または cDNA 配列から生成することができる。

〔 0 1 3 7 〕

プローブ及びプライマーを調製及び使用するための代表的な方法は、例えば、Molecular Cloning: A Laboratory Manual, 2nd Ed., Vol. 1-3, ed. Sambrook et al., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. 1989(以下、「Sambrook et al., 1989」)、Current Protocols in Molecular Biology, ed. Ausubel et al., Greene Publishing and Wiley-Interscience, New York, 1992(定期的更新あり)(以下、「Ausubel et al., 1992」)、及びInnis et al., PCR Protocols: A Guide to Methods and Applications, Academic Press: San Diego, 1990)に記載されている。PCRプライマー対は、既知の配列から派生させることができ、例えばVector NTIバージョン10のPCRプライマー分析ツール(Informax Inc., Bethesda Md)、PrimerSelect(DNASTAR Inc., Mad

10

20

30

40

50

ison, Wis.)、及びPrimer3(バージョン0.4.0. COPYRIGHT., 1991, Whitehead Institute for Biomedical Research, Cambridge, Mass.)など、その目的を意図したコンピュータープログラムを使用して行うことができる。さらに、既知のガイドラインを使用して、配列を視覚的にスキャンし、プライマーを手動で同定することができる。

#### 【0138】

以下でさらに詳細に説明するように、任意の従来の核酸ハイブリダイゼーションまたは増幅または配列決定方法を使用して、バリアント型B4GALT1遺伝子座の存在及び/またはバリアント型B4GALT1 mRNAまたはcDNAのレベルを特異的に検出すことができる。いくつかの実施形態では、核酸分子を、B4GALT1核酸の領域を増幅するためのプライマーとして使用でき、あるいは核酸分子を、バリアント型B4GALT1遺伝子座を含む核酸分子またはバリアント型B4GALT1 mRNAもしくはcDNAを含む核酸分子にストリンジエントな条件下でハイブリダイズするプローブとして使用することができる。

#### 【0139】

例えば、核酸配列決定、核酸ハイブリダイゼーション、及び核酸増幅を含む、様々な核酸技術が公知である。核酸配列決定技術の用例として、連鎖停止剤(サンガー)配列決定法及びダイターミネーター配列決定法が挙げられるが、これらに限定されない。

#### 【0140】

他の方法は、精製DNA、増幅DNA、及び固定細胞調製物(蛍光in situハイブリダイゼーション)に対する標識プライマーまたはプローブの使用を含む、配列決定以外の核酸ハイブリダイゼーション方法を含む。いくつかの方法では、標的核酸を、検出の前または検出と同時に増幅してもよい。核酸増幅技術の用例として、ポリメラーゼ連鎖反応(PCR)、リガーゼ連鎖反応(LCR)、鎖置換型増幅法(SDA)、及び核酸配列ベースの増幅法(NASBA)が挙げられるが、これらに限定されない。他の方法として、リガーゼ連鎖反応、鎖置換型増幅法、及び好熱性SDA(tSDA)が挙げられるが、これらに限定されない。

#### 【0141】

非増幅または増幅ポリヌクレオチドの検出に対して、例えば、ハイブリダイゼーション保護アッセイ(HPA)、リアルタイムでの増幅プロセスの定量的評価、及びリアルタイム増幅に基づかない、試料中に最初に存在する標的配列の量の測定を含む、任意の方法を使用することができる。

#### 【0142】

配列増幅を必ずしも必要とせず、例えば、サザン(DNA:DNA)プロットハイブリダイゼーション、in situハイブリダイゼーション(ISH)、及び適切なプローブを使用した染色体物質の蛍光in situハイブリダイゼーション(FISH)の公知の方法に基づく核酸の同定方法も提供する。サザンプロッティングを使用して、特定の核酸配列を検出することができる。そのような方法では、試料から抽出した核酸を断片化し、マトリックスゲル上で電気泳動的に分離し、メンブレンフィルターに移す。フィルターに結合した核酸を、目的の配列に相補的な標識プローブとのハイブリダイゼーションに供する。フィルターに結合したハイブリダイズしたプローブを検出する。

#### 【0143】

ハイブリダイゼーション技術において、プローブまたはプライマーがその標的に特異的にハイブリダイズするように、ストリンジエントな条件を用いることができる。いくつかの実施形態では、ストリンジエントな条件下でのポリヌクレオチドプライマーまたはプローブは、その標的配列(例えば、バリアント型B4GALT1遺伝子座、mRNA、またはcDNA)に対し、他の配列(例えば、対応する野生型B4GALT1遺伝子座、mRNA、またはcDNA)に比べて検出可能に高い程度、例えば、バックグラウンドの少なくとも2倍またはバックグラウンドの10倍でハイブリダイズする。ストリンジエントな条件は、配列に依存し、異なる状況で異なるであろう。ハイブリダイゼーション及び/ま

10

20

30

40

50

たは洗浄条件のストリンジエンシーを制御することにより、プローブに 100 % 相補的な標的配列を同定することができる（相同的探索）。あるいは、より低い程度の同一性を検出するために、ストリンジエンシー条件を調整して配列の一部のミスマッチを許容することができる（非相同的探索）。一般的に、プローブは、長さが約 1000 ヌクレオチド未満または長さが約 500 ヌクレオチド未満である。

#### 【0144】

DNAハイブリダイゼーションを促進する適切なストリンジエンシー条件、例えば、約 45 での  $6 \times$  塩化ナトリウム / クエン酸ナトリウム (SSC) 、その後の 50 での  $2 \times$  SSC の洗浄は、公知であるか、または Current Protocols in Molecular Biology , John Wiley & Sons , N.Y. (1989) , 6.3.1 - 6.3.6 に見出すことができる。通常、ハイブリダイゼーション及び検出のストリンジエントな条件は、pH 7.0 ~ 8.3 での塩濃度が約 1.5 M Na イオン未満、通常約 0.01 ~ 1.0 M Na イオン濃度（または他の塩）であり、温度が、短いプローブ（例えば、10 ~ 50 ヌクレオチド）では少なくとも約 30 、長いプローブ（例えば、50 ヌクレオチド超）では少なくとも約 60 である。ストリンジエントな条件はまた、ホルムアミドなどの不安定化剤の添加により達成してもよい。例示的な低ストリンジエンシー条件として、30 ~ 35 % ホルムアミド、1 M NaCl 、 1 % SDS (ドデシル硫酸ナトリウム) の緩衝液中での 37 でのハイブリダイゼーション、及び  $1 \times \sim 2 \times$  SSC ( $20 \times$  SSC = 3.0 M NaCl / 0.3 M クエン酸三ナトリウム) 中での 50 ~ 55 での洗浄が挙げられる。例示的な中程度のストリンジエンシー条件として、40 ~ 45 % ホルムアミド、1.0 M NaCl 、 1 % SDS 中での 37 でのハイブリダイゼーション、及び  $0.5 \times \sim 1 \times$  SSC 中での 55 ~ 60 での洗浄が挙げられる。例示的な高ストリンジエンシー条件として、50 % ホルムアミド、1 M NaCl 、 1 % SDS 中での 37 でのハイブリダイゼーション、及び  $0.1 \times$  SSC 中での 60 ~ 65 での洗浄が挙げられる。場合により、洗浄緩衝液には、約 0.1 % ~ 約 1 % の SDS を含ませてもよい。ハイブリダイゼーションの期間は、一般的に約 24 時間未満であり、通常は約 4 ~ 約 12 時間である。洗浄時間の期間は、少なくとも平衡に達するのに十分な時間である。

#### 【0145】

ハイブリダイゼーション反応において、特異性は通常、ハイブリダイゼーション後の洗浄の関数であり、重要な因子は最終洗浄溶液のイオン強度及び温度である。DNA - DNA ハイブリッドの場合、 $T_m$  は Meinkoth and Wahl , Anal. Biochem. , 1984 , 138 , 267 - 284 :  $T_m = 81.5 + 16.6 (\log M) + 0.41 (\% GC) - 0.61 (\% form) - 500 / L$  の式から概算することができ、式中、M は一価カチオンのモル濃度、% GC は DNA のグアノシン及びシトシンヌクレオチドの割合、% form はハイブリダイゼーション溶液中のホルムアミドの割合、ならびに L は塩基対のハイブリッドの長さである。 $T_m$  は、完全に一致するプローブに対して、相補的な標的配列の 50 % がハイブリダイズする温度（定義されたイオン強度及び pH の下で）である。 $T_m$  は、1 % のミスマッチごとに約 1 ずつ低下し、したがって、 $T_m$  、ハイブリダイゼーション、及び / または洗浄条件を調整して、所望の同一性の配列にハイブリダイズさせることができる。例えば、90 % の同一性を有する配列を探す場合、 $T_m$  を 10 低下させることができる。一般的に、ストリンジエントな条件は、定義されたイオン強度及び pH での特定の配列及びその相補鎖の熱融点 ( $T_m$ ) よりも約 5 低くなるように選択する。しかしながら、厳密にストリンジエントな条件では、熱融点 ( $T_m$ ) より 1 、 2 、 3 、または 4 低い温度でハイブリダイゼーション及び / または洗浄を利用ることができ、中程度にストリンジエントな条件では、熱融点 ( $T_m$ ) より 6 、 7 、 8 、 9 、または 10 低い温度でハイブリダイゼーション及び / または洗浄を利用でき、低ストリンジエンシー条件では、熱融点 ( $T_m$ ) より 11 、 12 、 13 、 14 、 15 、または 20 低い温度でハイブリダイゼーション及び / または洗浄を利用できる。当業者であれば、方程式、ハイブリダイゼー

10

20

30

40

50

ション組成物及び洗浄組成物、ならびに所望の  $T_m$  を用いて、ハイブリダイゼーション及び / または洗浄溶液のストリンジエンシーの変動が本質的に記載されることを理解するであろう。望ましい程度のミスマッチの結果、 $T_m$  が 45 (水溶液) または 32 (ホルムアミド溶液) 未満である場合、より高い温度を用いることができるよう S S C 濃度を上げることが最適である。

#### 【 0 1 4 6 】

また、例えばタンパク質配列決定及びイムノアッセイを含む、生体試料中のバリアント型 B 4 G A L T 1 ポリペプチドの存在またはレベルの検出方法を提供する。いくつかの実施形態では、ヒト対象における B 4 G A L T 1 A s n 3 5 2 S e r の存在の検出方法は、ヒト対象由来の生体試料に対して、その生体試料中の B 4 G A L T 1 A s n 3 5 2 S e r の存在を判定するアッセイを実施することを含む。

10

#### 【 0 1 4 7 】

タンパク質配列決定技術の非限定的な用例として、質量分析及びエドマン分解が挙げられるが、これらに限定されない。イムノアッセイの用例として、免疫沈降、ウェスタンプロット、免疫組織化学、E L I S A、免疫細胞化学、フローサイトメトリー、及び免疫 P C R が挙げられるが、これらに限定されない。様々な公知技術（例えば、比色分析、蛍光、化学発光、または放射性）を使用して検出可能に標識したポリクローナルまたはモノクローナル抗体は、イムノアッセイでの使用に適している。

#### 【 0 1 4 8 】

本開示はまた、対象の心血管病態の発症に対する感受性または心血管病態の発症リスクの判定方法を提供する。対象は、例えば、ヒト、非ヒト哺乳類、げっ歯類、マウス、またはラットを含む任意の生物であり得る。いくつかの実施形態では、方法は、対象由来の生体試料中のバリアント型 B 4 G A L T 1 ゲノム D N A、m R N A、または c D N A の存在を検出することを含む。集団内の遺伝子配列及びそのような遺伝子によってコードされる m R N A は、S N P などの多型により変化し得ることを理解されたい。B 4 G A L T 1 遺伝子、m R N A、c D N A、及びポリペプチドについて本明細書中で提供する配列は例示的な配列に過ぎず、他のそのような配列も可能である。

20

#### 【 0 1 4 9 】

心血管病態の非限定的な例として、1つ以上の血清脂質のレベルの上昇が挙げられる。血清脂質には、コレステロール、L D L、H D L、トリグリセリド、H D Lコレステロール、及び非 H D Lコレステロール、またはその任意の亜分画（例えば、H D L 2、H D L 2 a、H D L 2 b、H D L 2 c、H D L 3、H D L 3 a、H D L 3 b、H D L 3 c、H D L 3 d、L D L 1、L D L 2、L D L 3、リポタンパク質 A、L p a 1、L p a 1、L p a 3、L p a 4、または L p a 5）の1つ以上が含まれる。心血管病態には、冠動脈石灰化レベルの上昇が含まれ得る。心血管病態には、I I d型グリコシル化 ( C D G - I I d ) が含まれ得る。心血管病態には、心膜脂肪レベルの上昇が含まれ得る。心血管病態には、冠動脈疾患 ( C A D )、心筋梗塞 ( M I )、末梢動脈疾患 ( P A D )、脳卒中、肺塞栓症、深部静脈血栓症 ( D V T )、ならびに出血性素因及び凝固障害も含まれる。心血管病態には、アテローム血栓性病態が含まれ得る。アテローム血栓性病態には、フィブリノーゲンレベルの上昇が含まれ得る。アテローム血栓性病態には、フィブリノーゲン媒介性血餅が含まれ得る。心血管病態には、フィブリノーゲンレベルの上昇が含まれ得る。心血管病態には、フィブリノーゲン媒介性血餅が含まれ得る。心血管病態には、フィブリノーゲン活性の関与から形成される血餅が含まれ得る。フィブリノーゲン媒介性血餅またはフィブリノーゲン活性の関与から形成される血餅は、体内のどの静脈または動脈にも存在し得る。

30

#### 【 0 1 5 0 】

いくつかの実施形態では、心血管病態の発症に対するヒト対象の感受性の判定方法は：a ) ヒト対象由来の生体試料に対して、その生体試料中の核酸分子が、完全長 / 成熟型のバリアント型 B 4 G A L T 1 A s n 3 5 2 S e r ポリペプチドの 3 5 2 位に対応する位置にセリンをコードする核酸配列を含むかどうかを判定するアッセイを実施し、b ) 完全

40

50

長 / 成熟型のバリアント型 B 4 G A L T 1 A s n 3 5 2 S e r ポリペプチドの 3 5 2 位にセリンをコードする核酸配列を含む核酸分子が生体試料中で検出される場合、ヒト対象を、心血管病態を発症するリスクが低いと分類し、または完全長 / 成熟型のバリアント型 B 4 G A L T 1 A s n 3 5 2 S e r ポリペプチドの 3 5 2 位にセリンをコードする核酸配列を含む核酸分子が生体試料中で検出されない場合、ヒト対象を、心血管病態を発症するリスクが高いと分類することを含む。いくつかの実施形態では、バリアント型 B 4 G A L T 1 A s n 3 5 2 S e r ポリペプチドは、配列番号 8 を含む。いくつかの実施形態では、生体試料中の核酸分子は、ゲノム D N A 、 m R N A 、または c D N A である。

#### 【 0 1 5 1 】

いくつかの実施形態では、本開示は、心血管病態の発症に対するヒト対象の感受性の判定方法を提供し、方法は： a ) ヒト対象由来の生体試料に対して、その生体試料中の核酸分子が配列番号 2 の 5 3 7 5 7 ~ 5 3 5 7 7 位に対応する位置に配列番号 2 のヌクレオチド 5 3 7 5 7 ~ 5 3 5 7 7 を含むかどうかを判定するアッセイを実施し、 b ) 配列番号 2 の 5 3 7 5 7 ~ 5 3 5 7 7 位に対応する位置に配列番号 2 のヌクレオチド 5 3 7 5 7 ~ 5 3 5 7 7 を含む核酸分子が生体試料中で検出される場合、ヒト対象を、心血管病態を発症するリスクが低いと分類し、または配列番号 2 の 5 3 7 5 7 ~ 5 3 5 7 7 位に対応する位置に配列番号 2 のヌクレオチド 5 3 7 5 7 ~ 5 3 5 7 7 を含む核酸分子が生体試料中で検出されない場合、ヒト対象を、心血管病態を発症するリスクが高いと分類することを含む。

10

#### 【 0 1 5 2 】

いくつかの実施形態では、本開示は、心血管病態の発症に対するヒト対象の感受性の判定方法を提供し、方法は： a ) ヒト対象由来の生体試料に対して、その生体試料中の核酸分子が配列番号 4 の 1 2 4 3 ~ 1 2 4 5 位に対応する位置に配列番号 4 のヌクレオチド 1 2 4 3 ~ 1 2 4 5 を含むかどうかを判定するアッセイを実施し、 b ) 配列番号 4 の 1 2 4 3 ~ 1 2 4 5 位に対応する位置に配列番号 4 のヌクレオチド 1 2 4 3 ~ 1 2 4 5 を含む核酸分子が生体試料中で検出される場合、ヒト対象を、心血管病態を発症するリスクが低いと分類し、または配列番号 4 の 1 2 4 3 ~ 1 2 4 5 位に対応する位置に配列番号 4 のヌクレオチド 1 2 4 3 ~ 1 2 4 5 を含む核酸分子が生体試料中で検出されない場合、ヒト対象を、心血管病態を発症するリスクが高いと分類することを含む。

20

#### 【 0 1 5 3 】

いくつかの実施形態では、本開示は、心血管病態の発症に対するヒト対象の感受性の判定方法を提供し、方法は： a ) ヒト対象由来の生体試料に対して、その生体試料中の核酸分子が配列番号 6 の 1 0 5 4 ~ 1 0 5 6 位に対応する位置に配列番号 6 のヌクレオチド 1 0 5 4 ~ 1 0 5 6 を含むかどうかを判定するアッセイを実施し、 b ) 配列番号 6 の 1 0 5 4 ~ 1 0 5 6 位に対応する位置に配列番号 6 のヌクレオチド 1 0 5 4 ~ 1 0 5 6 を含む核酸分子が生体試料中で検出される場合、ヒト対象を、心血管病態を発症するリスクが低いと分類し、または配列番号 6 の 1 0 5 4 ~ 1 0 5 6 位に対応する位置に配列番号 6 のヌクレオチド 1 0 5 4 ~ 1 0 5 6 を含む核酸分子が生体試料中で検出されない場合、ヒト対象を、心血管病態を発症するリスクが高いと分類することを含む。

30

#### 【 0 1 5 4 】

いくつかの実施形態では、本方法は、生体試料中のバリアント型 B 4 G A L T 1 ゲノム D N A の存在を検出することを含む。いくつかの実施形態では、そのような方法は、対象の心血管病態の発症に対する感受性または心血管病態の発症リスクを判定することを含み、それは： a ) 対象からゲノム D N A を含む生体試料を取得し、 b ) バリアント型 B 4 G A L T 1 遺伝子（例えば、配列番号 2 参照）の 5 3 5 7 5 ~ 5 3 5 7 7 位に対応する D N A の位置を占有するヌクレオチドの同一性を判定するアッセイをゲノム D N A に対して実施し、 c ) バリアント型 B 4 G A L T 1 遺伝子の 5 3 5 7 5 ~ 5 3 5 7 7 位に対応するゲノム D N A の位置がアスパラギンではなくセリンをコードしている場合、対象を、心血管病態を発症するリスクが低いと分類することを含む。あるいは、バリアント型 B 4 G A L T 1 遺伝子の 5 3 5 7 5 ~ 5 3 5 7 7 位に対応するゲノム D N A の位置がアスパラギンではなくセリンをコードしていない場合、対象を、心血管病態を発症するリスクが高いと分

40

50

類することができる。

【0155】

いくつかの実施形態では、そのような方法は、心血管病態を有する対象を診断することを含み、それは：a) 対象からゲノムDNAを含む生体試料を取得し、b) バリアント型B4GALT1遺伝子（例えば、配列番号2参照）の53575～53577位に対応するDNAの位置を占有するヌクレオチドの同一性を判定するアッセイをゲノムDNAに対して実施し、c) バリアント型B4GALT1遺伝子の53575～53577位に対応するゲノムDNAの位置がアスパラギンではなくセリンをコードしている場合、対象を、心血管病態を有していると分類することを含む。あるいは、バリアント型B4GALT1遺伝子の53575～53577位に対応するゲノムDNAの位置がアスパラギンではなくセリンをコードしていない場合、対象を、心血管病態を有していないと分類することができる。

10

【0156】

いくつかの実施形態では、方法は、生体試料中のバリアント型B4GALT1 mRNAの存在を検出することを含む。いくつかの実施形態では、そのような方法は、対象の心血管病態の発症に対する感受性または心血管病態の発症リスクを判定することを含み、それは：a) 対象からmRNAを含む生体試料を取得し、b) バリアント型B4GALT1 mRNA（例えば、配列番号4参照）の1243～1245位に対応するmRNAの位置を占有するヌクレオチドの同一性を判定するアッセイをmRNAに対して実施し、c) バリアント型B4GALT1 mRNAの1243～1245位に対応するmRNAの位置がアスパラギンではなくセリンをコードしている場合、対象を、心血管病態を発症するリスクが低いと分類することを含む。あるいは、バリアント型B4GALT1 mRNAの1243～1245位に対応するmRNAの位置がアスパラギンではなくセリンをコードしていない場合、対象を、心血管病態を発症するリスクが高いと分類することができる。

20

【0157】

いくつかの実施形態では、そのような方法は、心血管病態を有する対象を診断することを含み、それは：a) 対象からmRNAを含む生体試料を取得し、b) バリアント型B4GALT1 mRNA（例えば、配列番号4参照）の1243～1245位に対応するmRNAの位置を占有するヌクレオチドの同一性を判定するアッセイをmRNAに対して実施し、c) バリアント型B4GALT1 mRNAの1243～1245位に対応するmRNAの位置がアスパラギンではなくセリンをコードしている場合、対象を、心血管病態を有していると分類することを含む。あるいは、バリアント型B4GALT1 mRNAの1243～1245位に対応するmRNAの位置がアスパラギンではなくセリンをコードしていない場合、対象を、心血管病態を有していないと分類することができる。

30

【0158】

いくつかの実施形態では、方法は、生体試料中のバリアント型B4GALT1 cDNAの存在を検出することを含む。いくつかの実施形態では、そのような方法は、対象の心血管病態の発症に対する感受性または心血管病態の発症リスクを判定することを含み、それは：a) 対象からcDNAを含む生体試料を取得し、b) バリアント型B4GALT1 cDNA（例えば、配列番号6参照）の1054～1056位に対応するcDNAの位置を占有するヌクレオチドの同一性を判定するアッセイをcDNAに対して実施し、c) バリアント型B4GALT1 cDNAの1054～1056位に対応するcDNAの位置がアスパラギンではなくセリンをコードしている場合、対象を、心血管病態を発症するリスクが低いと分類することを含む。あるいは、バリアント型B4GALT1 cDNAの1054～1056位に対応するcDNAの位置がアスパラギンではなくセリンをコードしていない場合、対象を、心血管病態を発症するリスクが高いと分類することができる。

40

【0159】

いくつかの実施形態では、そのような方法は、心血管病態を有する対象を診断することを含み、それは：a) 対象からcDNAを含む生体試料を取得し、b) バリアント型B4GALT1 cDNA（例えば、配列番号6参照）の1054～1056位に対応するc

50

D N A の位置を占有するヌクレオチドの同一性を判定するアッセイを c D N A に対して実施し、c ) バリアント型 B 4 G A L T 1 c D N A の 1 0 5 4 ~ 1 0 5 6 位に対応する c D N A の位置がアスパラギンではなくセリンをコードしている場合、対象を、心血管病態を有していると分類することを含む。あるいは、バリアント型 B 4 G A L T 1 c D N A の 1 0 5 4 ~ 1 0 5 6 位に対応する c D N A の位置がアスパラギンではなくセリンをコードしていない場合、対象を、心血管病態を有していないと分類することができる。

【 0 1 6 0 】

いくつかの実施形態では、アッセイは：ヒト対象由来の生体試料中の核酸分子の B 4 G A L T 1 ゲノム配列の部分を配列決定することを含み、配列決定する部分には、配列番号 2 の 5 3 5 7 5 ~ 5 3 5 7 7 位に対応する位置が含まれ、ヒト対象由来の生体試料中の核酸分子の B 4 G A L T 1 m R N A 配列の部分を配列決定することを含み、配列決定する部分には、配列番号 4 の 1 2 4 3 ~ 1 2 4 5 位に対応する位置が含まれ、または、ヒト対象由来の生体試料中の核酸分子の B 4 G A L T 1 c D N A 配列の部分を配列決定することを含み、配列決定する部分には、配列番号 6 の 1 0 5 4 ~ 1 0 5 6 位に対応する位置が含まれる。

【 0 1 6 1 】

いくつかの実施形態では、アッセイは：a ) 生体試料を：i ) 配列番号 2 の 5 3 5 7 5 ~ 5 3 5 7 7 位に対応する B 4 G A L T 1 ゲノム配列の位置に近接する B 4 G A L T 1 ゲノム配列の部分、i i ) 配列番号 4 の 1 2 4 3 ~ 1 2 4 5 位に対応する B 4 G A L T 1 m R N A の位置に近接する B 4 G A L T 1 m R N A 配列の部分、または i i i ) 配列番号 6 の 1 0 5 4 ~ 1 0 5 6 位に対応する B 4 G A L T 1 c D N A の位置に近接する B 4 G A L T 1 c D N A 配列の部分にハイブリダイズするプライマーと接触させること、b ) 少なくとも：i ) 5 3 5 7 5 ~ 5 3 5 7 7 位に対応する B 4 G A L T 1 ゲノム配列の位置、i i ) 1 2 4 3 ~ 1 2 4 5 位に対応する B 4 G A L T 1 m R N A の位置、または i i i ) 1 0 5 4 ~ 1 0 5 6 位に対応する B 4 G A L T 1 c D N A の位置を介してプライマーを伸長すること、及び c ) プライマーの伸長産物が、配列番号 8 の 3 5 2 位のセリンをコードするヌクレオチドを：i ) B 4 G A L T 1 ゲノム配列の 5 3 5 7 5 ~ 5 3 5 7 7 位に対応する、i i ) B 4 G A L T 1 m R N A の位 1 2 4 3 ~ 1 2 4 5 位に対応する、または i i i ) B 4 G A L T 1 c D N A の 1 0 5 4 ~ 1 0 5 6 位に対応する位置に含むかどうかを判定することを含む。

【 0 1 6 2 】

いくつかの実施形態では、アッセイは、ストリンジエントな条件下で、バリアント型 B 4 G A L T 1 ゲノム配列、m R N A 配列、または c D N A 配列に特異的にハイブリダイズし、対応する野生型 B 4 G A L T 1 配列にはハイブリダイズしないプライマーまたはプローブに生体試料を接触させること、及びハイブリダイゼーションが生じたかどうかを判定することを含む。いくつかの実施形態では、プライマーまたはプローブは、生体試料中で、配列番号 2 の 5 3 5 7 5 ~ 5 3 5 7 7 位に対応するゲノム D N A 内の位置に特異的にハイブリダイズする。いくつかの実施形態では、プライマーまたはプローブは、生体試料中で、配列番号 4 の 1 2 4 3 ~ 1 2 4 5 位に対応する m R N A 内の位置に特異的にハイブリダイズする。いくつかの実施形態では、プライマーまたはプローブは、生体試料中で、配列番号 6 の 1 0 5 4 ~ 1 0 5 6 位に対応する c D N A 内の位置に特異的にハイブリダイズする。

【 0 1 6 3 】

本明細書に開示する方法で使用することができる他のアッセイとして、例えば、逆転写ポリメラーゼ連鎖反応 ( R T - P C R ) または定量的 R T - P C R ( q R T - P C R ) が挙げられる。本明細書に開示する方法で使用することができるさらに他のアッセイとして、例えば、R N A 配列決定 ( R N A - S e q ) ならびにそれに続く生体試料中のバリアント型 m R N A または c D N A の存在及び量の決定が挙げられる。

【 0 1 6 4 】

本開示はまた、心血管病態の発症に対するヒト対象の感受性の判定方法または心血管病

10

20

30

40

50

態を有する対象の診断方法を提供し、方法は：a)ヒト対象由来の生体試料に対して、その生体試料中のB 4 G A L T 1ポリペプチドが配列番号8の352位に対応する位置にセリンを含むかどうかを判定するアッセイを実施し、b)配列番号8の352位に対応する位置にセリンを含むB 4 G A L T 1ポリペプチドが生体試料中で検出される場合、ヒト対象を、心血管病態を発症するリスクが低いと分類し、または配列番号8の352位に対応する位置にセリンを含むB 4 G A L T 1ポリペプチドが生体試料中で検出されない場合、ヒト対象を、心血管病態を発症するリスクが高いと分類することを含む。いくつかの実施形態では、方法は、対象から生体試料を取得することをさらに含む。

【0165】

いくつかの実施形態では、対象が、心血管病態を有しているか、または心血管病態を発症するリスクが高いと診断されている場合、心血管病態を治療または予防する治療薬または予防薬を対象に投与する。あるいは、方法は、特にLDLレベルが高い患者及び/または血栓性事象があった患者または血栓性事象のリスクの高い患者において、臨床的により進行した心血管病態の段階への進行に関連する1つ以上の症状を予防または軽減するよう10に調整された治療薬を投与することをさらに含み得る。

【0166】

本開示はまた、ヌクレアーゼ剤、外来性ドナー配列、転写活性化因子、転写抑制因子、アンチセンスRNA、siRNA、及びshRNAなどのアンチセンス分子、B 4 G A L T 1ポリペプチドまたはその断片、ならびに組換えB 4 G A L T 1遺伝子またはB 4 G A L T 1ポリペプチドをコードする核酸を発現するための発現ベクターの任意の組み合わせの使用による細胞の改変方法を提供する。方法は、in vitro、ex vivo、またはin vivoで行うことができる。ヌクレアーゼ剤、外来性ドナー配列、転写活性化因子、転写抑制因子、アンチセンスRNA、siRNA、及びshRNAなどのアンチセンス分子、B 4 G A L T 1ポリペプチドまたはその断片、ならびに発現ベクターは、本明細書の他の場所で記載するように任意の形態及び任意の手段で、ならびにすべてまたは一部を任意の組み合わせで同時にまたは連続して細胞に導入することができる。いくつかの方法は、細胞の内在性B 4 G A L T 1遺伝子を変化させることのみを含む。いくつかの方法は、転写活性化因子または抑制因子の使用またはアンチセンスRNA、siRNA、及びshRNAなどのアンチセンス分子の使用によって内在性B 4 G A L T 1遺伝子の発現を変化させることのみを含む。いくつかの方法は、組換えB 4 G A L T 1遺伝子またはB 4 G A L T 1ポリペプチドもしくはその断片をコードする核酸を細胞に導入することのみを含む。いくつかの方法は、B 4 G A L T 1ポリペプチドまたはその断片(例えば、本明細書に開示するB 4 G A L T 1ポリペプチドまたはその断片のいずれかまたは任意の組み合わせ)を細胞に導入することのみを含む。他の方法は、細胞の内在性B 4 G A L T 1遺伝子を改変すること、及びB 4 G A L T 1ポリペプチドもしくはその断片、あるいは組換えB 4 G A L T 1遺伝子またはB 4 G A L T 1ポリペプチドもしくはその断片をコードする核酸を細胞に導入することの両方を含む。他の方法は、細胞の内在性B 4 G A L T 1遺伝子の発現を変化させること、及びB 4 G A L T 1ポリペプチドもしくはその断片、あるいは組換えB 4 G A L T 1遺伝子またはB 4 G A L T 1ポリペプチドもしくはその断片をコードする核酸を細胞に導入することの両方を含む。30

【0167】

本開示は、ヌクレアーゼ剤及び/または外来性ドナー配列の使用による、細胞(例えば、多能性細胞または分化細胞)内のゲノムの内在性B 4 G A L T 1遺伝子の改変方法を提供する。方法は、in vitro、ex vivo、またはin vivoで行うことができる。ヌクレアーゼ剤は、単独で、または外来性ドナー配列と組み合わせて使用することができる。あるいは、外来性ドナー配列を、単独で、またはヌクレアーゼ剤と組み合わせて使用することができる。

【0168】

二本鎖切断(DSB)に応答する修復は、主に2つの保存されたDNA修復経路を介して起こる：非相同末端結合(NHEJ)及び相同組換え(HR)(Kasperek &

10

20

30

40

50

Humphrey, Seminars in Cell & Dev. Biol., 2011, 22, 886-897 参照)。外来性ドナー配列によって媒介される標的核酸(例えば、内在性 B4GALT1 遺伝子)の修復には、2つのポリヌクレオチド間の遺伝情報の交換の任意のプロセスが含まれ得る。例えば、NHEJ はまた、切断末端と外来性ドナー配列の末端との直接ライゲーション(すなわち、NHEJ ベースの捕捉)を通して外来性ドナー配列の標的化された組み込みをもたらし得る。修復は、相同組換え修復(HDR)または相同組換え(HR)によっても生じ得る。HDR または HR は、ヌクレオチド配列の相同性が必要であり得る核酸修復の形態を含み、「標的」分子(すなわち、二本鎖切断を受けた分子)の修復のテンプレートとして「ドナー」分子を使用し、ドナーからターゲットへの遺伝情報の転送をもたらす。

10

## 【0169】

ゲノム中の内在性 B4GALT1 遺伝子に対する標的化された遺伝子改変は、内在性 B4GALT1 遺伝子内の標的ゲノム遺伝子座で 5' 標的配列にハイブリダイズする 5' 相同性アーム及び内在性 B4GALT1 遺伝子内の標的ゲノム遺伝子座で 3' 標的配列にハイブリダイズする 3' 相同性アームを含む外来性ドナー配列と細胞を接触させることにより生成することができる。外来性ドナー配列を、標的ゲノム遺伝子座と組み換えて、内在性 B4GALT1 遺伝子に対する標的化された遺伝子改変を生成することができる。一例として、5' 相同性アームは、配列番号 1 の 53575 ~ 53577 位に対応する位置の 5' の標的配列にハイブリダイズすることができ、3' 相同性アームは、配列番号 1 の 53575 ~ 53577 位に対応する位置の 3' の標的配列にハイブリダイズすることができる。そのような方法は、例えば、そこから産生される完全長 / 成熟型ポリペプチドの 352 位に対応する位置にセリンをコードするヌクレオチド配列を含む B4GALT1 遺伝子をもたらし得る。外来性ドナー配列の例を、本明細書の他の場所に開示する。

20

## 【0170】

例えば、ゲノム中の内在性 B4GALT1 遺伝子に対する標的化された遺伝子改変を、細胞または細胞のゲノムを Cas タンパク質及び内在性 B4GALT1 遺伝子の標的ゲノム遺伝子座内の 1 つ以上のガイド RNA 認識配列にハイブリダイズする 1 つ以上のガイド RNA と接触させることにより生成することができる。例えば、そのような方法は、細胞を Cas タンパク質及び内在性 B4GALT1 遺伝子内のガイド RNA 認識配列にハイブリダイズするガイド RNA と接触させることを含み得る。いくつかの実施形態では、ガイド RNA 認識配列は、配列番号 1 のエクソン 5 に対応する領域内に位置する。いくつかの実施形態では、ガイド RNA 認識配列は、配列番号 1 の 53575 ~ 53577 位に対応する位置を含み得るか、またはその近傍にある。例えば、ガイド RNA 認識配列は、配列番号 1 の 53575 ~ 53577 位に対応する位置の、約 1000 以内、約 500 以内、約 400 以内、約 300 以内、約 200 以内、約 100 以内、約 50 以内、約 45 以内、約 40 以内、約 35 以内、約 30 以内、約 25 以内、約 20 以内、約 15 以内、約 10 以内、または約 5 ヌクレオチド以内とすることができます。さらに別の例として、ガイド RNA 認識配列は、内在性 B4GALT1 遺伝子の開始コドンまたは内在性 B4GALT1 遺伝子の終止コドンを含むか、またはその近傍とすることができます。例えば、ガイド RNA 認識配列は、開始コドンまたは終止コドンの、約 10 以内、約 20 以内、約 30 以内、約 40 以内、約 50 以内、約 100 以内、約 200 以内、約 300 以内、約 400 以内、約 500 以内、または約 1,000 ヌクレオチド以内とすることができます。Cas タンパク質とガイド RNA は複合体を形成し、Cas タンパク質はガイド RNA 認識配列を切断する。Cas タンパク質による切断は、二本鎖切断または一本鎖切断(例えば、Cas タンパク質がニッカーゼの場合)を引き起こし得る。そのような方法は、例えば、配列番号 1 のエクソン 5 に対応する領域が破壊されるか、開始コドンが破壊されるか、終止コドンが破壊されるか、またはコード配列を欠失した内在性 B4GALT1 遺伝子をもたらし得る。本方法で使用することができる Cas(例えば、Cas9) タンパク質及びガイド RNA の例及びバリエーションを、本明細書の他の場所に記載する。

30

## 【0171】

40

50

いくつかの実施形態では、2つ以上のヌクレアーゼ剤を使用することができる。例えば、配列番号1のエクソン5に対応する領域内、または配列番号1の53575～53577位に対応する位置を含むかまたはそれに近接する領域内（例えば、配列番号1の53575～53577位に対応する位置の、約1000以内、約500以内、約400以内、約300以内、約200以内、約100以内、約50以内、約45以内、約40以内、約35以内、約30以内、約25以内、約20以内、約15以内、約10以内、または約5ヌクレオチド以内）のヌクレアーゼ認識配列をそれぞれ標的とする2つのヌクレアーゼ剤を使用することができる。別の例として、それぞれが開始コドンを含むかまたはそれに近接するヌクレアーゼ認識配列を標的とする2つ以上のヌクレアーゼ剤を使用することができる。別の例として、1つが開始コドンを含むかまたはそれに近接するヌクレアーゼ認識配列を標的とし、もう1つが終止コドンを含むかまたはそれに近接するヌクレアーゼ認識配列を標的とする2つのヌクレアーゼ剤を使用することができ、これらのヌクレアーゼ剤による切断により、2つのヌクレアーゼ認識配列間のコード領域を欠失させることができる。さらに別の例として、1つ以上（例えば2つ）が開始コドンを含むかまたはそれに近接するヌクレアーゼ認識配列を標的とし、1つ以上（例えば2つ）が終止コドンを含むかまたはそれに近接するヌクレアーゼ認識配列を標的とする3つ以上のヌクレアーゼ剤を使用することができ、それらのヌクレアーゼ剤による切断により、開始コドンを含むかまたはそれに近接するヌクレアーゼ認識配列と、終止コドンを含むかまたはそれに近接するヌクレアーゼ認識配列の間のコード領域を欠失させることができる。

## 【0172】

いくつかの実施形態では、細胞を、内在性B4GALT1遺伝子の標的ゲノム遺伝子座内の追加のガイドRNA認識配列にハイブリダイズする1つ以上の追加のガイドRNAとさらに接触させることができる。細胞を1つ以上の追加のガイドRNA（例えば、第2のガイドRNA認識配列にハイブリダイズする第2のガイドRNA）と接触させることにより、Casタンパク質による切断で2つ以上の二本鎖切断または2つ以上の一一本鎖切断（例えば、Casタンパク質がニッカーゼの場合）を生じさせることができる。

## 【0173】

いくつかの実施形態では、内在性B4GALT1遺伝子の標的ゲノム遺伝子座と組み換えて、標的化された遺伝子改変を生成することができる1つ以上の外来性ドナー配列と、細胞をさらに接触させる。本方法で使用することができる外来性ドナー配列の例及びバリエーションを、本明細書の他の場所に開示する。

## 【0174】

本明細書の他の場所に記載するような任意の形態及び任意の手段により、Casタンパク質、ガイドRNA（複数可）、及び外来性ドナー配列（複数可）を細胞に導入することができ、Casタンパク質、ガイドRNA（複数可）、及び外来性ドナー配列（複数可）のすべてまたは一部を、任意の組み合わせで同時にまたは連続して導入することができる。

## 【0175】

いくつかの実施形態では、外来性ドナー配列による標的核酸（例えば、内在性B4GALT1遺伝子）の修復は、相同組換え修復（HDR）を介して生じる。相同組換え修復は、Casタンパク質が内在性B4GALT1遺伝子のDNAの両方の鎖を切断して二本鎖切断を生成する場合、Casタンパク質が標的核酸のDNAの1本の鎖を切断して一本鎖切断を生成するニッカーゼである場合、またはCasニッカーゼを使用して2つのオフセットニックによって形成される2本鎖切断を生成する場合に生じ得る。そのような方法では、外来性ドナー配列に、5'及び3'標的配列に対応する5'及び3'相同性アームを含ませる。ガイドRNA認識配列（複数可）または切断部位（複数可）は、5'標的配列の近傍のもの、3'標的配列の近傍のもの、5'標的配列と3'標的配列の両方の近傍のもの、または5'標的配列にも3'標的配列にも近接させないものとすることができます。いくつかの実施形態では、外来性ドナー配列に、5'及び3'相同性アームに隣接する核酸インサートをさらに含ませることができ、核酸インサートを5'及び3'標的配列の間に挿入する。核酸インサートが存在しない場合、外来性ドナー配列は、5'及び3'標的配列間のゲノム配列を

10

20

30

40

50

欠失させるように機能することができる。外来性ドナー配列の例を、本明細書の他の場所に開示する。

【0176】

あるいは、外来性ドナー配列によって媒介される内在性B4GALT1遺伝子の修復は、非相同末端結合(NHEJ)を介したライゲーションにより起こり得る。そのような方法では、外来性ドナー配列の少なくとも一方の末端は、内在性B4GALT1遺伝子のCasを介した切断により生じる少なくとも1つのオーバーハングに相補的な短い一本鎖領域を含む。外来性ドナー配列の相補的末端は、核酸インサートに隣接させることができる。例えば、外来性ドナー配列の各末端に、内在性B4GALT1遺伝子のCasを介した切断により生じるオーバーハングに相補的な短い一本鎖領域を含ませることができ、外来性ドナー配列のこれらの相補領域は核酸インサートに隣接させることができる。

10

【0177】

オーバーハング(すなわち、付着末端)は、Casを介した切断によって生じる二本鎖切断の平滑端の切除によって作成することができる。そのような切除は、断片の結合に必要なマイクロホモロジー領域を生成することができるが、これにより、B4GALT1遺伝子に望ましくないまたは制御不能な変化が生じ得る。あるいは、そのようなオーバーハングは、Casニッカーゼのペアで用いて作成することができる。例えば、細胞を、DNAの反対側の鎖を切断する第1及び第2ニッカーゼに接触させることができ、それによりゲノムを二重のニッキングによって変更する。これは、細胞を第1のCasタンパク質ニッカーゼ、内在性B4GALT1遺伝子の標的ゲノム遺伝子座内の第1のガイドRNA認識配列にハイブリダイズする第1のガイドRNA、第2のCasタンパク質ニッカーゼ、及び内在性B4GALT1遺伝子の標的ゲノム遺伝子座内の第2のガイドRNA認識配列にハイブリダイズする第2のガイドRNAと接触させることにより達成することができる。第1のCasタンパク質及び第1のガイドRNAは、第1の複合体を形成し、第2のCasタンパク質及び第2のガイドRNAは、第2の複合体を形成する。第1のCasタンパク質ニッカーゼは、第1のガイドRNA認識配列内のゲノムDNAの第1の鎖を切断し、第2のCasタンパク質ニッカーゼは、第2のガイドRNA認識配列内のゲノムDNAの第2の鎖を切断し、場合により、外来性ドナー配列が、内在性B4GALT1遺伝子の標的ゲノム遺伝子座と組み換わり、標的化された遺伝子変更を生じる。

20

【0178】

30

第1のニッカーゼは、ゲノムDNAの第1の鎖(すなわち、相補鎖)を切断することができ、第2のニッカーゼは、ゲノムDNAの第2の鎖(すなわち、非相補鎖)を切断することができる。第1及び第2のニッカーゼは、例えば、Cas9のRuvCドメインの触媒残基を変異させる(例えば、本明細書の他の場所に記載のD10A変異)か、またはCas9のHNHドメインの触媒残基を変異させる(例えば、本明細書の他の場所に記載のH840A変異)ことによって作成することができる。そのような方法では、二重のニッキングを用いて、付着末端(すなわち、オーバーハング)を有する二本鎖切断を作成することができる。切断部位を作成するように第1及び第2のガイドRNA認識配列を配置することができ、その結果、DNAの第1及び第2鎖上の第1及び第2のニッカーゼによって生じるニックにより、二本鎖切断が生じる。第1及び第2のCRISPR RNA認識配列内のニックをずらす場合には、オーバーハングが作成される。オフセットウインドウは、例えば、少なくとも約5bp、少なくとも約10bp、少なくとも約20bp、少なくとも約30bp、少なくとも約40bp、少なくとも約50bp、少なくとも約60bp、少なくとも約70bp、少なくとも約80bp、少なくとも約90bp、少なくとも約100bp、またはそれ以上であり得る。例えば、Ran et al., Cell, 2013, 154, 1380-1389, Mali et al., Nat. Biotech., 2013, 31, 833-838、及びShen et al., Nat. Methods, 2014, 11, 399-404参照。

40

【0179】

本明細書に記載の方法を使用して、様々なタイプの標的化された遺伝子変更を導入する

50

ことができる。そのような標的化された改変には、例えば、1つ以上のヌクレオチドの付加、1つ以上のヌクレオチドの欠失、1つ以上のヌクレオチドの置換、点突然変異、またはそれらの組み合わせが含まれ得る。例えば、少なくとも1、少なくとも2、少なくとも3、少なくとも4、少なくとも5、少なくとも7、少なくとも8、少なくとも9、もしくは少なくとも10、またはそれ以上のヌクレオチドを変化させ（例えば、欠失、挿入、または置換し）、標的化されたゲノム改変を形成することができる。

#### 【0180】

そのような標的化された遺伝子改変は、標的ゲノム遺伝子座の破壊をもたらし得る。破壊には、調節エレメント（プロモーターまたはエンハンサーなど）の変化、ミスセンス変異、ナンセンス変異、フレームシフト変異、トランケーション変異、ヌル変異、または少数のヌクレオチドの挿入または欠失（例えば、フレームシフト変異を引き起こすこと）が含まれ得るが、それは不活性化（すなわち機能の喪失）または対立遺伝子の喪失をもたらし得る。例えば、標的化された改変として、開始コドンがもはや機能しないようにするための内在性 B 4 G A L T 1 遺伝子の開始コドンの破壊が挙げられ得る。

10

#### 【0181】

いくつかの実施形態では、標的化された改変には、第1及び第2のガイドRNA認識配列またはCas切断部位間の欠失が含まれ得る。外来性ドナー配列（例えば、修復テンプレートまたはターゲティングベクター）を使用する場合、改変には、第1及び第2のガイドRNA認識配列またはCas切断部位間の欠失、ならびに5'及び3'標的配列間への核酸インサートの挿入が含まれ得る。

20

#### 【0182】

いくつかの実施形態では、外来性ドナー配列を単独でまたはヌクレアーゼ剤と組み合わせて使用する場合、改変には、5'及び3'標的配列間の欠失、ならびに第1及び第2の相同染色体のペアの5'及び3'標的配列間への核酸インサートの挿入、それにより、ホモ接合の改変ゲノムを生じることが含まれ得る。あるいは、外来性ドナー配列が核酸インサートのない5'及び3'相同性アームを有する場合、改変には5'及び3'標的配列間の欠失が含まれ得る。

#### 【0183】

第1及び第2のガイドRNA認識配列間の欠失または5'及び3'標的配列間の欠失は、欠失させる核酸が、第1及び第2のヌクレアーゼ切断部位間の核酸配列のみ、または5'及び3'標的配列間の核酸配列のみからなり、その結果、改変するゲノム標的遺伝子座に追加の欠失または挿入が存在しない正確な欠失であり得る。第1及び第2のガイドRNA認識配列間の欠失はまた、第1及び第2のヌクレアーゼ切断部位をはみ出す不正確な欠失であり得、それは非相同末端結合（NHEJ）による不正確な修復と一致し、改変するゲノム遺伝子座における追加の欠失及び/または挿入を生じる。例えば、欠失は、第1及び第2のCasタンパク質切断部位を、約1bp、約2bp、約3bp、約4bp、約5bp、約10bp、約20bp、約30bp、約40bp、約50bp、またはそれ以上はみ出しえる。同様に、改変するゲノム遺伝子座は、NHEJによる不正確な修復と一致する追加の挿入、例えば、約1bp、約2bp、約3bp、約4bp、約5bp、約10bp、約20bp、約30bp、約40bp、約50bp、またはそれ以上の挿入を含み得る。

30

#### 【0184】

標的化された遺伝子改変は、例えば、二対立遺伝子改変または単一対立遺伝子改変であり得る。二対立遺伝子の改変には、対応する相同染色体上の同じ遺伝子座に同じ改変を行うか（例えば、二倍体細胞の）、または対応する相同染色体上の同じ遺伝子座に異なる改変を行う事象が含まれる。いくつかの実施形態では、標的化された遺伝子改変は、単一対立遺伝子改変である。単一遺伝子改変には、1つの対立遺伝子のみに改変を行う事象（すなわち、2つの相同染色体のうちの1つのみの内在性 B 4 G A L T 1 遺伝子に対する改変）が含まれる。相同染色体には、同じ遺伝子座に同じ遺伝子を有するが、異なる対立遺伝

40

50

子である可能性のある染色体（例えば、減数分裂中にペアになる染色体）が含まれる。

【0185】

単一対立遺伝子変異は、標的化された B 4 G A L T 1 改変に関してヘテロ接合性である細胞をもたらし得る。ヘテロ接合性には、B 4 G A L T 1 遺伝子の 1 つの対立遺伝子（すなわち、両方の相同染色体上の対応する対立遺伝子）のみが標的化された改変を有する状況が含まれる。

【0186】

二対立遺伝子改変は、標的化された改変のホモ接合性をもたらし得る。ホモ接合性には、B 4 G A L T 1 遺伝子の両方の対立遺伝子（すなわち、両方の相同染色体上の対応する対立遺伝子）が標的化された改変を有する状況が含まれる。あるいは、二対立遺伝子改変は、標的化された改変に対して複合ヘテロ接合性（例えば、半接合性）をもたらし得る。複合ヘテロ接合性には、標的遺伝子座の両方の対立遺伝子（すなわち、両方の相同染色体上の対立遺伝子）が改変されるが、異なる方法で改変される状況（例えば、一方の対立遺伝子の標的化された改変及び他方の対立遺伝子の不活性化または破壊）が含まれる。

10

【0187】

本明細書に開示する方法は、改変された B 4 G A L T 1 遺伝子を有する細胞を同定することをさらに含み得る。欠失または挿入などの標的化された遺伝子改変を有する細胞を、様々な方法を使用して同定することができる。そのような方法は、B 4 G A L T 1 遺伝子に標的化された遺伝子改変を有する 1 つの細胞を同定することを含み得る。スクリーニングを実施して、改変されたゲノム遺伝子座を有するそのような細胞を同定することができる。スクリーニング工程は、親染色体の対立遺伝子の改変（M O A）を評価するための定量的アッセイ（例えば、対立遺伝子の喪失（L O A）及び／または対立遺伝子の獲得（G O A）アッセイ）を含み得る。

20

【0188】

適切な定量アッセイの他の例として、蛍光 in situ ハイブリダイゼーション（FISH）、比較ゲノムハイブリダイゼーション、等温DNA增幅、固定化プローブ（複数可）への定量ハイブリダイゼーション、IN VADER（登録商標）プローブ、T A Q M A N（登録商標）分子ビーコンプローブ、または E C L I P S E（商標）プローブ法が挙げられる。ロングレンジPCR、サザンプロットティング、またはサンガーフラスティングなどの標的修飾のスクリーニングのための従来のアッセイも使用することができる。そのようなアッセイは、通常、挿入されたターゲティングベクターと標的化されたゲノム遺伝子座との間の連結のエビデンスを得るために使用される。例えば、ロングレンジPCRアッセイでは、1つのプライマーが、挿入されたDNA内の配列を認識し、もう1つのプライマーが、ターゲティングベクターの相同性アームの末端をはみ出す標的ゲノム遺伝子座配列を認識する。

30

【0189】

次世代シーケンシング（NGS）もスクリーニングに使用することができる。次世代シーケンシングは「NGS」または「大規模並列シーケンシング」または「ハイスクローブットシーケンシング」とも呼ばれる。いくつかの実施形態では、選択マーカーを使用して標的細胞をスクリーニングする必要がない。例えば、本明細書に記載のM O A及びNGSアッセイは、選択カセットの使用に依拠しないことが可能である。

40

【0190】

本開示はまた、B 4 G A L T 1 ポリペプチドをコードする核酸の発現の変更方法を提供する。いくつかの実施形態では、本明細書の他の場所でさらに詳細に記載するように、ヌクレアーゼ剤による切断によって発現を変化させ、内在性 B 4 G A L T 1 ポリペプチドをコードする核酸の破壊を引き起こす。いくつかの実施形態では、転写活性化ドメインまたは転写抑制ドメインに融合または連結させたDNA結合タンパク質の使用により発現を変化させる。いくつかの実施形態では、アンチセンスRNA、s h RNA、またはs i RNAなどのRNA干渉組成物の使用により発現を変化させる。

【0191】

50

いくつかの実施形態では、内在性 B 4 G A L T 1 遺伝子または B 4 G A L T 1 ポリペプチドをコードする核酸の発現は、細胞または細胞内のゲノムを、内在性 B 4 G A L T 1 遺伝子または B 4 G A L T 1 ポリペプチドをコードする核酸内の標的ゲノム遺伝子座の認識配列に 1 つ以上のニックまたは二本鎖切断を誘導するヌクレアーゼ剤と接触させることにより改変することができる。そのような切断は、内在性 B 4 G A L T 1 遺伝子または B 4 G A L T 1 ポリペプチドをコードする核酸の発現の破壊をもたらし得る。例えば、ヌクレアーゼ認識配列は、内在性 B 4 G A L T 1 遺伝子の開始コドンを含むかまたはそれに近接させることができる。例えば、認識配列は、開始コドンの、約 10 以内、約 20 以内、約 30 以内、約 40 以内、約 50 以内、約 100 以内、約 200 以内、約 300 以内、約 400 以内、約 500 以内、または約 1,000 ヌクレオチド以内とすることができ、及びヌクレアーゼ剤による切断は、開始コドンを破壊することができる。いくつかの実施形態では、2 つ以上のヌクレアーゼ剤を使用することができ、それぞれが、開始コドンを含むかまたはそれに近接するヌクレアーゼ認識配列を標的とする。いくつかの実施形態では、2 つのヌクレアーゼ剤を使用することができ、1 つは開始コドンを含むかまたはそれに近接するヌクレアーゼ認識配列を標的とし、1 つは終止コドンを含むかまたはそれに近接するヌクレアーゼ認識配列を標的とし、ヌクレアーゼ剤による切断は、2 つのヌクレアーゼ認識配列間のコード領域の欠失をもたらし得る。いくつかの実施形態では、3 つ以上のヌクレアーゼ剤を使用することができ、1 つ以上（例えば 2 つ）は、開始コドンを含むかまたはそれに近接するヌクレアーゼ認識配列を標的とし、1 つ以上（例えば 2 つ）は、終止コドンを含むかまたはそれに近接するヌクレアーゼ認識配列を標的とし、ヌクレアーゼ剤による切断は、開始コドンを含むかまたはそれに近接するヌクレアーゼ認識配列と、終止コドンを含むかまたはそれに近接するヌクレアーゼ認識配列との間のコード領域の欠失をもたらし得る。内在性 B 4 G A L T 1 遺伝子または B 4 G A L T 1 ポリペプチドをコードする核酸を改変する他の例を、本明細書の他の場所に開示する。  
10

#### 【 0192 】

いくつかの実施形態では、内在性 B 4 G A L T 1 遺伝子または B 4 G A L T 1 ポリペプチドをコードする核酸の発現は、細胞または細胞内のゲノムを、内在性 B 4 G A L T 1 遺伝子内の標的ゲノム遺伝子座に結合する DNA 結合タンパク質と接触させることにより改変することができる。DNA 結合タンパク質は、例えば、転写活性化因子ドメインまたは転写抑制因子ドメインに融合したヌクレアーゼ不活性 Cas タンパク質であり得る。DNA 結合タンパク質の他の例として、転写活性化因子ドメインまたは転写抑制因子ドメインに融合したジンクフィンガータンパク質、または転写活性化因子ドメインまたは転写抑制因子ドメインに融合した T A L エフェクター ( T A L E ) タンパク質が挙げられる。そのようなタンパク質の例を、本明細書の他の場所に開示する。  
20

#### 【 0193 】

DNA 結合タンパク質の認識配列（例えば、ガイド RNA 認識配列）は、内在性 B 4 G A L T 1 遺伝子または B 4 G A L T 1 ポリペプチドをコードする核酸内の、発現を変化させるのに適した任意の場所とすることができます。いくつかの実施形態では、認識配列は、エンハンサーまたはプロモーターなどの調節エレメント内とすることができます、または調節エレメントに近接させることができます。例えば、認識配列は、内在性 B 4 G A L T 1 遺伝子の開始コドンを含むか、またはそれに近接させることができます。いくつかの実施形態では、認識配列は、開始コドンの、約 10 以内、約 20 以内、約 30 以内、約 40 以内、約 50 以内、約 100 以内、約 200 以内、約 300 以内、約 400 以内、約 500 以内、または約 1,000 ヌクレオチド以内とすることができます。  
30

#### 【 0194 】

いくつかの実施形態では、アンチセンス分子を使用して、内在性 B 4 G A L T 1 遺伝子または B 4 G A L T 1 ポリペプチドをコードする核酸の発現を変化させることができます。アンチセンス分子の例として、アンチセンス RNA、s i RNA、及び s h RNA が挙げられるが、これらに限定されない。そのようなアンチセンス RNA、s i RNA、または s h RNA は、m RNA の任意の領域を標的とするように設計することができる。例えば  
40

、アンチセンスRNA、siRNA、またはshRNAは、B4GALT1 mRNAに固有の領域を標的とするように設計することができる。

【0195】

本明細書に開示する核酸及びタンパク質は、任意の手段により細胞に導入することができる。いくつかの実施形態では、導入は任意の手段によって達成することができ、1つ以上の成分（例えば、2つの成分、またはすべての成分）を任意の組み合わせで同時にまたは連続して細胞に導入することができる。例えば、外来性ドナー配列を、ヌクレアーゼ剤の導入前に導入することができ、またはヌクレアーゼ剤の導入後に導入することができる（例えば、外来性ドナー配列を、ヌクレアーゼ剤の導入の、約1、約2、約3、約4、約8、約12、約24、約36、約48、または約72時間前または後に投与する）。細胞のゲノムをヌクレアーゼ剤または外来性ドナー配列と接触させることは、1つ以上のヌクレアーゼ剤またはヌクレアーゼ剤をコードする核酸（例えば、1つ以上のCasタンパク質または1つ以上のCasタンパク質をコードする核酸、及び1つ以上のガイドRNAまたは1つ以上のガイドRNAをコードする核酸（すなわち、1つ以上のCRISPR RNA及び1つ以上のtracrRNA））及び/または1つ以上の外来性ドナー配列を細胞内へ導入することを含み得る。細胞のゲノムとの接触（すなわち、細胞との接触）は、上記の成分のうちの1つのみ、1つ以上の成分、またはすべての成分を細胞に導入することを含み得る。

10

【0196】

ヌクレアーゼ剤は、タンパク質の形態で、またはヌクレアーゼ剤をコードする核酸、例えば、RNA（例えば、メッセンジャーRNA（mRNA））もしくはDNAの形態で細胞に導入することができる。DNAの形態で導入する場合、DNAを、細胞内で活性なプロモーターに作動可能に連結することができる。そのようなDNAを、1つ以上の発現構築物の中に存在させることができる。

20

【0197】

いくつかの実施形態では、Casタンパク質は、タンパク質、例えば、gRNAと複合化したCasタンパク質の形態で、またはCasタンパク質をコードする核酸、例えば、RNA（例えば、メッセンジャーRNA（mRNA））もしくはDNAの形態で細胞に導入することができる。ガイドRNAは、RNAの形態で、またはガイドRNAをコードするDNAの形態で細胞に導入することができる。DNAの形態で導入する場合、Casタンパク質及び/またはガイドRNAをコードするDNAを、細胞内で活性なプロモーターに作動可能に連結することができる。そのようなDNAを、1つ以上の発現構築物の中に存在させることができる。例えば、そのような発現構築物は、単一の核酸分子の成分であり得る。あるいは、それらは、2つ以上の核酸分子の間で任意の組み合わせで分離することができる（すなわち、1つ以上のCRISPR RNAをコードするDNA、1つ以上のtracrRNAをコードするDNA、及びCasタンパク質をコードするDNAを、別々の核酸分子の成分とすることができる）。

30

【0198】

いくつかの実施形態では、ヌクレアーゼ剤をコードするDNA（例えば、Casタンパク質及びガイドRNA）及び/または外来性ドナー配列をコードするDNAを、DNAミニサークルを介して細胞に導入することができる。DNAミニサークルは、非ウイルス性遺伝子導入に使用することができるスーパーコイルDNA分子であり、複製起点も抗生物質選択マーカーも有していない。したがって、DNAミニサークルは通常、プラスミドベクターよりもサイズが小さい。これらのDNAには細菌DNAが含まれていないため、細菌DNAに見出される非メチル化CpGモチーフがない。

40

【0199】

本明細書に記載の方法は、核酸またはタンパク質を細胞に導入するための特定の方法に依存せず、核酸またはタンパク質が少なくとも1つの細胞の内部にアクセスすることのみが得られる。核酸及びタンパク質を様々な細胞型に導入する方法は公知であり、安定的導入法、一過性導入法、及びウイルスを介した方法が挙げられるが、これらに限定されない。

50

## 【0200】

形質移入プロトコールならびに核酸またはタンパク質を細胞に導入するためのプロトコールは様々に異なり得る。非限定的な形質移入法として、リポソーム、ナノ粒子、カルシウム、デンドリマー、及びD E A E - デキストランまたはポリエチレンイミンなどのカチオン性ポリマーを使用した化学ベースの形質移入法が挙げられる。非化学的方法としては、電気穿孔法、ソノポレーション、及び光トランスフェクションが挙げられる。粒子ベースの形質移入法として、遺伝子銃の使用、または磁石補助型トランスフェクションが挙げられる。ウイルス法も形質移入に使用することができる。

## 【0201】

細胞への核酸またはタンパク質の導入は、電気穿孔法、細胞質内注射、ウイルス感染、アデノウイルス、アデノ随伴ウイルス、レンチウイルス、レトロウイルス、形質移入、脂質を介した形質移入、またはヌクレオフェクションによっても媒介され得る。ヌクレオフェクションは、核酸基質を細胞質だけでなく核膜を介して核内に送達できるようにする改良された電気穿孔法である。さらに、本明細書に開示する方法におけるヌクレオフェクションの使用は、一般的に、通常の電気穿孔法よりもはるかに少ない細胞しか必要としない（例えば、通常の電気穿孔法が700万であるのに対して約200万で済む）。いくつかの実施形態では、ヌクレオフェクションは、LONZA（登録商標）NUCLEOFECTOR（商標）系を用いて実施する。

10

## 【0202】

細胞への核酸またはタンパク質の導入は、マイクロインジェクションによって達成することもできる。mRNAのマイクロインジェクションは通常、細胞質内で実施し（例えば、mRNAを翻訳機構に直接送達するために）、一方、タンパク質、またはCasタンパク質をコードするDNAのマイクロインジェクションは通常、核内で実施する。あるいは、マイクロインジェクションを、核と細胞質の両方への注射により実施することができる：針を最初に核に導入し、第1の量を注射し、細胞から針を除去しながら第2の量を細胞質に注射することができる。ヌクレアーゼ剤タンパク質を細胞質に注射する場合、確実に核／前核へ送達するために、核局在化シグナルをタンパク質に含ませてもよい。

20

## 【0203】

核酸またはタンパク質を細胞に導入する他の方法として、例えば、ベクター送達、粒子を介した送達、エキソソームを介した送達、脂質ナノ粒子を介した送達、細胞透過ペプチドを介した送達、または埋め込み型装置を介した送達が挙げられる。核酸またはタンパク質を対象に投与してin vivoで細胞を変換する方法を、本明細書の他の場所に開示する。細胞への核酸及びタンパク質の導入は、流体力学的送達（HDD）によっても実現することができる。

30

## 【0204】

核酸またはタンパク質を細胞に導入する他の方法として、例えば、ベクター送達、粒子を介した送達、エキソソームを介した送達、脂質ナノ粒子を介した送達、細胞透過ペプチドを介した送達、または埋め込み型装置を介した送達が挙げられる。いくつかの実施形態では、核酸またはタンパク質を、ポリ（乳酸）（PLA）ミクロスフェア、ポリ（D,L-乳酸-コグリコール酸）（PLGA）ミクロスフェア、リポソーム、ミセル、逆ミセル、脂質コクリエート、または脂質微小管などの担体中で、細胞に導入することができる。

40

## 【0205】

細胞への核酸またはタンパク質の導入は、ある期間にわたって1回または複数回行うことができる。いくつかの実施形態では、導入は、ある期間にわたって少なくとも2回、ある期間にわたって少なくとも3回、ある期間にわたって少なくとも4回、ある期間にわたって少なくとも5回、ある期間にわたって少なくとも6回、ある期間にわたって少なくとも7回、ある期間にわたって少なくとも8回、ある期間にわたって少なくとも9回、ある期間にわたって少なくとも10回、少なくとも11回、ある期間にわたって少なくとも12回、ある期間にわたって少なくとも13回、ある期間にわたって少なくとも14回、ある期間にわたって少なくとも15回、ある期間にわたって少なくとも16回、ある期間に

50

わたって少なくとも 17 回、ある期間にわたって少なくとも 18 回、ある期間にわたって少なくとも 19 回、またはある期間にわたって少なくとも 20 回行うことができる。

【 0206 】

いくつかの実施形態では、方法及び組成物で使用する細胞は、それらのゲノムに安定的に組み込まれた DNA 構築物を有する。そのような場合、接触は、ゲノムに既に安定的に組み込まれた構築物を有する細胞を提供することを含み得る。いくつかの実施形態では、本明細書に開示する方法において使用する細胞は、そのゲノムに安定的に組み込まれた既存の Cas コード遺伝子を有していてもよい（すなわち、Cas 作動可能細胞）。いくつかの実施形態では、ポリヌクレオチドを細胞のゲノムに組み込み、それをその子孫に遺伝させることができる。DNA 構築物または標的化されたゲノム組み込み系の様々な成分の安定した組み込みに、任意のプロトコールを使用してもよい。

10

【 0207 】

所望の認識配列にニックまたは二本鎖切断を誘導する任意のヌクレアーゼ剤または所望の認識配列に結合する任意の DNA 結合タンパク質を、本明細書に開示する方法及び組成物において使用することができる。天然またはネイティブのヌクレアーゼ剤を、そのヌクレアーゼ剤が所望の認識配列にニックまたは二本鎖切断を誘導する限り、使用することができる。同様に、天然またはネイティブの DNA 結合タンパク質を、その DNA 結合タンパク質が所望の認識配列に結合する限り、使用することができる。あるいは、修飾型または改変型ヌクレアーゼ剤または DNA 結合タンパク質を使用することができる。改変型ヌクレアーゼ剤または DNA 結合タンパク質は、ネイティブの天然ヌクレアーゼ剤もしくは DNA 結合タンパク質に由来するものであるか、または人工的に作成もしくは合成することができる。改変型ヌクレアーゼ剤または DNA 結合タンパク質は、認識配列を認識することができ、例えば、その場合、認識配列は、ネイティブの（未改変または未修飾の）ヌクレアーゼ剤または DNA 結合タンパク質によって認識されるであろう配列ではない。ヌクレアーゼ剤または DNA 結合タンパク質の修飾は、タンパク質切断剤中のわずか 1 個のアミノ酸または核酸切断剤中のわずか 1 個のヌクレオチドとすることができます。

20

【 0208 】

ヌクレアーゼ剤の認識配列には、ヌクレアーゼ剤によりニックまたは二本鎖切断が誘導される DNA 配列が含まれる。同様に、DNA 結合タンパク質の認識配列には、DNA 結合タンパク質が結合する DNA 配列が含まれる。認識配列は、細胞にとって内在性（すなわちネイティブ）、または細胞にとって外来性であり得る。認識配列はまた、標的遺伝子座に配置することを望む目的のポリヌクレオチドに対して外来性であり得る。いくつかの実施形態では、認識配列は、宿主細胞のゲノム中に一箇所だけ存在する。

30

【 0209 】

例示する認識配列の活性型バリエント及び断片も提供する。そのような活性型バリエントは、所与の認識配列に対して、少なくとも 65 %、少なくとも 70 %、少なくとも 75 %、少なくとも 80 %、少なくとも 85 %、少なくとも 90 %、少なくとも 91 %、少なくとも 92 %、少なくとも 93 %、少なくとも 94 %、少なくとも 95 %、少なくとも 96 %、少なくとも 97 %、少なくとも 98 %、もしくは少なくとも 99 %、または 100 % の配列同一性を有し、活性型バリエントは生物活性を保持し、ヌクレアーゼ剤によって配列特異的な様式で認識され、切断され得る。ヌクレアーゼ剤による認識配列の二本鎖切断を測定するアッセイは、公知である（例えば、T A Q M A N (登録商標) q P C R アッセイ、F rendewey et al ., Methods in Enzymology , 2010 , 476 , 295 - 307 ）。

40

【 0210 】

認識配列の長さは様々に異なり得るが、例えば、ジンクフィンガータンパク質またはジンクフィンガーヌクレアーゼ (Z F N ) 対に対して約 30 ~ 約 36 bp (すなわち、各 Z F N に対して約 15 ~ 約 18 bp ) 、 T A L E タンパク質または T A L エフェクターヌクレアーゼ (T A L E N ) に対して約 36 bp 、または C R I S P R / Cas 9 ガイド RNA に対して約 20 bp の認識配列が挙げられる。

50

## 【0211】

DNA結合タンパク質またはヌクレアーゼ剤の認識配列は、標的ゲノム遺伝子座の内部または近傍の任意の場所に配置することができる。認識配列は、遺伝子（例えば、B4GALT1遺伝子）のコード領域内、または遺伝子の発現に影響を及ぼす調節領域内に配置することができる。DNA結合タンパク質またはヌクレアーゼ剤の認識配列は、イントロン、エクソン、プロモーター、エンハンサー、調節領域、または任意の非タンパク質コード領域に配置することができる。

## 【0212】

本明細書に開示する様々な方法及び組成物において使用することができるDNA結合タンパク質の1つのタイプは、TALEである。TALEは、例えば、エピジェネティック修飾ドメイン、転写活性化ドメイン、または転写抑制因子ドメインに融合させるか、または連結することができる。そのようなドメインの例は、Casタンパク質に関して以下に記載されており、例えば、PCT公開WO2011/145121にも見出すことができる。それに対応して、本明細書に開示する様々な方法及び組成物において使用することができるヌクレアーゼ剤の1つのタイプはTALENである。転写アクチベーター様（TAL）エフェクターヌクレアーゼは、原核生物または真核生物のゲノムの特定の標的配列において二本鎖切断を行うために使用することができる配列特異的ヌクレアーゼのクラスである。TALエフェクターヌクレアーゼは、FokIなどのエンドヌクレアーゼの触媒ドメインに、ネイティブ型もしくは変型TALエフェクター、またはその機能部分を融合することにより作成する。ユニークなモジュラー型TALエフェクターDNA結合ドメインにより、潜在的に任意のDNA認識特異性を有するタンパク質の設計が可能になる。したがって、TALエフェクターヌクレアーゼのDNA結合ドメインを、特定のDNA標的部位を認識するように設計し、したがって、所望の標的配列において二本鎖切断を行うために使用することができる。適切なTALヌクレアーゼの例、及び適切なTALヌクレアーゼの調製方法は、例えば、米国特許出願公開第2011/0239315号、第2011/0269234号、第2011/0145940号、第2003/0232410号、第2005/0208489号、第2005/0026157号、第2005/0064474号、第2006/0188987号、及び第2006/0063231号に開示されている。

## 【0213】

いくつかのTALENでは、TALENの各モノマーは、2つの超可変残基を介して単一の塩基対を認識する約33～約35のTALリピートを含む。いくつかのTALENでは、ヌクレアーゼ剤は、FokIエンドヌクレアーゼなどの独立したヌクレアーゼに作動可能に連結したTALリピートベースのDNA結合ドメインを含むキメラタンパク質である。例えば、ヌクレアーゼ剤は、第1のTALリピートベースのDNA結合ドメイン及び第2のTALリピートベースのDNA結合ドメインを含むことができ、第1及び第2のTALリピートベースのDNA結合ドメインのそれぞれは、FokIヌクレアーゼに作動可能に連結し、第1及び第2のTALリピートベースのDNA結合ドメインは、様々な長さ（約12～約20bp）のスペーサー配列によって分離された標的DNA配列の各鎖中の2つの隣接する標的DNA配列を認識し、FokIヌクレアーゼサブユニットは二量体化して、標的配列において二本鎖切断を行う活性型ヌクレアーゼを生成する。

## 【0214】

DNA結合タンパク質の別の例は、ジンクフィンガータンパク質である。そのようなジンクフィンガータンパク質は、例えば、エピジェネティック修飾ドメイン、転写活性化ドメイン、または転写抑制因子ドメインに連結するか、または融合させることができる。そのようなドメインの例は、Casタンパク質に関して以下に記載されており、例えば、PCT公開WO2011/145121にも見出すことができる。それに対応して、本明細書に開示する様々な方法及び組成物において使用することができるヌクレアーゼ剤の別の例はZFNである。いくつかのZFNでは、ZFNの各モノマーは3つ以上のジンクフィンガーベースのDNA結合ドメインを含み、各ジンクフィンガーベースのDNA結合ドメ

10

20

30

40

50

インは 3 b p の部分部位に結合する。他の ZFN では、Fok I エンドヌクレアーゼなどの独立したヌクレアーゼに作動可能に連結したジンクフィンガーベースの DNA 結合ドメインを含むキメラタンパク質である。例えば、ヌクレアーゼ剤は第 1 の ZFN 及び第 2 の ZFN を含むことができ、第 1 の ZFN 及び第 2 の ZFN のそれぞれは Fok I ヌクレアーゼサブユニットに作動可能に連結し、第 1 及び第 2 の ZFN は、約 5 ~ 約 7 b p のスペーサーにより分離された標的 DNA 配列の各鎖において 2 つの隣接する標的 DNA 配列を認識し、Fok I ヌクレアーゼサブユニットは二量体化して、二本鎖切断を行う活性型ヌクレアーゼを生成する。

【 0215 】

本明細書に記載の方法及び組成物において使用するための他の適切な DNA 結合タンパク質及びヌクレアーゼ剤として、本明細書の他の場所に記載する CRISPR - Cas 系が挙げられる。

10

【 0216 】

DNA 結合タンパク質またはヌクレアーゼ剤は、任意の公知の手段によって細胞に導入してもよい。DNA 結合タンパク質またはヌクレアーゼ剤をコードするポリペプチドを細胞に直接導入してもよい。あるいは、DNA 結合タンパク質またはヌクレアーゼ剤をコードするポリヌクレオチドを細胞に導入することができる。DNA 結合タンパク質またはヌクレアーゼ剤をコードするポリヌクレオチドを細胞に導入する場合、細胞内で DNA 結合タンパク質またはヌクレアーゼ剤を、一時的に、条件的に、または構成的に発現させることができる。例えば、DNA 結合タンパク質またはヌクレアーゼ剤をコードするポリヌクレオチドを発現力セットに含め、条件的プロモーター、誘導性プロモーター、構成的プロモーター、または組織特異的プロモーターに作動可能に連結することができる。そのようなプロモーターを、本明細書の他の場所でさらに詳細に論じる。いくつかの実施形態では、DNA 結合タンパク質またはヌクレアーゼ剤を、DNA 結合タンパク質またはヌクレアーゼ剤をコードする mRNA として細胞に導入することができる。

20

【 0217 】

DNA 結合タンパク質またはヌクレアーゼ剤をコードするポリヌクレオチドを、細胞のゲノムに安定的に組み込み、また、細胞内で活性なプロモーターに作動可能に連結することができる。あるいは、DNA 結合タンパク質またはヌクレアーゼ剤をコードするポリヌクレオチドを、ターゲティングベクター内、またはインサートのポリヌクレオチドを含むターゲティングベクターとは別のベクターもしくはプラスミド内に存在させることができる。

30

【 0218 】

DNA 結合タンパク質またはヌクレアーゼ剤をコードするポリヌクレオチドの導入を介して DNA 結合タンパク質またはヌクレアーゼ剤を細胞に提供する場合、DNA 結合タンパク質またはヌクレアーゼ剤をコードするそのようなポリヌクレオチドを、DNA 結合タンパク質またはヌクレアーゼ剤をコードする天然のポリヌクレオチド配列に比べて目的細胞内での使用頻度が高い置換コドンに改変することができる。いくつかの実施形態では、DNA 結合タンパク質またはヌクレアーゼ剤をコードするポリヌクレオチドを、細菌細胞、酵母細胞、ヒト細胞、非ヒト細胞、哺乳類細胞、げっ歯類細胞、マウス細胞、ラット細胞または任意の他の目的の宿主細胞を含む、所定の目的の原核細胞または真核細胞において天然のポリヌクレオチド配列に比べて使用頻度の高い置換コドンに改変することができる。

40

【 0219 】

本明細書に開示する方法は、細胞内のゲノムを改変するために、Clustered Regularly Interspersed Short Palindromic Repeats (CRISPR) / CRISPR 関連 (Cas) 系またはそのような系の構成要素を利用することができる。CRISPR - Cas 系は、Cas 遺伝子の発現に関与するか、またはその活性を誘導する転写産物及び他のエレメントを含む。CRISPR - Cas 系は、I 型、II 型、または III 型の系であり得る。あるいは、CRISPR

50

/ Cas 系は、例えば、V型の系（例えば、サブタイプV - A またはサブタイプV - B ）であり得る。本明細書に開示する方法及び組成物は、核酸の部位特異的切断のために CRISPR 複合体（Cas タンパク質と複合体化したガイド RNA (gRNA) を含む）を利用することにより CRISPR - Cas 系を使用することができる。

#### 【0220】

本明細書に開示する方法において使用する CRISPR - Cas 系は、非天然である。例えば、いくつかの CRISPR - Cas 系では、天然では一緒に存在しない gRNA と Cas タンパク質を含む非天然の CRISPR 複合体を使用する。

#### 【0221】

Cas タンパク質は、一般的に、ガイド RNA (gRNA、以下により詳細に記載する) と相互作用することができる少なくとも 1 つの RNA 認識ドメインまたは RNA 結合ドメインを含む。Cas タンパク質はまた、ヌクレアーゼドメイン（例えば、DNase または RNase ドメイン）、DNA 結合ドメイン、ヘリカーゼドメイン、タンパク質 - タンパク質相互作用ドメイン、二量体化ドメイン、及び他のドメインを含み得る。ヌクレアーゼドメインは、核酸分子の共有結合の切断を含む、核酸切断のための触媒活性を保有している。切断により、平滑末端または付着末端を生成することができ、それは一本鎖または二本鎖であり得る。野生型 Cas9 タンパク質は、通常、平滑切断産物を生成する。あるいは、野生型 Cpf1 タンパク質（例えば、FnCpf1）は、非標的鎖の PAM 配列から 18 番目の塩基対の後、及び標的鎖の 23 番目の塩基の後に切断が生じた 5' ヌクレオチドの 5' オーバーハングを有する切断産物をもたらし得る。Cas タンパク質は、内在性 B4GALT1 遺伝子に二本鎖切断（例えば、平滑末端を有する二本鎖切断）を生じる完全な切断活性を有しているか、または内在性 B4GALT1 遺伝子に一本鎖切断を生じるニッカーゼであり得る。

10

20

30

#### 【0222】

Cas タンパク質の例として、Cas1、Cas1B、Cas2、Cas3、Cas4、Cas5、Cas5e (CasD)、Cas6、Cas6e、Cas6f、Cas7、Cas8a1、Cas8a2、Cas8b、Cas8c、Cas9 (Csn1 または Csx12)、Cas10、Cas10d、CasF、CasG、CasH、CasY1、CasY2、CasY3、CasE1 (CasA)、CasE2 (CasB)、CasE3 (CasE)、CasE4 (CasC)、CasC1、CasC2、CasA5、Csn2、Csm2、Csm3、Csm4、Csm5、Csm6、Cmr1、Cmr3、Cmr4、Cmr5、Cmr6、CasB1、CasB2、CasB3、Csx17、Csx14、Csx10、Csx16、CasA X、Csx3、Csx1、Csx15、CasF1、CasF2、CasF3、CasF4、及び Cas1966、ならびにその相同体または変形体が挙げられるが、これらに限定されない。

40

#### 【0223】

いくつかの実施形態では、Cas タンパク質は、Cas9 タンパク質であるか、または II 型 CRISPR - Cas 系の Cas9 タンパク質に由来する。Cas9 タンパク質は、II 型 CRISPR - Cas 系に由来し、通常、保存されたアーキテクチャを有する 4 つの重要なモチーフを共有している。モチーフ 1、2、及び 4 は RuvC 様モチーフであり、モチーフ 3 は HNH モチーフである。例示的な Cas9 タンパク質として、*Streptococcus pyogenes*、*Streptococcus thermophilus*、*Streptococcus* 種、*Staphylococcus aureus*、*Nocardiosis dassonvillei*、*Streptomyces pristinaespiralis*、*Streptomyces viridochromogenes*、*Streptomyces viridochromogenes*、*Streptosporangium roseum*、*Streptosporangium roseum*、*Alicyclobacillus acidocaldarius*、*Bacillus pseudomycoides*、*Bacillus selenite*、*Exiguobacterium sibiricum*、*Lact*

50

obacillus delbrueckii、Lactobacillus salivarius、Microscilla marina、Burkholderiales bacterium、Polaromonas naphthalenivorans、Polaromonas種、Crocospaera watsonii、Cyanothecae種、Microcystis aeruginosa、Synechococcus種、Acetohalobium arabaticum、Ammonifex degensii、Caldicelulosiruptor beccii、Candidatus Desulfurovibrio、Clostridium botulinum、Clostridium difficile、Finegoldia magna、Natranaerobius thermophilus、Pelotomaculum thermopropionicum、Acidithiobacillus caldus、Acidithiobacillus ferrooxidans、Allochromatium vinosum、Marinobacter種、Nitrosococcus halophilus、Nitrosococcus watsoni、Pseudalteromonas haloplanktis、Ktedonobacter racemifer、Methanohalobium evestigatum、Anabaena variabilis、Nodularia spumigena、Nostoc種、Arthrospira maxima、Arthrospira plateensis、Arthrospira種、Lyngbya種、Microcoleus chthonoplastes、Oscillatoria種、Petrotoga mobilis、Thermosiphon africanus、またはAcaryochloris marina由来のものが挙げられるが、これらに限定されない。Cas9ファミリーメンバーのさらなる例は、PCT公開WO 2014/131833に記載されている。S. pyogenes由来のCas9（割り当てられたSwissProt登録番号Q99ZW2）は、例示的な酵素である。S. aureus由来のCas9（割り当てられたUniProt登録番号J7RUA5）は、別の例示的な酵素である。

## 【0224】

Casタンパク質の別の例は、Cpf1 (Prevotella及びFrancisella 1由来のCRISPR) タンパク質である。Cpf1は大きなタンパク質（約1300アミノ酸）であり、Cas9の対応するドメインに相同なRuvC様のヌクレアーゼドメインと、Cas9の特徴的なアルギニンリッチクラスターの対応物を有する。しかしながら、Cpf1には、Cas9タンパク質に存在するHNHヌクレアーゼドメインがなく、RuvC様ドメインは、HNHドメインを含む長いインサートを含むCas9とは対照的に、Cpf1配列内で連続している。例示的なCpf1タンパク質として、Francisella tularensis 1、Francisella tularensis亜種novicida、Prevotella albensis、Lachnospiraceae bacterium MC2017 1、Butyryrivibrio proteoclasticus、Peregrinibacteriabacterium GW2011\_GWA2\_33\_10、Parcubacteriabacterium GW2011\_GWC2\_44\_17、Smithella種SCADC、Acidaminococcus種BV3L6、Lachnospiraceae bacterium MA2020、Candidatus Methanoplasma thermatum、Eubacterium eligens、Moraxella bovoculi 237、Leptospira inadai、Lachnospiraceae bacterium ND2006、Porphyromonas creviori canis 3、Prevotella disiens、及びPorphyromonas macacae由来のものが挙げられるが、これらに限定されない。Francisella novicida U112由来のCpf1 (FnCpf1；割り当てられたUniProt登録番号A0Q7Q2) は、例示的な酵素である。

## 【0225】

10

20

30

40

50

Casタンパク質は、野生型タンパク質（すなわち、自然界に存在するもの）、改変型Casタンパク質（すなわち、Casタンパク質バリアント）、または野生型もしくは改変型Casタンパク質の断片であり得る。Casタンパク質はまた、野生型または改変型Casタンパク質の活性型バリアントまたは断片であり得る。活性型バリアントまたは断片は、野生型もしくは改変型Casタンパク質またはその部分と、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、または少なくとも99%、または100%の配列同一性を有し得るが、その場合、活性型バリアントは所望の切断部位で切断する能力を保持しており、したがって、ニック誘導活性または二本鎖切断誘導活性を保持している。ニック誘導活性または二本鎖切断誘導活性のアッセイは公知であり、一般的にそれは、切断部位を含むDNA基質上のCasタンパク質の全体的な活性と特異性を測定する。

#### 【0226】

Casタンパク質は、DNaseドメインなどの少なくとも1つのヌクレアーゼドメインを含み得る。例えば、野生型Cpf1タンパク質は一般的に、おそらくは二量体構成で、標的DNAの両方の鎖を切断するRuvC様ドメインを含む。Casタンパク質は、DNaseドメインなどの少なくとも2つのヌクレアーゼドメインを含み得る。例えば、野生型Cas9タンパク質は一般的に、RuvC様ヌクレアーゼドメインとHNH様ヌクレアーゼドメインを含む。RuvC及びHNHドメインはそれぞれ、二本鎖DNAの異なる鎖を切断して、DNAの二本鎖切断を行うことができる。

#### 【0227】

Casタンパク質（例えば、ヌクレアーゼ活性Casタンパク質またはヌクレアーゼ不活性Casタンパク質）を、融合タンパク質として異種ポリペプチドに作動可能に連結することもできる。例えば、Casタンパク質を、切断ドメイン、エピジェネティック修飾ドメイン、転写活性化ドメイン、または転写抑制因子ドメインに融合させることができる。転写活性化ドメインの例として、単純ヘルペスウイルスVP16活性化ドメイン、VP64（VP16の四量体誘導体）、NFBp65活性化ドメイン、p53活性化ドメイン1及び2、CREB（cAMP応答エレメント結合タンパク質）活性化ドメイン、E2A活性化ドメイン、ならびにNFAT（活性化T細胞の核因子）活性化ドメインが挙げられる。他の例として、Oct1、Oct-2A、SP1、AP-2、CTF1、P300、CBP、PCAF、SRC1、PvALF、ERF-2、OsgAI、HALF-1、C1、AP1、ARF-5、ARF-6、ARF-7、ARF-8、CPRF1、CPRF4、MYC-RP/GP、TRAB1PC4、及びHSF1由来の活性化ドメインが挙げられるが、これらに限定されない。例えば、米国特許出願公開第2016/0237456号、欧州特許第EP3045537号、及びPCT公開WO2011/145121参照。

#### 【0228】

いくつかの実施形態では、MS2-p65-HSF1と対になったdCas9-VP64融合タンパク質を含む転写活性化系を使用することができる。そのような系のガイドRNAは、sgRNAテトラループに付加したアプタマー配列と、二量体化MS2バクテリオファージコートタンパク質を結合するように設計したステムループ2を使用して設計することができる。例えば、Konermann et al., *Nature*, 2015, 517, 583-588参照。転写抑制因子ドメインの例として、誘導性cAMP初期抑制因子（ICER）ドメイン、Kruppel関連ボックスA（Krab-A）抑制因子ドメイン、YY1グリシンリッチ抑制因子ドメイン、Sp1様抑制因子、E(spl)抑制因子、抑制因子、及びMeCP2が挙げられる。他の例として、A/B、KOX、TGF-誘導性初期遺伝子（TIEG）、v-erbA、SID、SID4X、MBD2、MBD3、DNMT1、DNMG3A、DNMT3B、Rb、ROM2由来の転写抑制ドメインが挙げられるが、これらに限定されず、例えば、欧州特許第EP3045537号及びPCT公開WO2011/145121参照。Casタンパク質はまた、安定

性の増加または減少を提供する異種ポリペプチドに融合させることができる。融合ドメインまたは異種ポリペプチドは、Casタンパク質のN末端、C末端、または内部に配置することができる。

【0229】

Cas融合タンパク質の例は、細胞内局在化を提供する異種ポリペプチドに融合したCasタンパク質である。そのような異種ポリペプチドは、例えば、1つ以上の核局在化シグナル(NLS)、例えば、核を標的とするSV40 NLS、ミトコンドリアを標的とするミトコンドリア局在化シグナル、ER保持シグナルなどを含み得る。そのような細胞内局在シグナルは、N末端、C末端、またはCasタンパク質内の任意の場所に配置することができる。NLSには、一連の塩基性アミノ酸を含ませることができ、一本鎖配列または二本鎖配列であり得る。

10

【0230】

Casタンパク質はまた、細胞透過性ドメインに作動可能に連結することができる。例えば、細胞透過性ドメインは、HIV-1 TATタンパク質、ヒトB型肝炎ウイルス由来TLR細胞透過性モチーフ、MPG、Pep-1、VP22、単純ヘルペスウイルス由来細胞透過性ペプチド、またはポリアルギニンペプチド配列由来のものとすることができる。細胞透過性ドメインは、N末端、C末端、またはCasタンパク質内の任意の場所に配置することができる。

【0231】

Casタンパク質はまた、追跡または精製を容易にするために、蛍光タンパク質、精製タグ、またはエピトープタグなどの異種ポリペプチドに作動可能に連結することができる。蛍光タンパク質の例として、緑色蛍光タンパク質(例えば、GFP、GFP-2、tagGFP、turboGFP、eGFP、Emerald、Azami Green、モノマー型Azami Green、CopGFP、AceGFP、ZsGreen1)、黄色蛍光タンパク質(例えば、YFP、eYFP、Citrine、Venus、YPet、PhiYFP、ZsYellow1)、青色蛍光タンパク質(eBFP、eBFP2、Azurite、mKalamal、GFPuv、Sapphire、T-Sapphire)、シアン蛍光タンパク質(例えば、eCFP、Cerulean、CyPet、AmCyan1、Midoriishi-Cyan)、赤色蛍光タンパク質(mKate、mKate2、mPlum、DsRedモノマー、mCherry、mRFP1、DsRed-Express、DsRed2、DsRed-Monomer、HcRed-Tandem、HcRed1、AsRed2、eqFP611、mRaspberry、mStrawberry、Jred)、橙色蛍光タンパク質(mOrange、mKO、Kusabira-Orange、モノマー型Kusabira-Orange、mTangerine、tdTomato)、及び任意の他の適切な蛍光タンパク質が挙げられる。タグの例として、グルタチオン-S-トランスフェラーゼ(GST)、キチン結合タンパク質(CBP)、マルトース結合タンパク質、チオレドキシン(TRX)、ポリ(NANP)、タンデムアフィニティー精製(TAP)タグ、myc、AcV5、AU1、AU5、E、ECS、E2、FLAG、赤血球凝集素(HA)、nus、Softag1、Softag3、Strep、SBP、Glu-Glu、HSV、KT3、S、S1、T7、V5、VSV-G、ヒスチジン(His)、ビオチンカルボキシルキアタンパク質(BCCP)、及びカルモジュリンが挙げられる。

20

タグの例として、グルタチオン-S-トランスフェラーゼ(GST)、キチン結合タンパク質(CBP)、マルトース結合タンパク質、チオレドキシン(TRX)、ポリ(NANP)、タンデムアフィニティー精製(TAP)タグ、myc、AcV5、AU1、AU5、E、ECS、E2、FLAG、赤血球凝集素(HA)、nus、Softag1、Softag3、Strep、SBP、Glu-Glu、HSV、KT3、S、S1、T7、V5、VSV-G、ヒスチジン(His)、ビオチンカルボキシルキアタンパク質(BCCP)、及びカルモジュリンが挙げられる。

30

【0232】

また、Cas9タンパク質を、外来性ドナー配列または標識した核酸に係留することができる。そのような係留(すなわち、物理的連結)は、共有相互作用または非共有相互作用を通じて達成することができ、係留は、直接的であり得るか(例えば、直接的融合によって、あるいはタンパク質のシステインもしくはリジン残基の修飾またはインテイン修飾により達成できる化学的結合によって)、またはストレプトアビシンもしくはアブタマーなどの1つ以上の介在リンカーもしくはアダプター分子を介して達成することができる。タンパク質-核酸コンジュゲートを合成するための非共有結合型の戦略として、ビオチン

40

50

- ストレプトアビシン法及びニッケル - ヒスチジン法が挙げられる。適切に官能化した核酸及びタンパク質を多種多様な化学を使用して結合させることにより、共有結合型のタンパク質 - 核酸コンジュゲートを合成することができる。これらの化学のいくつかでは、タンパク質表面のアミノ酸残基へのオリゴヌクレオチドの直接的な付加が伴う（例えば、リジンアミンまたはシステインチオール）が、一方で、他のより複雑なスキームでは、タンパク質の翻訳後修飾またはタンパク質の触媒ドメインもしくは反応性ドメインの関与が必要である。タンパク質を核酸に共有結合させる方法として、例えば、オリゴヌクレオチドと、タンパク質のリジンまたはシステイン残基との化学的架橋、発現させるタンパク質のライゲーション、化学酵素法、及び光アプタマーの使用が挙げられ得る。外来性ドナー配列または標識した核酸を、C末端、N末端、またはCas9タンパク質内の内部領域に係留することができる。いくつかの実施形態では、外来性ドナー配列または標識した核酸を、Cas9タンパク質のC末端またはN末端に係留する。同様に、Cas9タンパク質を、5'末端、3'末端、または外来性ドナー配列または標識した核酸内の内部領域に係留することができる。いくつかの実施形態では、Cas9タンパク質を、外来性ドナー配列または標識した核酸の5'末端または3'末端に係留する。

#### 【0233】

Casタンパク質は、任意の形態で提供することができる。例えば、Casタンパク質を、タンパク質、例えば、gRNAと複合体化したCasタンパク質の形態で提供することができる。あるいは、Casタンパク質を、Casタンパク質をコードする核酸、例えば、RNA（例えば、メッセンジャーRNA（mRNA））またはDNAの形態で提供することができる。いくつかの実施形態では、Casタンパク質をコードする核酸を、特定の細胞または生物においてタンパク質へ効率的に翻訳させるために、コドン最適化することができる。例えば、Casタンパク質をコードする核酸を、細菌細胞、酵母細胞、ヒト細胞、非ヒト細胞、哺乳類細胞、げっ歯類細胞、マウス細胞、ラット細胞、または任意の他の目的の宿主細胞において天然のポリヌクレオチド配列に比べて高い使用頻度を有する置換コドンに改変することができる。Casタンパク質をコードする核酸を細胞に導入する場合、Casタンパク質は、細胞内で、一時的に、条件的に、または構成的に発現する。

#### 【0234】

Casタンパク質をコードする核酸を、細胞のゲノムに安定的に組み込み、細胞内で活性なプロモーターに作動可能に連結することができる。あるいは、Casタンパク質をコードする核酸を、発現構築物のプロモーターに作動可能に連結することができる。発現構築物には、遺伝子または他の目的の核酸配列（例えば、Cas遺伝子）の発現を誘導することができ、そのような目的の核酸配列を標的細胞に移入することができる任意の核酸構築物が含まれる。例えば、Casタンパク質をコードする核酸を、核酸インサートを含むターゲティングベクター及び/またはgRNAをコードするDNAを含むベクター内に存在させることができる。あるいは、核酸インサートを含むターゲティングベクターとは別個の、及び/またはgRNAをコードするDNAを含むベクターとは別個のベクターまたはプラスミド内に存在させることができる。発現構築物に使用することができるプロモーターとして、例えば、真核細胞、ヒト細胞、非ヒト細胞、哺乳類細胞、非ヒト哺乳類細胞、げっ歯類細胞、マウス細胞、ラット細胞、ハムスター細胞、ウサギ細胞、多能性細胞、胚性幹（ES）細胞、または接合体の1つ以上で活性なプロモーターが挙げられる。そのようなプロモーターは、例えば、条件的プロモーター、誘導性プロモーター、構成的プロモーター、または組織特異的プロモーターであり得る。いくつかの実施形態では、プロモーターは、一方の向きへのCasタンパク質と、他方の向きへのガイドRNAの両方の発現を駆動する双方向性プロモーターであり得る。そのような双方向性プロモーターは：1) 3つの外部制御エレメント：遠位配列エレメント（DSE）、近位配列エレメント（PSE）、及びTATAボックスを含む完全な従来型の单方向性Pol IIIプロモーター、及び2) PSE及びDSEの5'末端に逆方向に融合したTATAボックスを含む第2の基本的なPol IIIプロモーターからなり得る。例えば、H1プロモーターでは、DSEはPSE及びTATAボックスに隣接しており、U6プロモーター由来のPSE及

10

20

30

40

50

び T A T A ボックスを追加することによって逆方向の転写を制御するハイブリッドプロモーターを作成することにより、プロモーターを双方向性にすることができる。双方向性プロモーターの使用により、Cas タンパク質及びガイド RNA をコードする遺伝子を同時に発現させ、コンパクトな発現カセットを生成して送達を促進することができる。

#### 【0235】

本開示はまた、Cas タンパク質（例えば、Cas9 タンパク質）に結合し、Cas タンパク質を標的 DNA（例えば、B4GALT1 遺伝子）内の特定の位置に標的化するガイド RNA（gRNA）を提供する。いくつかの実施形態では、ガイド RNA は、Cas 酵素を誘導して内在性 B4GALT1 遺伝子に結合または切断するのに有効であり、ガイド RNA は、例えば、配列番号 1 の 53575 ~ 53577 位を含むかまたはそれに近接する内在性 B4GALT1 遺伝子内のガイド RNA 認識配列にハイブリダイズする DNA ターゲティングセグメントを含む。例えば、ガイド RNA 認識配列は、配列番号 1 の 53575 ~ 53577 位の、約 5 以内、約 10 以内、約 15 以内、約 20 以内、約 25 以内、約 30 以内、約 35 以内、約 40 以内、約 45 以内、約 50 以内、約 100 以内、約 200 以内、約 300 以内、約 400 以内、約 500 以内、または約 1,000 ヌクレオチド以内とすることができる。他の例示的なガイド RNA は、配列番号 1 のエクソン 5 に対応する領域内にある内在性 B4GALT1 遺伝子内のガイド RNA 認識配列にハイブリダイズする DNA ターゲティングセグメントを含む。他の例示的なガイド RNA は、内在性 B4GALT1 遺伝子の開始コドンを含むかもしくはそれに近接するか、または内在性 B4GALT1 遺伝子の終止コドンを含むかもしくはそれに近接する内在性 B4GALT1 遺伝子内のガイド RNA 認識配列にハイブリダイズする DNA ターゲティングセグメントを含む。例えば、ガイド RNA 認識配列は、開始コドンの、約 5 以内、約 10 以内、約 15 以内、約 20 以内、約 25 以内、約 30 以内、約 35 以内、約 40 以内、約 45 以内、約 50 以内、約 100 以内、約 200 以内、約 300 以内、約 400 以内、約 500 以内、もしくは約 1,000 ヌクレオチド以内、または終止コドンの、約 5 以内、約 10 以内、約 15 以内、約 20 以内、約 25 以内、約 30 以内、約 35 以内、約 40 以内、約 45 以内、約 50 以内、もしくは約 1,000 ヌクレオチド以内とすることができる。内在性 B4GALT1 遺伝子は、任意の生物に由来する B4GALT1 遺伝子とすることができます。例えば、B4GALT1 遺伝子は、ヒト B4GALT1 遺伝子または非ヒト哺乳類、げっ歯類、マウス、もしくはラットなどの別の生物由来のオルソログとすることができます。

#### 【0236】

いくつかの実施形態では、ガイド RNA 認識配列は、ヒト B4GALT1 遺伝子の 5' 末端に存在する。いくつかの実施形態では、ガイド RNA 認識配列は、ヒト B4GALT1 遺伝子の転写開始部位（TSS）に隣接している。いくつかの実施形態では、ガイド RNA 認識配列は、ヒト B4GALT1 遺伝子の 3' 末端に存在する。いくつかの実施形態では、ガイド RNA 認識配列は、配列番号 1 の 53575 ~ 53577 位に近接している。配列番号 1 の 53575 ~ 53577 位に近接する例示的なガイド RNA 認識配列として、ATTAGTTTTAGAGGCCATGT（配列番号 9）及び GGCTCTCAGGCCAAAGTGTAT（配列番号 10）（いずれも配列番号 1 の 53575 ~ 53577 位の 5' 側）、ならびに TACTCCCTTCCCCCTTTAGGA（配列番号 11）及び GTCCGAGGCTCTGGGCCCTAG（配列番号 12）（いずれも配列番号 1 の 53575 ~ 53577 位の 3' 側）が挙げられるが、これらに限定されない。

#### 【0237】

ガイド RNA は、2 つのセグメント：DNA ターゲティングセグメント及びタンパク質結合セグメントを含み得る。いくつかの gRNA は、2 つの別々の RNA 分子を含む：アクチベーター RNA（例えば、tracrRNA）及びターゲター RNA（例えば、CRIISPR RNA すなわち crRNA）。他の gRNA は単一の RNA 分子（単一の RNA ポリヌクレオチド；単一分子の gRNA、シングルガイド RNA、すなわち sgRNA）である。例えば、Cas9 の場合、シングルガイド RNA は、tracrRNA に（例

10

20

30

40

50

えばリンカーを介して)融合させたc r RNAを含み得る。例えば、C p f 1の場合、切断を達成するためにc r RNAのみが必要である。g RNAには、2分子(すなわち、モジュラー型)g RNAと単一分子のg RNAの両方が含まれる。

#### 【0238】

所定のg RNAのDNAターゲティングセグメント(c r RNA)は、標的DNAの配列(すなわち、ガイドRNA認識配列)に相補的なヌクレオチド配列を含む。g RNAのDNAターゲティングセグメントは、ハイブリダイゼーション(すなわち、塩基の対合)を介して、配列特異的な様式で標的DNA(例えば、B 4 G A L T 1遺伝子)と相互作用する。そのため、DNAターゲティングセグメントのヌクレオチド配列は、様々に異なり得るとともに、g RNAと標的DNAが相互作用する標的DNA内の位置を決定付ける。対象g RNAのDNAターゲティングセグメントを、標的DNA内の任意の所望の配列にハイブリダイズするように改変することができる。天然のc r RNAは、C R I S P R - C a s 系及び生物によって異なるが、多くの場合、約21～約46ヌクレオチドの長さの2つの直列反復配列(D R)が隣接する約21～約72ヌクレオチド長のターゲティングセグメントを含む。S. pyogenesの場合、D Rは36ヌクレオチド長であり、ターゲティングセグメントは30ヌクレオチド長である。3'に位置するD Rは、対応するt r a c r RNAに相補的であり、対応するt r a c r RNAとハイブリダイズし、次いでC a s タンパク質に結合する。

#### 【0239】

DNAターゲティングセグメントは、少なくとも約12ヌクレオチド、少なくとも約15ヌクレオチド、少なくとも約17ヌクレオチド、少なくとも約18ヌクレオチド、少なくとも約19ヌクレオチド、少なくとも約20ヌクレオチド、少なくとも約25ヌクレオチド、少なくとも約30ヌクレオチド、少なくとも約35ヌクレオチド、または少なくとも約40ヌクレオチドの長さを有し得る。そのようなDNAターゲティングセグメントは、約12ヌクレオチド～約100ヌクレオチド、約12ヌクレオチド～約80ヌクレオチド、約12ヌクレオチド～約50ヌクレオチド、約12ヌクレオチド～約40ヌクレオチド、約12ヌクレオチド～約30ヌクレオチド、約12ヌクレオチド～約25ヌクレオチド、または約12ヌクレオチド～約20ヌクレオチドの長さを有し得る。例えば、DNAターゲティングセグメントは、約15ヌクレオチド～約25ヌクレオチド(例えば、約17ヌクレオチド～約20ヌクレオチド、または約17ヌクレオチド、約18ヌクレオチド、約19ヌクレオチド、もしくは約20ヌクレオチド)であり得る。例えば、米国出願公開第2016/0024523号参照。S. pyogenes由来のC a s 9の場合、一般的なDNAターゲティングセグメントは、約16～約20ヌクレオチド長または約17～約20ヌクレオチド長である。S. aureus由来のC a s 9の場合、一般的なDNAターゲティングセグメントは、約21～約23ヌクレオチド長である。C p f 1の場合、一般的なDNAターゲティングセグメントは、少なくとも約16ヌクレオチド長または少なくとも約18ヌクレオチド長である。

#### 【0240】

標的DNA内のDNAターゲティング配列とガイドRNA認識配列との間の相補性の割合は、少なくとも約60%、少なくとも約65%、少なくとも約70%、少なくとも約75%、少なくとも約80%、少なくとも約85%、少なくとも約90%、少なくとも約95%、少なくとも約97%、少なくとも約98%、少なくとも約99%、または100%であり得る。標的DNA内のDNAターゲティング配列とガイドRNA認識配列との間の相補性の割合は、約20連続ヌクレオチドにわたって少なくとも約60%であり得る。一例として、標的DNA内のDNAターゲティング配列とガイドRNA認識配列との間の相補性の割合は、標的DNAの相補鎖内のガイドRNA認識配列の5'末端の約14連続ヌクレオチドにわたって約100%であり、また、残部にわたっては約0%と低い。そのような場合、DNAターゲティング配列は約14ヌクレオチド長とみなすことができる。別の例として、標的DNA内のDNAターゲティング配列とガイドRNA認識配列との間の相補性の割合は、標的DNAの相補鎖内のガイドRNA認識配列の5'末端の7連続ヌクレオ

10

20

30

40

50

チドにわたって約 100 % であり、また、残部にわたっては約 0 % と低い。そのような場合、DNA ターゲティング配列は約 7 ヌクレオチド長とみなすことができる。いくつかのガイド RNA では、DNA - ターゲティング配列内の少なくとも約 17 ヌクレオチドが標的 DNA に相補的である。例えば、DNA ターゲティング配列を、約 20 ヌクレオチド長とすることができる、標的 DNA (ガイド RNA 認識配列) に対して 1 つ、2 つ、または 3 つのミスマッチを含み得る。いくつかの実施形態では、ミスマッチはプロトスペーサー隣接モチーフ (PAM) 配列に隣接していない (例えば、ミスマッチが DNA ターゲティング配列の 5' 末端にあるか、またはミスマッチが PAM 配列から、少なくとも 2、少なくとも 3、少なくとも 4、少なくとも 5、少なくとも 6、少なくとも 7、少なくとも 8、少なくとも 9、少なくとも 10、少なくとも 11、少なくとも 12、少なくとも 13、少なくとも 14、少なくとも 15、少なくとも 16、少なくとも 17、少なくとも 18、もしくは少なくとも 19 塩基対離れている)。

#### 【0241】

ガイド RNA には、追加の望ましい特徴 (例えば、変更または調節された安定性；細胞内ターゲティング；蛍光標識による追跡；タンパク質またはタンパク質複合体に対する結合部位など) を提供する修飾または配列を含ませることができる。そのような修飾の例として、例えば、5' キャップ (例えば、7 - メチルグアニル酸キャップ (m7G))；3' ポリアデニル化テール (すなわち、3' ポリ (A) テール)；リボスイッチ配列 (例えば、タンパク質及び / またはタンパク質複合体による調節された安定性及び / または調節されたアクセシビリティを可能にするための)；安定性制御配列；dsRNA 二重鎖を形成する配列 (すなわち、ヘアピン)；RNA を細胞内の位置 (例えば、核、ミトコンドリア、葉緑体など) に標的化する修飾または配列；追跡を提供する修飾または配列 (例えば、蛍光分子への直接結合、蛍光検出を促進する部分への結合、蛍光検出を可能にする配列など)；タンパク質に対する結合部位を提供する修飾または配列 (例えば、転写活性化因子、転写抑制因子、DNA メチルトランスフェラーゼ、DNA デメチラーゼ、ヒストンアセチルトランスフェラーゼ、ヒストンデアセチラーゼなどを含む、DNA に作用するタンパク質)；及びその組み合わせが挙げられる。

#### 【0242】

ガイド RNA は、任意の形態で提供することができる。例えば、gRNA は、2 つの分子 (別個の crRNA 及び tracrRNA) または 1 つの分子 (sgRNA) のいずれかの RNA の形態で、及び場合により Cas タンパク質との複合体の形態で提供することができる。例えば、gRNA は、例えば T7 RNA ポリメラーゼを使用した *in vitro* 転写によって調製することができる。ガイド RNA は、化学合成によって調製することもできる。

#### 【0243】

gRNA はまた、gRNA をコードする DNA の形態で提供することができる。gRNA をコードする DNA は、単一の RNA 分子 (sgRNA) または別個の RNA 分子 (例えば、別個の crRNA 及び tracrRNA) をコードし得る。後者の場合、gRNA をコードする DNA は、1 つの DNA 分子として、または crRNA と tracrRNA をそれぞれコードする別々の DNA 分子として提供することができる。gRNA を DNA の形で提供する場合、gRNA を細胞内で、一時的に、条件的に、または構成的に発現させることができる。gRNA をコードする DNA を、細胞のゲノムに安定的に組み込み、細胞内で活性なプロモーターに作動可能に連結することができる。あるいは、gRNA をコードする DNA を、発現構築物のプロモーターに作動可能に連結することができる。例えば、gRNA をコードする DNA を、異種核酸を含むベクター内に存在させることができる。ベクターには外来性ドナー配列をさらに含ませることができ、及び / またはベクターには Cas タンパク質をコードする核酸をさらに含ませることができる。あるいは、gRNA をコードする DNA を、外来性ドナー配列を含むベクター及び / または Cas タンパク質をコードする核酸を含むベクターとは別個のベクターまたはプラスミドに存在させることができる。そのような発現構築物において使用することができるプロモーターとし

て、例えば、真核細胞、ヒト細胞、非ヒト細胞、哺乳類細胞、非ヒト哺乳類細胞、げっ歯類細胞、マウス細胞、ラット細胞、ハムスター細胞、ウサギ細胞、多能性細胞、胚性幹細胞、または接合体の1つ以上において活性なプロモーターが挙げられる。そのようなプロモーターは、例えば、条件的プロモーター、誘導性プロモーター、構成的プロモーター、または組織特異的プロモーターであり得る。そのようなプロモーターはまた、例えば、双方向性プロモーターであり得る。適切なプロモーターの具体例として、RNAポリメラーゼIIプロモーター、例えば、ヒトU6プロモーター、ラットU6ポリメラーゼIIプロモーター、またはマウスU6ポリメラーゼIIプロモーターが挙げられる。

#### 【0244】

本開示はまた、本明細書に開示する1つ以上のガイドRNA（例えば、1、2、3、4、またはそれ以上のガイドRNA）及び単離された核酸またはタンパク質の安定性を高める（例えば、所定の保管条件（例えば、-、20、4、または周囲温度）で、分解産物が閾値未満（出発核酸またはタンパク質の0.5重量%未満など）に留まる期間を延長するか；またはin vivoでの安定性を増加させる）担体を含む組成物も提供する。そのような担体の例として、ポリ（乳酸）（PLA）ミクロスフェア、ポリ（D,L-乳酸-コグリコール酸）（PLGA）ミクロスフェア、リポソーム、ミセル、逆ミセル、脂質コクリエート、及び脂質微小管が挙げられるが、これらに限定されない。そのような組成物は、Cas9タンパク質などのCasタンパク質、またはCasタンパク質をコードする核酸をさらに含み得る。そのような組成物は、本明細書の別の場所に開示する、1つ以上（例えば、1、2、3、4、またはそれ以上）の外来性ドナー配列及び/または1つ以上（例えば、1、2、3、4、またはそれ以上）のターゲティングベクター及び/または1つ以上（例えば、1、2、3、4、またはそれ以上）の発現ベクターをさらに含み得る。

#### 【0245】

ガイドRNA認識配列には、結合に十分な条件が存在する場合、gRNAのDNAターゲティングセグメントが結合する標的DNA（例えば、B4GALT1遺伝子）に存在する核酸配列が含まれる。例えば、ガイドRNA認識配列には、ガイドRNAが相補性を有するように設計した配列が含まれ、その場合、ガイドRNA認識配列とDNAターゲティング配列との間のハイブリダイゼーションが、CRISPR複合体の形成を促進する。ハイブリダイゼーションを引き起こし、CRISPR複合体の形成を促進するのに十分な相補性があれば、必ずしも完全な相補性は必要ない。ガイドRNA認識配列には、Casタンパク質の切断部位も含まれ、これについては以下で詳しく説明する。ガイドRNA認識配列は、例えば、細胞の核もしくは細胞質内、またはミトコンドリアもしくは葉緑体などの細胞のオルガネラ内に配置することができる任意のポリヌクレオチドを含み得る。

#### 【0246】

標的DNA内のガイドRNA認識配列は、Casタンパク質またはgRNAにより標的化する（すなわち、結合されるか、またはハイブリダイズされるか、または相補的である）ことができる。適切なDNA/RNA結合条件には、細胞に通常存在する生理学的条件が含まれる。他の適切なDNA/RNA結合条件は公知である。

#### 【0247】

Casタンパク質は、gRNAのDNAターゲティングセグメントが結合する標的DNAに存在する核酸配列の内部または外部の部位で核酸を切断することができる。「切断部位」には、Casタンパク質が一本鎖切断または二本鎖切断を生じる核酸の位置が含まれる。例えば、CRISPR複合体（ガイドRNA認識配列にハイブリダイズし、Casタンパク質と複合体を形成するgRNAを含む）の形成により、gRNAのDNAターゲティングセグメントが結合する標的DNAに存在する核酸配列内もしくはその近傍（例えば、1以内、2以内、3以内、4以内、5以内、6以内、7以内、8以内、9以内、10以内、20以内、もしくは50以内、またはそれ以上の塩基対）において、一方または両方の鎖が切断される。切断部位は、核酸の片方の鎖のみ、または両方の鎖上にあり得る。切断部位は、核酸の両方の鎖の同じ位置とできるか（平滑末端を生成）、または

10

20

30

40

50

各鎖の異なる部位に存在させることができる（付着末端（すなわち、オーバーハング）を生成）。いくつかの実施形態では、第1鎖上のニッカーゼのガイドRNA認識配列は、第2鎖上のニッカーゼのガイドRNA認識配列から、少なくとも2、少なくとも3、少なくとも4、少なくとも5、少なくとも6、少なくとも7、少なくとも8、少なくとも9、少なくとも10、少なくとも15、少なくとも20、少なくとも25、少なくとも30、少なくとも40、少なくとも50、少なくとも75、少なくとも100、少なくとも250、少なくとも500、または少なくとも1,000塩基対、離れている。

#### 【0248】

Casタンパク質による標的DNAの部位特異的切断は、標的DNA内の、i) gRNAと標的DNAとの間の塩基対相補性及びii) プロトスペーサー隣接モチーフ (PAM) と呼ばれる短いモチーフの両方によって決定される位置で起こり得る。PAMはガイドRNA認識配列に隣接させることができる。いくつかの実施形態では、ガイドRNA認識配列に、PAMを3'末端で隣接させることができる。あるいは、ガイドRNA認識配列に、PAMを5'末端で隣接させることができる。例えば、Casタンパク質の切断部位を、PAM配列の上流または下流における約1～約10、または約2～約5塩基対（例えば3塩基対）とすることができる。いくつかの場合（例えば、*S. pyogenes*由来のCas9または密接に関連するCas9を使用する場合）では、非相補鎖のPAM配列は5' - N<sub>1</sub>GG - 3' とすることができる、式中、Nは任意のDNAヌクレオチドであり、標的DNAの非相補鎖のガイドRNA認識配列の直近の3'側である。したがって、相補鎖のPAM配列は5' - C C N<sub>2</sub> - 3' であるだろうが、式中、N<sub>2</sub>は任意のDNAヌクレオチドであり、標的DNAの相補鎖のガイドRNA認識配列の直近の5'側である。いくつかのそのような場合では、N<sub>1</sub>とN<sub>2</sub>を相補的とすることができる、N<sub>1</sub> - N<sub>2</sub>塩基対を任意の塩基対とすることができる（例えば、N<sub>1</sub> = C及びN<sub>2</sub> = G；N<sub>1</sub> = G及びN<sub>2</sub> = C；N<sub>1</sub> = A及びN<sub>2</sub> = T；またはN<sub>1</sub> = T、及びN<sub>2</sub> = A）。*S. aureus*由来のCas9の場合、PAMはNNGRRT（配列番号13）またはNNGRR（配列番号14）とすることができる、式中、Nは、A、G、C、またはTとすることができる、Rは、GまたはAとすることができる。いくつかの場合（例えば、FnCpf1の場合）では、PAM配列は、5'末端の上流に存在し、配列5' - TTN - 3'を有し得る。

#### 【0249】

ガイドRNA認識配列の例として、gRNAのDNAターゲティングセグメントに相補的なDNA配列、またはPAM配列に加えてそのようなDNA配列が挙げられる。例えば、標的モチーフは、GN<sub>19</sub>NGG（配列番号15）またはN<sub>20</sub>NGG（配列番号16）などのCas9タンパク質によって認識されるNGGモチーフの直前の20ヌクレオチドのDNA配列であり得る（例えばPCT公開WO 2014/165825参照）。5'末端のグアニンは、細胞内のRNAポリメラーゼによる転写を促進することができる。ガイドRNA認識配列の他の例は、in vitroでT7ポリメラーゼによる効率的な転写を促進するために、5'末端に2つのグアニンヌクレオチド（例えば、GGN<sub>0</sub>NGG；配列番号17）を含み得る。例えば、PCT公開WO 2014/065596参照。他のガイドRNA認識配列は、5' GまたはGG及び3' GGまたはNGGを含む、約4～約22ヌクレオチド長を有し得る。いくつかの実施形態では、ガイドRNA認識配列は、約14～約20ヌクレオチド長を有し得る。

#### 【0250】

ガイドRNA認識配列は、細胞に対して内在性または外来性の任意の核酸配列であり得る。ガイドRNA認識配列は、遺伝子産物（例えば、タンパク質）をコードする配列もしくは非コーディング配列（例えば、調節配列）とすることができます、またはその両方を含ませることができる。

#### 【0251】

いくつかの実施形態では、ガイドRNA認識配列を、配列番号1のエクソン5に対応する領域内に存在させることができる。いくつかの実施形態では、ガイドRNA認識配列は、配列番号1の53575～53577位を含むことができるか、またはそれに近接する

10

20

30

40

50

。例えば、ガイドRNA認識配列は、配列番号1の53575～53577位に対応する位置の、約1000以内、約500以内、約400以内、約300以内、約200以内、約100以内、約50以内、約45以内、約40以内、約35以内、約30以内、約25以内、約20以内、約15以内、約10以内、または約5ヌクレオチド以内とすることができる。いくつかの実施形態では、ガイドRNA認識配列は、内在性B4GALT1遺伝子の開始コドンまたは内在性B4GALT1遺伝子の終止コドンを含むか、またはそれに近接させることができる。例えば、ガイドRNA認識配列は、開始コドンまたは終止コドンの、約10以内、約20以内、約30以内、約40以内、約50以内、約100以内、約200以内、約300以内、約400以内、約500以内、または約1,000ヌクレオチド以内とすることができる。

10

#### 【0252】

本明細書に開示する方法及び組成物は、内在性B4GALT1遺伝子を切断することなく、またはヌクレアーゼ剤による内在性B4GALT1遺伝子の切断後に、外来性ドナー配列（例えば、ターゲティングベクターまたは修復テンプレート）を利用して内在性B4GALT1遺伝子を改変することができる。外来性ドナー配列とは、標的配列との部位特異的組換えを可能にするために必要なエレメントを含む任意の核酸またはベクターを指す。ヌクレアーゼ剤と組み合わせて外来性ドナー配列を使用することにより、相同組換え修復の促進による内在性B4GALT1遺伝子内により正確な改変がもたらされ得る。

#### 【0253】

そのような方法においては、ヌクレアーゼ剤が内在性B4GALT1遺伝子を切断して一本鎖切断（ニック）または二本鎖切断を生じさせ、外来性ドナー配列が非相同末端結合（NHEJ）を介したライゲーション、または相同組換え修復事象によって内在性B4GALT1遺伝子と組み換わる。外来性ドナー配列による修復は、ヌクレアーゼ切断部位を除去または破壊する場合があり、それにより、標的化された対立遺伝子はヌクレアーゼ剤によって再標的化され得ない。

20

#### 【0254】

外来性ドナー配列には、デオキシリボ核酸（DNA）またはリボ核酸（RNA）を含ませることができ、それらは一本鎖または二本鎖とすることでき、それらは直鎖または環状形態とすることができます。例えば、外来性ドナー配列を、一本鎖オリゴデオキシヌクレオチド（ssODN）とすることができる。例示的な外来性ドナー配列は、長さが約50ヌクレオチド～約5kb、長さが約50ヌクレオチド～約3kb、または長さが約50～約1,000ヌクレオチドである。他の例示的な外来性ドナー配列は、長さが約40～約200ヌクレオチドである。例えば、外来性ドナー配列を、約50～約60、約60～約70、約70～約80、約80～約90、約90～約100、約100～約110、約110～約120、約120～約130、約130～約140、約140～約150、約150～約160、約160～約170、約170～約180、約180～約190、または約190～約200ヌクレオチド長とすることができます。あるいは、外来性ドナー配列を、約50～約100、約100～約200、約200～約300、約300～約400、約400～約500、約500～約600、約600～約700、約700～約800、約800～約900、または約900～約1,000ヌクレオチド長とすることができます。あるいは、外来性ドナー配列を、約1kb～約1.5kb、約1.5kb～約2kb、約2kb～約2.5kb、約2.5kb～約3kb、約3kb～約3.5kb、約3.5kb～約4kb、約4kb～約4.5kb、または約4.5kb～約5kb長とすることができます。あるいは、外来性ドナー配列を、例えば、約5kb以下、約4.5kb以下、約4kb以下、約3.5kb以下、約3kb以下、約2.5kb以下、約2kb以下、約1.5kb以下、約1kb以下、約900ヌクレオチド以下、約800ヌクレオチド以下、約700ヌクレオチド以下、約600ヌクレオチド以下、約500ヌクレオチド以下、約400ヌクレオチド以下、約300ヌクレオチド以下、約200ヌクレオチド以下、約100ヌクレオチド以下、または約50ヌクレオチド以下の長さとすることができます。

30

#### 【0255】

40

50

いくつかの実施形態では、外来性ドナー配列は、長さが約 80 ヌクレオチド～約 200 ヌクレオチド（例えば、長さが約 120 ヌクレオチド）の ssODN である。別の例では、外来性ドナー配列は、長さが約 80 ヌクレオチド～約 3kb の ssODN である。そのような ssODN には、例えば、それぞれ約 40 ヌクレオチド～約 60 ヌクレオチドの長さの相同性アームを含ませることができる。そのような ssODN にはまた、例えば、それぞれ約 30 ヌクレオチド～100 ヌクレオチドの長さの相同性アームを含ませることもできる。相同性アームは、対称的（例えば、それぞれ約 40 ヌクレオチド長またはそれぞれ約 60 ヌクレオチド長）または非対称的（例えば、長さが約 36 ヌクレオチドの 1 つの相同性アーム、及び長さが約 91 ヌクレオチドの 1 つの相同性アーム）とすることができる。

10

#### 【0256】

外来性ドナー配列には、追加の望ましい特徴（例えば、改変または調節された安定性；蛍光標識による追跡または検出；タンパク質またはタンパク質複合体に対する結合部位など）を提供する修飾または配列を含ませることができる。外来性ドナー配列は、1つ以上の蛍光標識、精製タグ、エピトープタグ、またはそれらの組み合わせを含ませることができる。例えば、外来性ドナー配列には、1つ以上の蛍光標識（例えば、蛍光タンパク質もしくは他のフルオロフォアまたは色素）、例えば、少なくとも 1 つ、少なくとも 2 つ、少なくとも 3 つ、少なくとも 4 つ、または少なくとも 5 つの蛍光標識を含ませることができる。例示的な蛍光標識として、フルオレセイン（例えば、6-カルボキシフルオレセイン（6-FAM）、Texas Red、HEX、Cy3、Cy5、Cy5.5、Pacific Blue、5-（及び-6）-カルボキシテトラメチルローダミン（TAMRA）及び Cy7 などのフルオロフォアが挙げられる。オリゴヌクレオチドを標識するために、広範囲の蛍光色素が市販されている（例えば、Integrated DNA Technologies 製）。そのような蛍光標識（例えば、内部蛍光標識）を使用して、例えば、外来性ドナー配列の末端と適合する突出末端を有する切断された内在性 B4GALT1 遺伝子に直接組み込まれた外来性ドナー配列を検出することができる。標識またはタグは、5' 末端、3' 末端、または外来性ドナー配列の内部に配置することができる。例えば、外来性ドナー配列は、5' 末端で Integrated DNA Technologies の IR700 フルオロフォア（5' IRDYE（登録商標）700）と結合することができる。

20

#### 【0257】

外来性ドナー配列にはまた、内在性 B4GALT1 遺伝子に組み込ませる DNA のセグメントを含む核酸インサートを含ませることができる。内在性 B4GALT1 遺伝子に核酸インサートを組み込むことにより、内在性 B4GALT1 遺伝子における目的の核酸配列の追加、内在性 B4GALT1 遺伝子における目的の核酸配列の欠失、または内在性 B4GALT1 遺伝子における目的の核酸配列の置換（すなわち、欠失及び挿入）をもたらすことができる。いくつかの外来性ドナー配列は、内在性 B4GALT1 遺伝子の対応する欠失を伴わずに内在性 B4GALT1 遺伝子に核酸インサートを挿入するように設計される。他の外来性ドナー配列は、対応する核酸インサートの挿入を伴わずに、内在性 B4GALT1 遺伝子内の目的の核酸配列を欠失させるように設計される。他の外来性ドナー配列は、内在性 B4GALT1 遺伝子内の目的の核酸配列を欠失させ、それを核酸インサートで置換するように設計される。

30

#### 【0258】

核酸インサート、及び内在性 B4GALT1 遺伝子内の欠失及び／または置換させる対応する核酸は、様々な長さであり得る。例示的な核酸インサート、または欠失及び／または置換させる内在性 B4GALT1 遺伝子内の対応する核酸は、約 1 ヌクレオチド～約 5 kb の長さまたは約 1 ヌクレオチド～約 1,000 ヌクレオチドの長さである。例えば、核酸インサート、または内在性 B4GALT1 遺伝子内の欠失及び／または置換させる対応する核酸は、約 1～約 10、約 10～約 20、約 20～約 30、約 30～約 40、約 40～約 50、約 50～約 60、約 60～約 70、約 70～約 80、約 80～約 90、約 9

40

50

0～約100、約100～約110、約110～約120、約120～約130、約130～約140、約140～約150、約150～約160、約160～約170、約170～約180、約180～約190、または約190～約200ヌクレオチド長とすることができる。同様に、核酸インサート、または内在性B4GALT1遺伝子中の欠失及び／または置換させる対応する核酸は、約1～約100、約100～約200、約200～約300、約300～約400、約400～約500、約500～約600、約600～約700、約700～約800、約800～約900、または約900～約1,000ヌクレオチド長とすることができる。同様に、核酸インサート、または内在性B4GALT1遺伝子内の欠失及び／または置換させる対応する核酸は、約1kb～約1.5kb、約1.5kb～約2kb、約2kb～約2.5kb、約2.5kb～約3kb、約3kb～約3.5kb、約3.5kb～約4kb、約4kb～約4.5kb、または約4.5kb～約5kb長とすることができる。

#### 【0259】

核酸インサートには、ゲノムDNAまたは任意の他のタイプのDNAを含ませることができる。例えば、核酸インサートにcDNAを含ませることができる。

核酸インサートには、内在性B4GALT1遺伝子の全部または一部（例えば、B4GALT1ポリペプチドの特定のモチーフまたは領域をコードする遺伝子の部分）に相同な配列を含ませることができる。例えば、核酸インサートに、内在性B4GALT1遺伝子における置換のために標的化された配列と比較して、1つ以上（例えば、1、2、3、4、5、またはそれ以上）の点変異または1つ以上のヌクレオチド挿入もしくは欠失を含む配列を含ませることができる。

#### 【0260】

核酸インサート、または内在性B4GALT1遺伝子内の欠失及び／または置換させる対応する核酸は、エクソンなどのコード領域、イントロン、非翻訳領域、もしくは調節領域などの非コード領域（例えば、プロモーター、エンハンサー、もしくは転写抑制因子結合エレメント）、またはその任意の組み合わせとことができる。

#### 【0261】

核酸インサートにはまた、選択マーカーをコードするポリヌクレオチドを含ませることもできる。あるいは、核酸インサートには、選択マーカーをコードするポリヌクレオチドを含ませないこともできる。選択マーカーには、選択カセットを含ませることができる。いくつかの実施形態では、選択カセットを自己欠失性カセットとすることができます。一例として、自己欠失性カセットには、マウスP<sub>r</sub>m1プロモーターに作動可能に連結したCre遺伝子（イントロンにより分断された、Creリコンビナーゼをコードする2つのエクソンを含む）、及びヒトユビキチンプロモーターに作動可能に連結したネオマイシン耐性遺伝子を含ませることができる。例示的な選択マーカーとして、ネオマイシンホスホトランスクエラーゼ(neo<sup>r</sup>)、ハイグロマイシンBホスホトランスクエラーゼ(hyg<sup>r</sup>)、ピューロマイシン-N-アセチルトランスクエラーゼ(puro<sup>r</sup>)、プラストサイジンSデアミナーゼ(bsr<sup>r</sup>)、キサンチン/グアニンホスホリボシルトランスクエラーゼ(gpt)、または単純ヘルペスウイルスチミジンキナーゼ(HSV-k)、またはその組み合わせが挙げられる。選択マーカーをコードするポリヌクレオチドを、標的化されている細胞内で活性なプロモーターに作動可能に連結することができる。プロモーターの例を、本明細書の他の場所に記載する。

#### 【0262】

核酸インサートにはまた、レポーター遺伝子を含ませることもできる。例示的なレポーター遺伝子として、ルシフェラーゼ、-ガラクトシダーゼ、緑色蛍光タンパク質(GFP)、高感度緑色蛍光タンパク質(eGFP)、シアノ蛍光タンパク質(CFP)、黄色蛍光タンパク質(YFP)、高感度黄色蛍光タンパク質(eYFP)、青色蛍光タンパク質(BFP)、高感度青色蛍光タンパク質(eBFP)、DsRed、ZsGreen、MmGFP、mPlum、mCherry、tdTomato、mStrawberry、J-Red、mOrange、mKO、mCitrine、Venus、YPet、E

10

20

30

40

50

merald、CyPet、Cerulean、T-Sapphire、及びアルカリホスファターゼをコードするものが挙げられる。そのようなレポーター遺伝子を、標的化されている細胞内で活性なプロモーターに作動可能に連結することができる。プロモーターの例を、本明細書の他の場所に記載する。

#### 【0263】

核酸インサートにはまた、1つ以上の発現カセットまたは欠失カセットを含ませることができる。特定のカセットには、発現に影響を及ぼす様々な調節成分とともに、目的のヌクレオチド配列、選択マーカーをコードするポリヌクレオチド、及びレポーター遺伝子の1つ以上を含ませることができる。含ませることができる選択可能なマーカー及びレポーター遺伝子の例を、本明細書の他の場所で詳細に論じる。

10

#### 【0264】

核酸インサートには、部位特異的組換え標的配列に隣接する核酸を含ませることができる。あるいは、核酸インサートには、1つ以上の部位特異的組換え標的配列を含ませることができる。そのような部位特異的組換え標的配列は核酸インサート全体に隣接させることができるが、核酸インサート内の目的の任意の領域または個々のポリヌクレオチドも、そのような部位に隣接させることができる。核酸インサートまたは核酸インサート内の目的のポリヌクレオチドに隣接させることができる部位特異的組換え標的配列として、例えば、 $1oxP$ 、 $1ox511$ 、 $1ox2272$ 、 $1ox66$ 、 $1ox71$ 、 $1oxM2$ 、 $1ox5171$ 、 $FRT$ 、 $FRT11$ 、 $FRT71$ 、 $attP$ 、 $att$ 、 $FRT$ 、 $rox$ 、またはその組み合わせが挙げられる。いくつかの実施形態では、部位特異的組換え部位を、核酸インサート内に含まれる選択マーカー及び/またはレポーター遺伝子をコードするポリヌクレオチドに隣接させる。内在性 B4GALT1 遺伝子への核酸インサートの組み込みに続いて、部位特異的組換え部位間の配列を除去することができる。いくつかの実施形態では、それぞれが部位特異的組換え部位を含む核酸インサートを含む2つの外来性ドナー配列を使用することができる。外来性ドナー配列は、目的の核酸に隣接する5'及び3'領域を標的とすることができる。標的ゲノム遺伝子座への2つの核酸インサートの組み込みに続いて、挿入された2つの部位特異的組換え部位の間の目的の核酸を除去することができる。

20

#### 【0265】

核酸インサートにはまた、I型、II型、III型、及びIV型エンドヌクレアーゼを含む制限エンドヌクレアーゼ（すなわち、制限酵素）の1つ以上の制限酵素部位を含ませることができる。I型及びII型の制限エンドヌクレアーゼは、特定の認識配列を認識するが、通常、ヌクレアーゼ結合部位からの様々な位置で切断し、これは、切断部位（認識配列）から数百塩基対離れている場合がある。II型系では、制限酵素活性はメチラーゼ活性とは無関係であり、通常、切断は結合部位内またはその近くの特定の部位で生じる。ほとんどのII型酵素はパリンドローム配列を切断するが、IIa型酵素は非パリンドローム認識配列を認識し、認識配列の外側を切断し、IIb型酵素は、認識配列の外側の両方の部位で二重に配列を切断し、IIs型酵素は、非対称の認識配列を認識し、認識配列から約1~約20ヌクレオチドの定義された距離で一方の側で切断する。IV型制限酵素は、メチル化DNAを標的とする。

30

#### 【0266】

いくつかの実施形態では、外来性ドナー配列は、5'末端及び/または3'末端に、標的ゲノム遺伝子座（例えば、B4GALT1 遺伝子内）におけるヌクレアーゼを介した、またはCasタンパク質を介した切断により生じる1つ以上のオーバーハングに相補的な短い一本鎖領域を有する。これらのオーバーハングは、5'及び3'相同性アームとも呼ばれる。例えば、いくつかの外来性ドナー配列は、5'末端及び/または3'末端に、標的ゲノム遺伝子座の5'及び/または3'標的配列におけるCasタンパク質を介した切断によって作成される1つ以上のオーバーハングに相補的な短い一本鎖領域を有する。いくつかの実施形態では、そのような外来性ドナー配列は、5'末端のみまたは3'末端のみに相補的領域を有する。例えば、そのような外来性ドナー配列のいくつかは、標的ゲノム遺伝子座

40

50

の 5' 標的配列に作成されるオーバーハングに相補的な 5' 末端にのみ、または標的ゲノム遺伝子座の 3' 標的配列に作成されるオーバーハングに相補的な 3' 末端にのみ相補的な領域を有する。他のそのような外来性ドナー配列は、5' 及び 3' 末端の両方に相補的な領域を有する。例えば、他のそのような外来性ドナー配列は、5' 及び 3' 末端の両方に、例えば、標的ゲノム遺伝子座における Cas を介した切断により生成される、それぞれ第 1 及び第 2 のオーバーハングに相補的な相補的領域を有する。例えば、外来性ドナー配列が二本鎖である場合、一本鎖相補領域は、ドナー配列の上側鎖の 5' 末端及びドナー配列の下側鎖の 5' 末端から延伸し、それぞれの末端に 5' オーバーハングを形成することができる。あるいは、一本鎖相補領域は、ドナー配列の上側鎖の 3' 末端及び鋸型の下側鎖の 3' 末端から延伸し、3' オーバーハングを形成することができる。

10

#### 【 0 2 6 7 】

相補的領域は、外来性ドナー配列と内在性 B 4 G A L T 1 遺伝子とのライゲーションを促進するのに十分な任意の長さとすることができます。例示的な相補的領域は、長さが約 1 ~ 約 5 ヌクレオチド、長さが約 1 ~ 約 2 5 ヌクレオチド、または長さが約 5 ~ 約 1 5 0 ヌクレオチドである。例えば、相補的領域は、少なくとも約 1、少なくとも約 2、少なくとも約 3、少なくとも約 4、少なくとも約 5、少なくとも約 6、少なくとも約 7、少なくとも約 8、少なくとも約 9、少なくとも約 1 0、少なくとも約 1 1、少なくとも約 1 2、少なくとも約 1 3、少なくとも約 1 4、少なくとも約 1 5、少なくとも約 1 6、少なくとも約 1 7、少なくとも約 1 8、少なくとも約 1 9、少なくとも約 2 0、少なくとも約 2 1、少なくとも約 2 2、少なくとも約 2 3、少なくとも約 2 4、または少なくとも約 2 5 ヌクレオチド長とすることができます。あるいは、相補的領域は、約 5 ~ 約 1 0、約 1 0 ~ 約 2 0、約 2 0 ~ 約 3 0、約 3 0 ~ 約 4 0、約 4 0 ~ 約 5 0、約 5 0 ~ 約 6 0、約 6 0 ~ 約 7 0、約 7 0 ~ 約 8 0、約 8 0 ~ 約 9 0、約 9 0 ~ 約 1 0 0、約 1 0 0 ~ 約 1 1 0、約 1 1 0 ~ 約 1 2 0、約 1 2 0 ~ 約 1 3 0、約 1 3 0 ~ 約 1 4 0、約 1 4 0 ~ 約 1 5 0 ヌクレオチド長、またはそれ以上とすることができます。

20

#### 【 0 2 6 8 】

そのような相補的領域は、ニッカーゼの 2 つのペアによって作成されるオーバーハングに相補的にすることができます。DNA の正反対の側の鎖を切断する第 1 及び第 2 のニッカーゼを使用して第 1 の二重鎖切断を作成し、DNA の正反対の側の鎖を切断する第 3 及び第 4 のニッカーゼを使用して第 2 の二重鎖切断を作成することにより、付着末端を有する 2 つの二重鎖切断を作成することができます。例えば、Cas タンパク質を使用して、第 1、第 2、第 3、及び第 4 のガイド RNA に対応する第 1、第 2、第 3、及び第 4 のガイド RNA 認識配列にニックを導入することができる。DNA の第 1 及び第 2 の鎖上の第 1 及び第 2 のニッカーゼによって作成されるニックが二本鎖切断を作成するように第 1 及び第 2 のガイド RNA 認識配列を配置して、第 1 の切断部位を作成することができる（すなわち、第 1 の切断部位には、第 1 及び第 2 のガイド RNA 認識配列内のニックが含まれる）。同様に、DNA の第 1 及び第 2 の鎖上の第 3 及び第 4 のニッカーゼによって作成されるニックが二本鎖切断を作成するように第 3 及び第 4 のガイド RNA 認識配列を配置して、第 2 の切断部位を作成することができる（すなわち、第 2 の切断部位には、第 3 及び第 4 のガイド RNA 認識配列内のニックが含まれる）。いくつかの実施形態では、第 1 及び第 2 のガイド RNA 認識配列及び / または第 3 及び第 4 のガイド RNA 認識配列内のニックは、オーバーハングを作成するオフセットニックとすることができます。オフセットウインドウは、例えば、少なくとも約 5 bp、少なくとも約 1 0 bp、少なくとも約 2 0 bp、少なくとも約 3 0 bp、少なくとも約 4 0 bp、少なくとも約 5 0 bp、少なくとも約 6 0 bp、少なくとも約 7 0 bp、少なくとも約 8 0 bp、少なくとも約 9 0 bp、もしくは少なくとも約 1 0 0 bp またはそれ以上とすることができます。そのような実施形態では、二本鎖外来性ドナー配列を、第 1 及び第 2 のガイド RNA 認識配列内のニックによって、及び第 3 及び第 4 のガイド RNA 認識配列内のニックによって作成されるオーバーハングに相補的な一本鎖相補領域を用いて設計することができる。次いで、そのような外来性ドナー配列を、非相同末端結合を介したライゲーションにより挿入することができる。

30

40

50

## 【0269】

いくつかの実施形態では、外来性ドナー配列（すなわち、ターゲティングベクター）は相同性アームを含む。外来性ドナー配列が核酸インサートも含む場合、相同性アームを核酸インサートに隣接させることができる。参照を容易にするために、相同性アームを、本明細書では5'及び3'（すなわち、上流及び下流）相同性アームと呼ぶ。この用語は、外来性ドナー配列内の核酸インサートに対する相同性アームの相対的位置に関するものである。

## 【0270】

相同性アームと標的配列は、その2つの領域が相同組換え反応の基質として作用するのに十分なレベルの配列同一性を互いにに対して共有している場合、互いに対応する。特定の標的配列と、外来性ドナー配列に見出される対応する相同性アームとの間の配列同一性は、相同組換えの発生を可能にする任意の程度の配列同一性とすることができます。例えば、外来性ドナー配列の相同性アーム（またはその断片）及び標的配列（またはその断片）によって共有される配列同一性の量は、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、または100%の配列同一性とすることができます。その結果、配列は相同組換えを受ける。さらに、相同性アームと対応する標的配列との間の対応する相同性領域は、相同組換えを促進するのに十分な任意の長さとすることができます。例示的な相同性アームは、約25ヌクレオチド～約2.5kbの長さ、約25ヌクレオチド～約1.5kbの長さ、または約25～約500ヌクレオチドの長さである。例えば、所定の相同性アーム（または各相同性アーム）及び/または対応する標的配列には、約25～約30、約30～約40、約40～約50、約50～約60、約60～約70、約70～約80、約80～約90、約90～約100、約100～約150、約150～約200、約200～約250、約250～約300、約300～約350、約350～約400、約400～約450、または約450～約500ヌクレオチド長の対応する相同性領域を含めることができます。それにより、相同性アームは、内在性B4GALT1遺伝子内の対応する標的配列と相同組換えを行うのに十分な相同性を有する。あるいは、特定の相同性アーム（または各相同性アーム）及び/または対応する標的配列には、約0.5kb～約1kb、約1kb～約1.5kb、約1.5kb～約2kb、または約2kb～約2.5kbの長さの対応する相同性領域を含めることができます。例えば、相同性アームを、それぞれ約750ヌクレオチド長とすることができます。相同性アームは、対称（長さがほぼ同じサイズ）とすることも、非対称（一方が他方より長い）とすることもできる。

## 【0271】

相同性アームは、細胞に固有の遺伝子座（例えば、標的遺伝子座）に対応させることができる。あるいは、それらを、例えば導入遺伝子、発現カセット、またはDNAの異種または外来性領域を含む、細胞のゲノムに組み込んだDNAの異種または外来性セグメントの領域に対応させることができる。いくつかの実施形態では、ターゲティングベクターの相同性アームは、酵母人工染色体（YAC）、細菌人工染色体（BAC）、ヒト人工染色体の領域、または適切な宿主細胞に含まれる他の変更した領域に対応させることができる。いくつかの実施形態では、ターゲティングベクターの相同性アームは、BACライブラリー、コスミドライブラリー、もしくはP1ファージライブラリーの領域に対応するか、もしくはそれに由来するものとすることができます。または合成DNAに由来するものとすることができます。

## 【0272】

外来性ドナー配列と組み合わせてヌクレアーゼ剤を使用する場合、5'及び3'標的配列は一般的に、ヌクレアーゼ切断部位に十分に近接して位置しており、それにより、ヌクレ

10

20

30

40

50

アーゼ切断部位での一本鎖切断（ニック）または二本鎖切断時の標的配列と相同性アームとの間の相同組換え事象の発生を促進する。ヌクレアーゼ切断部位には、ヌクレアーゼ剤（例えば、ガイドRNAと複合体化したCas9タンパク質）によってニックまたは二本鎖切断が生じるDNA配列が含まれる。外来性ドナー配列の5'及び3'相同性アームに対応する内在性B4GALT1遺伝子内の標的配列は、距離が、ヌクレアーゼ切断部位での一本鎖切断または二本鎖切断時の5'及び3'標的配列と相同性アームとの間の相同組換え事象の発生を促進するようなものである場合、ヌクレアーゼ切断部位に「十分に近接して位置する」。したがって、外来性ドナー配列の5'及び/または3'相同性アームに対応する標的配列は、例えば、所定のヌクレアーゼ切断部位の少なくとも1ヌクレオチド以内、または特定のヌクレアーゼ切断部位の少なくとも10ヌクレオチド～約1,000ヌクレオチド以内とすることができます。いくつかの実施形態では、ヌクレアーゼ切断部位を、標的配列の少なくとも一方または両方に直接隣接させることができます。10

#### 【0273】

外来性ドナー配列の相同性アーム及びヌクレアーゼ切断部位に対応する標的配列の空間的関係は様々に異なり得る。いくつかの実施形態では、標的配列をヌクレアーゼ切断部位の5'に位置させることができるか、標的配列をヌクレアーゼ切断部位の3'に位置させることができるか、または標的配列をヌクレアーゼ切断部位に隣接させることができる。20

#### 【0274】

本開示はまた、内在性B4GALT1遺伝子の発現を改変または変化させるための本明細書に開示する方法を用いた、疾患を有するかまたは有するリスクのある対象における心血管病態の治療方法及び処置または予防方法を提供する。本開示はまた、内在性B4GALT1mRNAの発現を低下させる方法を使用するか、または対象にB4GALT1ポリペプチドをコードする組換え核酸を提供するか、B4GALT1ポリペプチドをコードするmRNAを提供するか、もしくはB4GALT1ポリペプチドを提供する方法を使用して、疾患を有するかまたは有するリスクのある対象における心血管病態の治療方法及び処置または予防方法を提供する。本方法は、1つ以上の核酸分子またはタンパク質を対象内に、対象の器官内に、または対象の細胞内に（例えば、in vivoまたはex vivoで）導入することを含み得る。20

#### 【0275】

いくつかの実施形態では、本開示は、治療に使用するためのB4GALT1ポリペプチドをコードするmRNA（例えば、本明細書で論じるポリヌクレオチド、例えば配列番号4の配列を含むmRNA）を提供する。いくつかのそのような実施形態では、療法は、心血管病態を治療するかまたは予防することである。30

#### 【0276】

いくつかの実施形態では、本開示は、治療に使用するためのB4GALT1ポリペプチド（例えば、本明細書で論じるポリペプチド、例えば配列番号8の配列を含むポリペプチド）を提供する。いくつかのそのような実施形態では、療法は、心血管病態を治療するかまたは予防することである。

#### 【0277】

対象には、予防的または治療的処置を受けるヒト及び他の哺乳類対象（例えば、ネコ、イヌ、げっ歯類、マウス、またはラット）または非哺乳類対象（例えば、家禽類）が含まれる。そのような対象は、例えば、バリアント型B4GALT1の保因者ではなく（またはバリアント型B4GALT1のヘテロ接合保因者に過ぎず）、心血管病態を発症しているかまたは発症しやすい対象（例えば、ヒト）であり得る。40

#### 【0278】

心血管病態の非限定的な例として、1つ以上の血清脂質のレベルの上昇が挙げられる。血清脂質には、コレステロール、LDL、HDL、トリグリセリド、HDLコレステロール、及び非HDLコレステロール、またはその任意の亜分画（例えば、HDL2、HDL2a、HDL2b、HDL2c、HDL3、HDL3a、HDL3b、HDL3c、HDL3d、LDL1、LDL2、LDL3、リポタンパク質A、Lp a1、Lp a1、Lp

a 3、L p a 4、またはL p a 5)の1つ以上が含まれる。心血管病態には、冠動脈石灰化レベルの上昇が含まれ得る。心血管病態には、I I d型グリコシル化(C D G - I I d)が含まれ得る。心血管病態には、心膜脂肪のレベルの上昇が含まれ得る。心血管病態には、アテローム血栓性病態が含まれ得る。アテローム血栓性病態には、フィブリノーゲンレベルの上昇が含まれ得る。アテローム血栓性病態には、フィブリノーゲン媒介性血餅が含まれ得る。心血管病態には、フィブリノーゲンレベルの上昇が含まれ得る。心血管病態には、フィブリノーゲン媒介性血餅が含まれ得る。心血管病態には、フィブリノーゲン活性の関与から形成される血餅が含まれ得る。フィブリノーゲン媒介性血餅またはフィブリノーゲン活性の関与から形成される血餅は、体内のどの静脈または動脈にも存在し得る。

## 【0279】

10

そのような方法は、ゲノム編集または遺伝子治療を含み得る。例えば、バリアント型B 4 G A L T 1ではない内在性B 4 G A L T 1遺伝子を改変して、バリアント型B 4 G A L T 1に関連する変異(すなわち、完全長/成熟型B 4 G A L T 1ポリペプチドの352位に対応する位置でのアスパラギンからセリンへの置換)を含ませることができる。別の例として、バリアント型B 4 G A L T 1ではない内在性B 4 G A L T 1遺伝子をノックアウトまたは不活性化することができる。同様に、バリアント型B 4 G A L T 1ではない内在性B 4 G A L T 1遺伝子をノックアウトまたは不活性化することができ、バリアント型B 4 G A L T 1に関連する改変を含むB 4 G A L T 1遺伝子(例えば、完全なバリアント型B 4 G A L T 1または改変を含むミニ遺伝子)を導入して発現させることができる。同様に、バリアント型B 4 G A L T 1ではない内在性B 4 G A L T 1遺伝子をノックアウトまたは不活性化することができ、B 4 G A L T 1バリアントポリペプチドをコードする組換えDNAを導入し、発現させることができ、B 4 G A L T 1バリアントポリペプチドをコードするmRNAを導入し、発現させることができ(例えば、細胞内タンパク質置換療法)、及び/またはバリアント型B 4 G A L T 1ポリペプチドを導入することができる(例えば、タンパク質置換療法)。

## 【0280】

20

いくつかの実施形態では、方法は、バリアント型B 4 G A L T 1ではない内在性B 4 G A L T 1遺伝子をノックアウトまたは不活性化することなく、B 4 G A L T 1 rs551564683バリアント(例えば、完全なバリアント型B 4 G A L T 1または改変を含むミニ遺伝子)に関連する改変を含む組換えB 4 G A L T 1遺伝子を導入し、発現させること、バリアント型B 4 G A L T 1ポリペプチドまたはその断片をコードする組換え核酸(例えば、DNA)を導入し、発現させること、バリアント型B 4 G A L T 1ポリペプチドまたはその断片をコードする1つ以上のmRNAを導入し、発現させること(例えば、細胞内タンパク質置換療法)、またはバリアント型B 4 G A L T 1ポリペプチドまたはその断片を導入すること(例えば、タンパク質置換療法)を含む。いくつかの実施形態では、そのような方法を、アンチセンスRNA、siRNA、またはshRNAの使用などにより、バリアント型B 4 G A L T 1ではない内在性B 4 G A L T 1 mRNAを発現低下させるための標的とする方法と組み合わせて実施することもできる。

30

## 【0281】

40

B 4 G A L T 1遺伝子もしくはミニ遺伝子もしくはバリアント型B 4 G A L T 1ポリペプチドをコードするDNAまたはその断片は、ゲノムを改変しない発現ベクターの形態で導入し、発現させることができ、内在性B 4 G A L T 1遺伝子座にゲノム的に組み込むようにターゲティングベクターの形態で導入することができ、またはセーフハーバー遺伝子座などの内在性B 4 G A L T 1遺伝子座以外の遺伝子座にゲノム的に組み込むように導入することができる。ゲノムに組み込むB 4 G A L T 1遺伝子は、B 4 G A L T 1プロモーターまたは組み込み部位における内在性プロモーターなどの別のプロモーターに作動可能に連結することができる。セーフハーバー遺伝子座は、遺伝子構造または発現に悪影響を与えることなく、目的のすべての組織で導入遺伝子を安定かつ確実に発現させることができる染色体部位である。セーフハーバー遺伝子座は、例えば、次の特徴の1つ以上またはすべてを有し得る: 1)任意の遺伝子の5'末端から約50kbを上回る距離、がん関連遺

50

伝子から約 300 kb を超える距離、マイクロ RNA から約 300 kb を超える距離、遺伝子転写ユニットの外側、及び超保存領域の外側。適切なセーフハーバー遺伝子座の例として、アデノ随伴ウイルス部位 1 (AAVS1)、ケモカイン (CCモチーフ) 受容体 5 (CCR5) 遺伝子座、及びマウス ROSA26 遺伝子座のヒトオーソログが挙げられるが、これらに限定されない。

【0282】

いくつかの実施形態では、方法は、バリアント型 B4GALT1 の保因者ではなく（またはバリアント型 B4GALT1 のヘテロ接合保因者に過ぎず）、心血管病態を発症しているかまたは発症しやすい対象の治療方法を含み、方法は、対象内へ、または対象の細胞内へ：a) 内在性 B4GALT1 遺伝子内のヌクレアーゼ認識配列に結合するヌクレアーゼ剤（またはそれをコードする核酸）を導入することを含み、その場合、ヌクレアーゼ認識配列は、配列番号 1 の 53575 ~ 53577 位を含むかまたはそれに近接しており、ならびに b) 配列番号 1 の 53575 ~ 53577 位の 5' の標的配列にハイブリダイズする 5' 相同性アーム、ならびにセリンをコードする核酸配列を含み、5' 相同性アーム及び 3' 相同性アームに隣接する核酸インサートを含む外来性ドナー配列を導入することを含む。ヌクレアーゼ剤は、対象の細胞の内在性 B4GALT1 遺伝子を切断することができ、外来性ドナー配列は、細胞の内在性 B4GALT1 遺伝子と組み換わることができ、外来性ドナー配列の内在性 B4GALT1 遺伝子との組換え時に、配列番号 1 の 53575 ~ 53577 位に対応するヌクレオチドにおいて、セリンをコードする核酸配列が挿入される。そのような方法で使用することができるヌクレアーゼ剤（例えば、Cas9タンパク質及びガイド RNA）の例を、本明細書の他の場所に開示する。

10

【0283】

いくつかの実施形態では、方法は、バリアント型 B4GALT1 の保因者ではなく（またはバリアント型 B4GALT1 のヘテロ接合保因者に過ぎず）、心血管病態を発症しているかまたは発症しやすい対象の治療方法を含み、方法は、対象内へ、または対象の細胞内へ、配列番号 1 の 53575 ~ 53577 位に対応する位置の 5' の標的配列にハイブリダイズする 5' 相同性アーム、配列番号 1 の 53575 ~ 53577 位の 3' の標的配列にハイブリダイズする 3' 相同性アーム、ならびにセリンをコードするヌクレオチド配列を含み、5' 相同性アーム及び 3' 相同性アームに隣接する核酸インサートを含む外来性ドナー配列を導入することを含む。外来性ドナー配列は、細胞の内在性 B4GALT1 遺伝子と組み換わることができ、外来性ドナー配列の内在性 B4GALT1 遺伝子との組換え時に、配列番号 1 の 53575 ~ 53577 位に対応するヌクレオチドにおいて、セリンをコードするヌクレオチド配列が挿入される。

20

【0284】

いくつかのそのような方法は、バリアント型 B4GALT1 の保因者ではなく（またはバリアント型 B4GALT1 のヘテロ接合保因者に過ぎず）、心血管病態を発症しているかまたは発症しやすい対象の治療方法を含み、方法は、対象内へ、または対象の細胞内へ：a) 内在性 B4GALT1 遺伝子内のヌクレアーゼ認識配列に結合するヌクレアーゼ剤（またはそれをコードする核酸）を導入することを含み、その場合、ヌクレアーゼ認識配列は、内在性 B4GALT1 遺伝子の開始コドンを含むか、または開始コドンの、約 100、約 200、約 300、約 400、約 500、約 1000、約 2000、約 3000、約 4000、約 5000、または約 1,000 ヌクレオチド以内であるか、または配列番号 9 ~ 12 から選択される。ヌクレアーゼ剤は、対象の細胞の内在性 B4GALT1 遺伝子を切断し、発現を破壊することができる。

30

【0285】

いくつかの実施形態では、方法は、バリアント型 B4GALT1 の保因者ではなく（またはバリアント型 B4GALT1 のヘテロ接合保因者に過ぎず）、心血管病態を発症しているかまたは発症しやすい対象の治療方法を含み、方法は、対象内へ、または対象の細胞内へ：a) 内在性 B4GALT1 遺伝子内のヌクレアーゼ認識配列に結合するヌクレアーゼ剤（またはそれをコードする核酸）を導入することを含み、その場合、ヌクレアーゼ認

40

50

識配列は、内在性 B 4 G A L T 1 遺伝子の開始コドンを含むか、または、開始コドンの、約 10 以内、約 20 以内、約 30 以内、約 40 以内、約 50 以内、約 100 以内、約 200 以内、約 300 以内、約 400 以内、約 500 以内、または約 1,000 ヌクレオチド以内コドンであるか、または配列番号 9 ~ 12 から選択され、b ) 完全長 / 成熟型 B 4 G A L T 1 ポリペプチドの 352 位に対応する位置でセリンをコードするヌクレオチド配列を 53575 ~ 53577 位に含む組換え B 4 G A L T 1 遺伝子を含む発現ベクターを導入することを含む。発現ベクターは、ゲノムに組み込まれないベクターとすることができる。あるいは、完全長 / 成熟型 B 4 G A L T 1 ポリペプチドの 352 位に対応する位置でセリンをコードするヌクレオチド配列を 53575 ~ 53577 位に含む組換え B 4 G A L T 1 遺伝子を含むターゲティングベクター（すなわち、外来性ドナー配列）を導入することができる。ヌクレアーゼ剤は、対象の細胞内の B 4 G A L T 1 遺伝子内を切断し、発現を破壊することができ、発現ベクターは、対象の細胞内で組換え B 4 G A L T 1 遺伝子を発現することができる。あるいは、ゲノムに組み込んだ組換え B 4 G A L T 1 遺伝子を、対象の細胞内で発現させることができる。そのような方法で使用することができるヌクレアーゼ剤（例えば、ヌクレアーゼ活性 C a s 9 タンパク質及びガイド R N A ）の例を、本明細書の他の場所に開示する。適切なガイド R N A 及びガイド R N A 認識配列の例もまた、本明細書の他の場所に開示する。工程 b ) は、バリアント型 B 4 G A L T 1 A s n 352 S e r ポリペプチドまたはその断片と、少なくとも 90 %、少なくとも 95 %、少なくとも 96 %、少なくとも 97 %、少なくとも 98 %、少なくとも 99 %、または 100 % 同一である B 4 G A L T 1 ポリペプチドをコードする核酸（例えば、D N A ）を含み、及び / またはバリアント型 B 4 G A L T 1 m R N A またはその断片と、少なくとも 90 %、少なくとも 95 %、少なくとも 96 %、少なくとも 97 %、少なくとも 98 %、少なくとも 99 %、または 100 % 同一である配列を含む発現ベクターまたはターゲティングベクターを導入することを代替的に含み得る。同様に、工程 b ) はまた、バリアント型 B 4 G A L T 1 A s n 352 S e r ポリペプチドまたはその断片と、少なくとも 90 %、少なくとも 95 %、少なくとも 96 %、少なくとも 97 %、少なくとも 98 %、少なくとも 99 %、または 100 % 同一であり、及び / またはバリアント型 B 4 G A L T 1 m R N A またはその断片と、少なくとも 90 %、少なくとも 95 %、少なくとも 96 %、少なくとも 97 %、少なくとも 98 %、少なくとも 99 %、または 100 % 同一である相補的 D N A （またはその部分）を有する B 4 G A L T 1 A s n 352 S e r ポリペプチドをコードする m R N A を導入することを含み得る。同様に、工程 b ) はまた、バリアント型 B 4 G A L T 1 A s n 352 S e r ポリペプチドまたはその断片と、少なくとも 90 %、少なくとも 95 %、少なくとも 96 %、少なくとも 97 %、少なくとも 98 %、少なくとも 99 %、または 100 % 同一であるアミノ酸配列を含むタンパク質を導入することを含み得る。

#### 【 0286 】

いくつかの実施形態では、さらに第 2 のヌクレアーゼ剤を対象へ、または対象の細胞内へ導入し、第 2 のヌクレアーゼ剤は、内在性 B 4 G A L T 1 遺伝子内の第 2 のヌクレアーゼ認識配列に結合し、第 2 のヌクレアーゼ認識配列は、内在性 B 4 G A L T 1 遺伝子の終止コドンを含むか、または終止コドンの、約 10 以内、約 20 以内、約 30 以内、約 40 以内、約 50 以内、約 100 以内、約 200 以内、約 300 以内、約 400 以内、約 500 以内、または約 1,000 ヌクレオチド以内であるか、または配列番号 9 ~ 12 から選択され、その場合、ヌクレアーゼ剤は、細胞の内在性 B 4 G A L T 1 遺伝子を、第 1 のヌクレアーゼ認識配列及び第 2 のヌクレアーゼ認識配列の両方において切断し、第 1 のヌクレアーゼ認識配列と第 2 のヌクレアーゼ認識配列との間に欠失を含むように細胞が改変される。いくつかの実施形態では、第 2 のヌクレアーゼ剤を、C a s 9 タンパク質及びガイド R N A とすることができる。終止コドンに近接した適切なガイド R N A 及びガイド R N A 認識配列を、本明細書の他の場所に開示する。

#### 【 0287 】

いくつかの実施形態では、方法はまた、バリアント型 B 4 G A L T 1 の保因者ではなく

10

20

30

40

50

(またはバリアント型B4GALT1のヘテロ接合保因者に過ぎず)、心血管病態を発症しているかまたは発症しやすい対象の治療方法を含むことができ、方法は、対象内へ、または対象の細胞内へ：内在性B4GALT1 mRNA内の領域内の配列にハイブリダイズするアンチセンスRNA、siRNA、またはshRNAを導入することを含む。例えば、アンチセンスRNA、siRNA、またはshRNAは、配列番号3(B4GALT1 mRNA)のエクソン5の領域内の配列にハイブリダイズし、対象の細胞内でのB4GALT1 mRNAの発現を低下させることができる。いくつかの実施形態では、そのような方法は、配列番号2の53575～53577位に挿入されたセリンをコードするヌクレオチド配列を含む組換えB4GALT1遺伝子を含む発現ベクターを対象に導入することをさらに含み得る。発現ベクターは、ゲノムに組み込まれないベクターとすることができます。あるいは、配列番号2の53575～53577位に対応する位置にセリンをコードする核酸配列を含む組換えB4GALT1遺伝子を含むターゲティングベクター(すなわち、外来性ドナー配列)を導入することができる。発現ベクターを使用する方法では、発現ベクターは、対象の細胞内で組換えB4GALT1遺伝子を発現することができる。あるいは、組換えB4GALT1遺伝子をゲノムに組み込む方法では、組換えB4GALT1遺伝子を対象の細胞内で発現させることができる。

#### 【0288】

いくつかの実施形態では、そのような方法は、バリアント型B4GALT1 Asn352Serポリペプチドまたはその断片と、少なくとも90%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、または100%同一であるB4GALT1ポリペプチドをコードする核酸(例えばDNA)を含み、及び/またはバリアント型B4GALT1 mRNAまたはその断片と、少なくとも90%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、または100%同一である配列を含む発現ベクターまたはターゲティングベクターを導入することを代替的に含み得る。同様に、そのような方法は、バリアント型B4GALT1 Asn352Serポリペプチドまたはその断片と、少なくとも90%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、または100%同一であり、及び/またはバリアント型B4GALT1 mRNAまたはその断片と、少なくとも90%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、または100%同一である相補的DNA(またはその部分)を有する、ポリペプチドをコードするmRNAを導入することを代替的に含み得る。同様に、そのような方法は、バリアント型B4GALT1 Asn352Serポリペプチドまたはその断片と、少なくとも90%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、または100%同一である配列を含むポリペプチドを導入することを代替的に含み得る。

#### 【0289】

いくつかの実施形態では、そのような方法は、バリアント型B4GALT1の保因者ではなく(またはバリアント型B4GALT1のヘテロ接合保因者に過ぎず)、心血管病態を発症しているかまたは発症しやすい対象の治療方法を含むことができ、方法は、対象内へ、または対象の細胞内へ発現ベクターを導入することを含み、その場合、発現ベクターは、完全長/成熟型B4GALT1ポリペプチドの352位に対応する位置にセリンをコードするヌクレオチド配列を53575～53577位に含む組換えB4GALT1遺伝子を含み、発現ベクターは、対象の細胞内で組換えB4GALT1遺伝子を発現する。発現ベクターは、ゲノムに組み込まれないベクターとすることができます。あるいは、完全長/成熟型B4GALT1ポリペプチドの352位に対応する位置にセリンをコードするヌクレオチド配列を配列番号2の53575～53577位に含む組換えB4GALT1遺伝子を含むターゲティングベクター(すなわち、外来性ドナー配列)を導入することができる。発現ベクターを使用する方法では、発現ベクターは、対象の細胞内で組換えB4GALT1遺伝子を発現することができる。あるいは、組換えB4GALT1遺伝子をゲノムに組み込む方法では、組換えB4GALT1遺伝子を対象の細胞内で発現させることができます。

10

20

30

40

50

できる。

【0290】

そのような方法は、バリアント型 B 4 G A L T 1 A s n 3 5 2 S e r ポリペプチドまたはその断片と、少なくとも 9 0 %、少なくとも 9 5 %、少なくとも 9 6 %、少なくとも 9 7 %、少なくとも 9 8 %、少なくとも 9 9 %、または 1 0 0 % 同一であり、及び / またはバリアント型 B 4 G A L T 1 m R N A またはその断片と、少なくとも 9 0 %、少なくとも 9 5 %、少なくとも 9 6 %、少なくとも 9 7 %、少なくとも 9 8 %、少なくとも 9 9 %、または 1 0 0 % 同一である配列を含む、B 4 G A L T 1 ポリペプチドをコードする核酸（例えば、D N A）を含む発現ベクターまたはターゲティングベクターを導入することを代替的に含み得る。同様に、そのような方法は、バリアント型 B 4 G A L T 1 ポリペプチドまたはその断片と、少なくとも 9 0 %、少なくとも 9 5 %、少なくとも 9 6 %、少なくとも 9 7 %、少なくとも 9 8 %、少なくとも 9 9 %、または 1 0 0 % 同一であり、及び / またはバリアント型 B 4 G A L T 1 m R N A またはその断片と、少なくとも 9 0 %、少なくとも 9 5 %、少なくとも 9 6 %、少なくとも 9 7 %、少なくとも 9 8 %、少なくとも 9 9 %、または 1 0 0 % 同一である相補的 D N A（またはその部分）を有する、ポリペプチドをコードする m R N A を導入することを代替的に含み得る。同様に、そのような方法は、バリアント型 B 4 G A L T 1 A s n 3 5 2 S e r ポリペプチドまたはその断片と、少なくとも 9 0 %、少なくとも 9 5 %、少なくとも 9 6 %、少なくとも 9 7 %、少なくとも 9 8 %、少なくとも 9 9 %、または 1 0 0 % 同一である配列を含むタンパク質を導入することを代替的に含み得る。

10

【0291】

上記の方法のいずれかで使用するのに適した発現ベクター及び組換え B 4 G A L T 1 遺伝子を、本明細書の他の場所に開示する。例えば、組換え B 4 G A L T 1 遺伝子は、完全な B 4 G A L T 1 バリアント遺伝子であり得るか、または対応する野生型 B 4 G A L T 1 遺伝子に対して遺伝子の 1 つ以上の非必須セグメントを欠失させた B 4 G A L T 1 ミニ遺伝子であり得る。例として、欠失させるセグメントには 1 つ以上のイントロン配列を含ませることができ、ミニ遺伝子にはエクソン 1 ~ 6 を含ませることができる。完全な B 4 G A L T 1 バリアント遺伝子の例は、配列番号 2 と、少なくとも 9 0 %、少なくとも 9 5 %、少なくとも 9 6 %、少なくとも 9 7 %、少なくとも 9 8 %、少なくとも 9 9 %、または 1 0 0 % 同一である遺伝子である。

20

【0292】

いくつかの実施形態では、そのような方法は、心血管病態を発症しているか、または心血管病態を発症しやすい対象の細胞の改変方法を含む。そのような方法では、ヌクレアーゼ剤及び / または外来性ドナー配列及び / または組換え発現ベクターを、発症を遅らせ、重症度を軽減し、さらなる悪化を抑制し、及び / または治療中の心血管病態の少なくとも 1 つの徴候または症状を改善する用量、投与経路及び投与頻度を意味する有効なレジームでの投与を介して、細胞内に導入することができる。用語「症状」とは、対象によって知覚される疾患の主観的エビデンスを指し、「徴候」とは、医師によって観察される疾患の客観的エビデンスを指す。対象が既に疾患に罹患している場合、そのレジームは治療有効レジームと呼ばれ得る。対象が、一般的な集団に比べて疾患のリスクが高いが、まだ症状を経験していない場合、レジームは予防有効レジームと呼ばれ得る。いくつかの例では、同じ対象の過去の対照または過去の経験と比較して、個々の患者における治療または予防の有効性を観察することができる。他の例では、治療または予防の有効性を、未治療の対象の対照集団に対する、治療対象の集団における前臨床試験または臨床試験において示すことができる。

30

【0293】

送達は、本明細書の他の箇所で開示するように、任意の適切な方法とすることができます。例えば、ヌクレアーゼ剤または外来性ドナー配列または組換え発現ベクターは、例えば、ベクター送達、ウイルス送達、粒子を介した送達、ナノ粒子を介した送達、リポソームを介した送達、エキソソームを介した送達、脂質を介した送達、脂質ナノ粒子を介した送

40

50

達、細胞透過性ペプチドを介した送達、または埋め込み可能装置を介した送達により送達することができる。特定の例には、流体力学的送達、ウイルスを介した送達、及び脂質ナノ粒子を介した送達が含まれる。

【0294】

投与は、非経口、静脈内、経口、皮下、動脈内、頭蓋内、くも膜下腔内、腹腔内、局所、鼻腔内、または筋肉内を含むがこれらに限定されない任意の適切な経路によるものとすることができる。例えば、タンパク質補充療法によく使用される特定の例は、静脈内注入である。投与頻度及び投与回数は、数ある要因の中でも、スクレアーゼ剤または外来性ドナー配列または組換え発現ベクターの半減期、対象の病態、及び投与経路に依存し得る。投与用の医薬組成物は、望ましくは無菌であり、実質的に等張であり、GMP条件下で製造される。医薬組成物は、単位投与形態（すなわち、単回投与の用量）で提供することができる。医薬組成物は、1つ以上の生理学的及び薬学的に許容される担体、希釈剤、賦形剤または補助剤を使用して製剤化することができる。製剤は、選択した投与経路に依存する。用語「薬学的に許容される」とは、担体、希釈剤、賦形剤、または補助剤が製剤の他の成分と適合し、そのレシピエントに実質的に有害ではないことを意味する。

【0295】

他のそのような方法は、心血管病態を発症しているかまたは発症しやすい対象由来の細胞におけるex vivo法を含む。標的化された遺伝子改変を施した細胞を、その後、対象に移植して戻すことができる。

【0296】

本開示は、本明細書に記載の方法のいずれかにより、内在性の野生型B4GALT1の発現を低下させるか、またはB4GALT1 Asn352Serの発現を増加させることにより、LDLの低下を必要とする対象のLDLを低下させる方法を提供する。本開示は、本明細書に記載の方法のいずれかにより、内在性の野生型B4GALT1の発現を低下させるか、またはB4GALT1 Asn352Serの発現を増加させることにより、総コレステロールの低下を必要とする対象の総コレステロールを低下させる方法を提供する。本開示は、本明細書に記載の方法のいずれかにより、内在性の野生型B4GALT1の発現を低下させるか、またはB4GALT1 Asn352Serの発現を増加させることにより、フィブリノーゲンの低下を必要とする対象のフィブリノーゲンを低下させる方法を提供する。本開示は、本明細書に記載の方法のいずれかにより、内在性の野生型B4GALT1の発現を低下させるか、またはB4GALT1 Asn352Serの発現を増加させることにより、eGFRを低下させる方法を提供する。本開示は、本明細書に記載の方法のいずれかにより、内在性の野生型B4GALT1の発現を低下させるか、またはB4GALT1 Asn352Serの発現を増加させることにより、ASTの増加を必要とする対象のASTを増加させるがALTは増加させない方法を提供する。本開示は、本明細書に記載の方法のいずれかにより、内在性の野生型B4GALT1の発現を低下させるか、またはB4GALT1 Asn352Serの発現を増加させることにより、クレアチニンの増加を必要とする対象のクレアチニンを増加させる方法を提供する。

【0297】

本開示はまた、心血管病態の発症リスクの診断、または心血管病態の発症リスクの診断及び治療を必要とする対象における心血管病態の発症リスクの診断方法、または心血管病態の発症リスクの診断方法及び治療方法を提供し、方法は：本明細書に記載のバリアント型B4GALT1遺伝子、mRNA、cDNA、またはポリペプチドの存在または非存在に関する対象由来試料の分析結果を提供する試験を要求し、バリアント型B4GALT1遺伝子、mRNA、cDNA、またはポリペプチドを有さない対象において、本明細書に記載するような治療薬を対象に投与することを含む。バリアント型B4GALT1遺伝子、mRNA、cDNA、またはポリペプチドの存在または非存在を判定する本明細書に記載の試験のいずれかを使用することができる。

【0298】

10

20

30

40

50

本開示はまた、LDLの低下、総コレステロールの低下、フィブリノーゲンの低下、eGFRの低下、ASTの増加、及びクレアチニンの増加を必要とする対象において、LDLを低下させ、総コレステロールを低下させ、フィブリノーゲンを低下させ、eGFRを低下させ、ASTを増加させ（ただしALTは増加させない）、及びクレアチニンを増加させるための医薬品の製造における、本明細書に開示するバリアント型B4GALT1遺伝子、mRNA、cDNA、ポリペプチド、及びハイブリダイズする核酸分子のいずれかの使用を提供する。本開示はまた、冠動脈疾患、冠動脈石灰化、及び関連障害を治療するための医薬品の製造における、バリアント型B4GALT1遺伝子、mRNA、cDNA、ポリペプチド、及びハイブリダイズする核酸分子のいずれかの使用を提供する。

## 【0299】

10

本開示はまた、LDLの低下、総コレステロールの低下、フィブリノーゲンの低下、eGFRの低下、ASTの増加、及びクレアチニンの増加を必要とする対象において、LDLを低下させ、総コレステロールを低下させ、フィブリノーゲンを低下させ、eGFRを低下させ、ASTを増加させ（ただしALTは増加させない）、及びクレアチニンを増加させるための、本明細書に開示するバリアント型B4GALT1遺伝子、mRNA、cDNA、ポリペプチド、及びハイブリダイズする核酸分子のいずれかの使用を提供する。

## 【0300】

本開示はまた、冠動脈疾患、冠動脈石灰化、IId型グリコシル化（CDG-IId）、及び関連障害を治療するための、バリアント型B4GALT1遺伝子、mRNA、cDNA、ポリペプチド、及びハイブリダイズする核酸分子のいずれかの使用を提供する。

20

## 【0301】

本開示はまた、B4GALT1遺伝子の変更を必要とする対象の細胞においてB4GALT1遺伝子を変更するための、本明細書に開示するバリアント型B4GALT1遺伝子、mRNA、cDNA、ポリペプチド、及びハイブリダイズする核酸分子のいずれかの使用を提供する。

## 【0302】

本開示はまた、B4GALT1遺伝子の発現の変化を必要とする対象の細胞においてB4GALT1遺伝子の発現を変化させるための、本明細書に開示するバリアント型B4GALT1遺伝子、mRNA、cDNA、ポリペプチド、及びハイブリダイズする核酸分子のいずれかの使用を提供する。

30

## 【0303】

本開示はまた、本明細書に開示する心血管病態のいずれかの発症リスクを診断するための、本明細書に開示するバリアント型B4GALT1遺伝子、mRNA、cDNA、ポリペプチド、及びハイブリダイズする核酸分子のいずれかの使用を提供する。

## 【0304】

本開示はまた、本明細書に開示する心血管病態のいずれかを発症している対象を診断するための、本明細書に開示するバリアント型B4GALT1遺伝子、mRNA、cDNA、ポリペプチド、及びハイブリダイズする核酸分子のいずれかの使用を提供する。

## 【0305】

上記または下記で引用するすべての特許文書、ウェブサイト、他の刊行物、登録番号などは、あたかも個々の項目をそのように参照により援用することを具体的かつ個別に示すことと同じ程度に、あらゆる目的のためにその全体を参照により援用するものとする。異なるバージョンの配列が異なる時点での登録番号に関連付けられている場合、本出願の有効出願日における登録番号に関連付けられているバージョンを意図するものとする。有効出願日とは、実際の出願日または優先出願の出願日のうち早い方を意味し、該当する場合は登録番号が参照される。同様に、異なるバージョンの刊行物、ウェブサイトなどが異なる時間に公開される場合、特に明記しない限り、出願の有効出願日において最後に公開されたバージョンを意味するものとする。本開示の任意の特徴、工程、要素、実施形態、または態様は、特に明記しない限り、他の特徴、工程、要素、実施形態、または態様と組み合わせて使用することができる。本開示を、明確化及び理解を目的として、例示及び例とし

40

50

てある程度詳細に説明してきたが、特定の変更及び改変を添付の特許請求の範囲内で実施し得ることは明らかであろう。

【0306】

本明細書に挙げるヌクレオチド及びアミノ酸配列は、ヌクレオチド塩基の標準文字略語、及びアミノ酸の1文字コードを使用して示される。ヌクレオチド配列は、配列の5'末端で始まり、3'末端へ向かって前方に進む（すなわち、各行の左から右へ）という標準の慣習に従う。各ヌクレオチド配列の1本の鎖のみを示すが、表示する鎖の参照によって相補鎖が含まれるものと理解される。アミノ酸配列は、配列のアミノ末端で始まり、カルボキシ末端へ向かって前方に進む（すなわち、各行の左から右へ）という標準的な慣習に従う。

【0307】

2018年4月18日に出願された米国出願第62/659,344号、2017年8月25日に出願された米国出願第62/550,161号、及び2017年6月5日に出願された米国出願第62/515,140号は、各々、その全体を参照により本明細書に援用する。

【0308】

以下の実施例は、実施形態をより詳細に説明するために提供する。それらは、特許請求する実施形態を例示することを意図しており、限定することを意図していない。

【実施例】

【0309】

実施例1：ゲノムワイドの統計的有意性での血清脂質形質に関連する染色体9p.21上の新規遺伝子座の決定

材料及び方法：

チップ・ジェノタイピング及びQC：OOAの個体の全血からゲノムDNAを抽出し、picogreenを使用して定量した。University of Maryland Biopolymer Core Facilityで、Affymetrix 500K及び6.0チップを使用してゲノムワイド・ジェノタイピングを実施した。遺伝子型判定にBRLMMアルゴリズムを使用した。判定率<0.93、高レベルのメンデルの法則のエラー、または性別の不一致のサンプルは除外した。判定率<0.95、HWE pval<1.0E-6、またはMAF<0.01のSNPは除外した。X染色体及びY染色体、ならびにミトコンドリアゲノム上のSNPも除外した。

【0310】

WGS及びQC：ライブラリー調製及び全ゲノム配列決定を、MIT及びハーバードのプロード研究所によって実施した。ミシガン大学のNHLBI Informatics Resource Coreは、すべてのTOPMedサンプルのアライメント、塩基判定、及び配列品質スコアリングを実行し、分析に使用した少なくとも10の読み取り深度を有するすべての品質フィルターを通過するすべてのバリエントのbcfファイルを配信した。LCRまたはX染色体のすべての部位を除去することを含めて、このファイルにさらなるQCを適用した。>5%の欠損率、HWE p値<1.0E-09、及びMAF<0.1%のバリエントも削除した。サンプルのQCを実行して、>5%の欠損率、高レベルのメンデルの法則のエラー（いくつかの場合では）、または同一の（MZ）双生児（各ペアの1つ）のサンプルを削除した。

【0311】

WES及びQC：エクソームの捕捉及び配列決定を、以下により詳細に記載するよう、Regeneron Genetics Center (RGC)において実施した。簡潔に述べると、捕捉したライブラリーを、Illumina HiSeq 2500プラットフォーム上で、v4ケミストリーで75bpのペアエンドリードを使用して配列決定した。捕捉した塩基のペアエンド配列決定を、塩基の>85%が20倍以上でカバーされるように実施したが、これは、ほとんどの標的化された塩基でヘテロ接合バリエントを判定するのに十分である。RGC DNAseq分析パイプラインで実装されているBWA-MEM及びGATKを使用して、リードアライメント及びバリエント判定を実施した。判

10

20

30

40

50

定率 < 0.90、高レベルのメンデルの法則のエラー、同一の (M Z) 双生児 (各ペアの 1つ)、または性別の不一致のサンプルは除外した。判定率 < 0.90 の S N P、及び単形性の S N P も除外した。染色体 X 及び Y、及びミトコンドリアゲノムの S N P も除外した。

### 【0312】

相関解析：空腹時血液試料を採取し、脂質分析に用いた。フリーデヴァルト式を用いて L D L を計算し、脂質低下薬の対象を対象としたいいくつかの分析では、L D L レベルを 0.7 で除算して調整した。血統に基づく血縁関係マトリックス及び / または W E S から血縁関係を推定する家族性補正を用いた家族相関を説明するために線形混合モデルを使用して遺伝的相関解析を実施した。分析はまた、年齢、年齢の二乗、性別、コホート、及び A P O B R 3 5 2 7 Q 遺伝子型について調整した。A P O B R 3 5 2 7 Q はアーミッシュにおいて濃縮されており、L D L レベルに強い影響を与える (58 mg / d l) ことが以前に特定されており (Shen et al., Arch Intern. Med., 2010, 170, 1850 - 1855)、したがって L D L 分析におけるこのバリエントの影響を考慮した。ゲノムワイドの補正された p 値 5.0 E - 0 8 を有意性の閾値として使用した。

### 【0313】

ゲノムワイド相関解析 (G W A S) を用いた染色体 9 p 領域と L D L との相関性の特定：心血管リスク因子に関連する新規遺伝子の原因バリエントを同定するために、A f f y m e t r i x 5 0 0 K 及び 6.0 チップを用いて遺伝子型決定した 1852 名のオールド・オーダー・アーミッシュの対象を使用して、ゲノムワイド相関解析を実施した。これらの参加者の基本的な特徴を表 1 に示す。

### 【0314】

#### 【表 1】

表 1：調査対象集団の基本的な特徴

	G W A S ディスカバリー	W G S 微細マッピング	W E S による確認
人数	1852	1083	4565
男性 (%)	48	50	43
年齢 (歳)	51.1 ± 16.3	50.4 ± 16.8	41.7 ± 15.2
B M I (kg/m <sup>2</sup> )	27.4 ± 5.0	26.9 ± 4.5	26.6 ± 4.9
S B P (mmHg)	121.1 ± 16.0	120.9 ± 15.6	115.1 ± 16.1
D B P (mmHg)	73.6 ± 9.4	74.4 ± 9.6	71.6 ± 9.6
コレステロール (mg/dl)	210.6 ± 46.3	211.8 ± 46.9	208.2 ± 49.2
H D L (mg/dl)	56.1 ± 14.8	55.9 ± 15.6	60.9 ± 16.4
L D L (mg/dl)	138.2 ± 42.1	140.4 ± 43.2	132.7 ± 44.9
トリグリセリド (mg/dl)	80.4 ± 53.0	77.7 ± 48.8	72.1 ± 45.6
コレステロール低下薬 (%)	2.4	3.2	1.9
糖尿病 (%)	2.6	2.4	2.2

ほとんどすべての W G S 微細マッピングサンプル (96%) が G W A S ディスカバリーサンプルに含まれていた。

W E S サンプルの 30 %のみが G W A S または W G S サンプルに含まれていた。

### 【0315】

図 1 に示すように、L D L と染色体 9 p 上の遺伝子座との間の強力な新規相関性シグナルが発見された。リード関連 S N P は r s 8 5 5 4 5 3 (p = 2.2 E - 0 8) であり、

10

20

30

40

50

出現頻度はアーミッシュで 15%、一般集団で 25% であった。マイナーな「T」対立遺伝子は、10 mg/dl 低い LDL レベルと関連していた。したがって、この GWAS SNP はアーミッシュと非アーミッシュの両方で一般的であり、効果のサイズは大きいが、大規模な GWAS メタ分析のいずれにおいても同定されなかった。これらの特徴は、以前の研究 (APOC3 及び LIPF) の特徴と一致しており、これに基づき、この GWAS SNP はこの領域における原因 / 機能的バリアントではなく、一般的な集団ではまれであるがアーミッシュ集団では一般的である、別のバリアントとの連鎖不平衡 (LD) であると結論付けられた。さらに、複数の系統の 5 つの独立した交配に基づく複数の研究により、ラット第 5 染色体上に位置するラットゲノムのシンテニー領域が、血清コレステロール及びトリグリセリドレベルの QTL を保有していることが見出された (The Rat Genome Database (RGD))。S112.26.35.44, 54 及び S128)。 10

#### 【0316】

全エクソーム配列決定 (WES) を用いた確認 :

次いで、基本的な特徴を表 1 に示す 4,565 名のアーミッシュ個体の高品質 QC の WES を使用した。LDL の混合モデル全エクソーム分析の結果、B4GALT1\_rs551564683 ミスセンスバリアントが  $3.3E-18$  の  $p$  値と 14.7 mg/dl 低い LDL の効果量との最も重要な相関性として同定された。rs551564683 バリアントのアーミッシュの MAF は 6% であったが、一般集団では極めてまれであった。バリアントは、出現頻度または集団に関する情報のない dbSNP にあり、ExAC データベース (60,000 サンプル) には存在せず、NHBLI Trans-Omics for Precision Medicine (TOPMed) データセットの 15,387 名の非アーミッシュ由来の WGS において 1 コピーのみしか見出されなかった。さらに、研究者が利用できる他の集団コホートの集合データセット (合計 125,401 名) では、このバリアントのヘテロ接合体は 79 体、ホモ接合体は 5 体しか見出されなかった (アーミッシュ集団において 1,000 倍超の濃縮を示す)。このミスセンスバリアントは、GWAS バリアントから 500 Kb 離れており、LD の  $r^2$  推定値は 0.5 である。rs551564683 と完全に相関するバリアントは存在せず、実際、次に重要な SNP は、 $p$  値が  $E-14$  の rs149557496 である。したがって、rs551564683 の相関性の強さから、第 9 染色体 GWAS 遺伝子座が本物であることだけでなく、rs551564683 が原因バリアントに期待されるすべての特徴を備えていることが確認される。 20

#### 【0317】

全ゲノム配列決定 (WGS) を使用した染色体 9 p 領域の微細マッピング :

より小さいサンプルで利用可能な WGS を使用して、エクソーム配列決定のギャップを埋め、rs551564683 が原因であるというさらなるエビデンスを提供した。1083 個の OOA の WGS データを、TOPMed プログラムの一部として生成した。WGS サンプルの基本的な特徴を表 1 に示す。WGS は、目的の領域の上位のバリアントと関する可能性のあるコーディング及び非コーディングの両方のすべての SNP 及びインデル (挿入 / 削除) を捕捉する。上位のバリアントの出現頻度は約 6% であるため、配列読み取りが不十分でバリアントの判定元がバリアントの見逃しを引き起こすことはほとんどないであろう。しかしながら、QC 手順中に除外されるバリアントが存在し得る。QC に合格しなかったバリアントを調査することにより、2 つの追加のバリアントを分析に追加した。相関解析により、B4GALT1 遺伝子のミスセンス SNP (N352S) rs551564683 が、この領域において LDL と最も有意に相関するバリアントとして同定され、 $p$  値は  $2.9E-06$ 、効果量は  $-16.4 \text{ mg/dl}$  であった (表 2 参照)。 30

#### 【0318】

## 【表2】

表2: OOAのrs551564683含有遺伝子型による平均 (n)  
LDLレベル (mg/dl)

コホート	TT	TC	CC	p値
WESによる確認 (n=4,565)	135 (n=4025)	118 (n=529)	103 (n=12)	$3.3 \times 10^{-18}$
WGS微細マッピング (n=1,083)	144 (n=952)	128 (n=130)	87 (n=1)	$2.9 \times 10^{-6}$

10

## 【0319】

TOPMed WGSデータセットは、2.9E-06 ~ 2.5E-05のp値を有するLDLに相關する20個のバリアントを提供し、上位ヒットrs551564683 ( $r^2 = 0.83 \sim 0.94$ )と完全にではないが高度に相關している(図2の赤色参照)。rs551564683を調整する条件分析は、20のバリアントの相關シグナルを完全に消失させ、この領域の他のシグナルを明らかにせず、単一の原因バリアントを強く示唆した。

## 【0320】

これらの20個のバリアント(図2の赤色参照)を注意深く調査することにより、バリアントを2つの群:影付きの三角形内の7個の赤色のバリアント及び13個の影付きではない赤色のバリアントに分割した。影付きの三角形内の7個の赤色のバリアントは、互いにほぼ完全に相關しており、 $r^2$ は0.83であり、上位ヒットはrs551564683であった。これらの7個のバリアントは、3つの理由に基づいて原因/機能バリアントとしては安全に除外された:1) それらはOOA以外では比較的一般的であり( $maf > 1\%$ )、2) TOPMed内のフレーミンハム心疾患研究(FHS)由来の3877サンプルにおいて、それらはLDLとの相關を示さず、そして3)これら7個のバリアントの1つは、4,565名のOOA対象のWESデータにおける上位ヒットrs551564683での3.3E-18に対して、6.3E-14のLDL相關性のp値を有していた。

20

## 【0321】

図2の影付き長方形内の別のバリアント群も、相關性のp値は約10E-6に過ぎず、互いに完全に相關しており、 $r^2$ は0.68であり、上位ヒットはrs551564683であった。この群もまた、OOA以外では一般的であり( $maf$ は約4%)、また、TOPMed内のFHS由来の3877サンプルにおいてLDLとの相關性を示さなかったため、原因/機能バリアントとしては除外された。

30

## 【0322】

第9染色体の31.5Mb ~ 35.5Mbの短いアーム上で4Mbにわたって延在する図2の上位ヒットrs551564683及び13個の影付きではない赤色のバリアントが残った。上記のように、これらの13個のバリアントは互いにほぼ完全に相關しており、 $r^2$ は0.91 ~ 0.94であり、上位ヒットはrs551564683であった。これらのバリアントの中で、上位ヒットのrs551564683が唯一のコーディングバリアントであり、タンパク質機能に対するバリアントの影響を予測する9つのアルゴリズムのうち5つにより、損傷性または有害性として分類された。上位ヒットのrs551564683及びこれらの13個のバリアントは、OOAにおいて6%のmafを有していたが、一般集団ではほとんど存在していなかった。

40

## 【0323】

## ハプロタイプ分析:

別個の遺伝子座間の不完全な $r^2$ は、組換え事象の結果である。主要な14-SNPハプロタイプの詳細な分析を行った。図3は、この4Mb領域の3つの主要なハプロタイプを示す。14個のSNPで同一の遺伝子型を有していたハプロタイプAの115の対象(1つのホモ接合体、及び114のヘテロ接合体)が存在し、どのSNPが原因であるかに

50

についての情報は提供されなかった。6つの対象は、rs551564683と4つの上流のSNPでヘテロ接合遺伝子型を含むハプロタイプBを有し、7つの対象は、rs551564683と9つの下流のSNPでヘテロ接合遺伝子型を含むハプロタイプCを有していた。組換ハプロタイプB及びCは、関連する対象にクラスター化され、遺伝子型判定エラーの人为的結果ではないというエビデンスを提供した。表3は、ハプロタイプB及びCの個体を単一の群に追加した後のrs551564683のp値を、ハプロタイプAの個体と比較して示す。

【0324】

【表3】

10

表3：ハプロタイプ分析結果

	A	B	C	B + C
保因者	115	7	6	13
全数	1063	1070	1069	1076
rs551564683	3.43E-05	1.40E-05	1.18E-05	4.82E-06

【0325】

ハプロタイプB及びCを個々に追加すると、p値が向上し、両方を追加するとp値がさらに向上した。p値の向上は、ハプロタイプBとCの両方が原因対立遺伝子を保有していることを示した。BとCに共通する唯一のSNPがrs551564683であり、これが原因バリエントであると考えた。

20

【0326】

グリコシル化のB4GALT1先天性障害は、rs551564683の機能的役割をサポートする：

フェノムワイド相関解析(PhewAS)を実施して、rs551564683とアーミッシュデータベースのすべての形質との相関性を試験した。LDL( $p = 3.3 \times 10^{-18}$ )及び総コレステロール( $p = 3.0 \times 10^{-18}$ )の後の最も強い相関性は、アスパラギン酸トランスアミナーゼ(AST)( $p = 3.0 \times 10^{-8}$ )で見出され、その場合、マイナー対立遺伝子のホモ接合体は、野生型のホモ接合体に対してASTレベルが2倍増加していた。高いASTは、B4GALT1においてフレームシフト挿入によって短縮型の機能不全のタンパク質が生じたことにより引き起こされる先天性グリコシル化障害(CGD)の症例で以前に報告された。さらに、フィブリノーゲンレベルとの強い相関性が観察され( $p = 5.0 \times 10^{-4}$ )、マイナーホモ接合体レベルは野生型よりも約20%低く、同じCDG患者の血液凝固不全と一致していた。さらに、小規模な実験では、13のマイナー対立遺伝子のホモ接合体において、13の野生型ホモ接合体に比べて、クレアチニナーゼ血清レベルの50%の増加( $p = 0.02$ )が認められた。B4GALT1におけるミスセンスSNPに関連し、トランケート挿入に起因する表現型のこの一貫性は、B4GALT1 rs551564683 SNPがこの領域の原因/機能遺伝子及びバリエントであるというエビデンスをさらに増強する。

30

【0327】

脂質亜分画とrs551564683との相関性を、759名のアーミッシュ個体のサンプルセットで試験し、表4に示すように、有意または有意でないp値を有するほぼすべての亜分画の低いレベルとの相関性が見出された。

40

【0328】

冠動脈石灰化スコア、大動脈石灰化スコア、及び心膜脂肪は、低いレベルと相関する傾向を示したが、有意なp値ではなかった。

PhewASはまた、rs551564683が、高いクレアチニン及び低いeGFR、ならびに高いヘマトクリット及び低い好塩基球と相関することを見出した。

50

【0329】

【表4】

表4：759名のOOA個体におけるrs551564683と脂質亜分画との相関性

形質	効果量	p値
コレステロール	-1.66E+01	3.79E-04
HDL	-4.16E+00	8.72E-03
HDL2	-1.51E+00	4.53E-02
HDL2a	-9.26E-01	9.93E-02
HDL2b	-1.94E-01	2.96E-01
HDL2c	-2.64E-01	2.14E-01
HDL3	-2.64E+00	3.98E-03
HDL3a	-1.51E+00	2.00E-02
HDL3b	-1.68E-01	4.16E-01
HDL3c	-5.93E-01	1.47E-02
HDL3d	-4.44E-01	2.48E-02
IDL	-7.31E-01	4.92E-01
IDL1	-1.19E-02	9.73E-01
IDL2	-7.65E-01	3.37E-01
LDL	-1.23E+01	2.37E-03
LDL1	-2.22E+00	7.20E-02
LDL2	-5.64E+00	3.99E-02
LDL3	-3.81E+00	1.32E-01
LDL4	-3.96E-02	9.65E-01
LDLReal	-1.12E+01	9.53E-04
Lpa	-2.15E-01	6.34E-01
Lpa1	-2.91E-01	3.00E-01
Lpa2	4.67E-02	8.27E-01
Lpa3	2.31E-01	5.04E-01
Lpa4	-2.91E-02	9.19E-01
Lpa5	-2.48E-01	3.11E-01
レムナントリポタンパク質	-7.23E-01	5.97E-01
TC/HDL比	-3.29E-02	7.68E-01
総非HDL	-1.24E+01	3.97E-03
総VLDL	-1.03E-01	8.70E-01
トリグリセリド	2.19E+00	6.46E-01
VLDL1+2	-4.10E-02	8.86E-01
VLDL3	6.15E-03	9.86E-01
VLDL3a	2.28E-02	8.97E-01
VLDL3b	-6.57E-02	7.30E-01

10

20

30

40

【0330】

実施例2：試料調製及び配列決定

アーミッシュの対象からゲノムDNAの試料濃縮物を取得し、その後、社内施設に移し、配列分析まで - 80 (Liconic Tube Store) で保管した。蛍光 (Life Technologies) によって試料の量を測定し、100ngの試料を2%プレキャストアガロースゲル (Life Technologies) で電気泳動することによって品質を評価した。

【0331】

50

D N A 試料を正規化し、それぞれの試料を、集束音響エネルギー ( C o v a r i s L E 2 2 0 ) を使用して 1 5 0 塩基対の平均断片長にせん断した。せん断したゲノム D N A を、 K a p a B i o s y s t e m s のカスタム試薬キットを用い、社内開発した完全自動化アプローチを使用して、エクソーム捕捉用に調製した。ライブラリーの調製中に、固有の 6 塩基対のバーコードを各 D N A 断片に追加し、多重化エクソーム捕捉及び配列決定を容易にできるようにした。いくつかの修正を加えた、 I D T から入手可能な x G e n d e s i g n 上でのエクソーム捕捉の前に等量の試料をプールした。 I l l u m i n a v 4 H i S e q 2 5 0 0 上で 7 5 b p のペアエンド配列決定法を使用して、多重化試料を配列決定した。

## 【 0 3 3 2 】

I l l u m i n a H i S e q 2 5 0 0 プラットフォーム上で生成した生配列データを、 D N A n e x u s ( D N A n e x u s I n c . , M o u n t a i n V i e w , C A ) の高性能コンピューティングリソースにアップロードし、自動化ワークフローによって生の . b c l ファイルから注釈付きバリエント判定結果へと処理した。試料特異的バーコードに基づいた分析のために、生の読み取り結果を、 C A S A V A ソフトウェア ( I l l u m i n a I n c . , S a n D i e g o , C A ) を使用して適切な試料に割り当てた。

## 【 0 3 3 3 】

次いで、試料特異的読み取り結果を、 B W A - m e m ( L i a n d D u r b i n , B i o i n f o r m a t i c s , 2 0 0 9 , 2 5 , 1 7 5 4 - 1 7 6 0 ) を使用して参照配列にアライメントした。これにより、特定の試料の読み取り結果すべてと、各読み取り結果をマッピングしたゲノム座標を含む各試料のバイナリアライメントファイル ( B A M ) を生成した。アライメント完了後、試料の読み取り結果を評価し、 P i c a r d M a r k D u p l i c a t e s ツール ( p i c a r d . s o u r c e f o r g e . n e t ) を使用して重複した読み取り結果を識別してフラグを立て、重複した各読み取り結果がマークされたアライメントファイル ( d u p l i c a t e s M a r k e d . B A M ) を作成した。

## 【 0 3 3 4 】

次いで、ゲノム解析ツールキット ( G A T K ) ( V a n d e r A u w e r a , C u r . P r o t o c o l s i n B i o i n f o r m a t i c s , 2 0 1 3 , 1 1 , 1 1 - 3 3 , M c K e n n a , G e n o m e R e s . , 2 0 1 0 , 2 0 , 1 2 9 7 - 1 3 0 3 ) を使用して、アライメントして重複をマークした各試料の読み取り結果のローカルアライメントを実施した。次いで、 G A T K H a p l o t y p e C a l l e r を使用して、アライメントして重複をマークした読み取り結果を処理し、単一ヌクレオチド変異及び I N D E L を含めて、試料がゲノム参照と異なっているすべてのエクソン位置、ならびに特定の試料が参照と異なっている任意の位置での試料中のバリエントの接合性を同定した。

## 【 0 3 3 5 】

参照及び代替対立遺伝子の両方に割り当てられた読み取りカウント、遺伝子型判定の信頼性を表す遺伝子型品質、及びその位置でのバリエント判定の全体的な品質を含む、関連する測定基準を、すべてのバリエント部位で出力した。次いで、 G A T K の V a r i a n t Q u a l i t y S c o r e R e c a l i b r a t i o n ( V Q S R ) を使用して、トレーニングデータセットを用いて試料のバリエントの全体的な品質スコアを評価し、このスコアを評価し、再計算して特異性を高めた。各試料のメトリック統計を捕捉して、捕捉性能、アライメント性能、及びバリエント判定を評価した。コホートの配列決定の完了後、 G A T K を使用したジョイントジェノタイピングによりプロジェクトレベルの V C F を生成し、コホート内の試料が参照ゲノムのバリエントを保持している部位におけるすべての試料の遺伝子型及び関連するメトリック情報を生成した。下流の統計解析に使用したのは、このプロジェクトレベルの V C F であった。 V Q S R に加えて、 G A T K を使用した Q u a l i t y B y D e p t h ( Q D ) メトリックを用いてバリエントに注釈を加え、 Q D > 2 . 0 、欠損率 < 1 % 、及び H a r d y - W e i n b e r g 平衡 p 値 > 1 . 0 × 1 0 - 6 を有する両対立遺伝子のバリエントを、さらなる分析のために保持した。

## 【 0 3 3 6 】

10

20

30

40

50

下流の配列データ解析の前に、遺伝的に判定した性別と報告された性別が不一致の試料、高いヘテロ接合率、低い配列カバレッジ（標的化された塩基の75%未満の20×カバレッジとして定義される）、またはあいまいな相関性の程度が異常に高い試料、及び遺伝的に同定された試料の重複は除外した。

【0337】

ANNOVAR (Wang et al., Nuc. Acids Res., 2010, 38, e164) 及び注釈及び解析のための他のカスタマイズされたアルゴリズムを使用する注釈パイプラインを使用して配列バリアントに注釈を加えた。バリアントを、それらの潜在的な機能的効果に従って分類し、その後、一般的に利用可能な集団対照データベース、及び一般的な多型と高頻度の、おそらく良性のバリアントを除外するためのデータベースにおいて観察された頻度でフィルタリングした。複数種のアライメントに基づく保存スコアとともに、バリアントの機能的効果の生物情報学的予測のためのアルゴリズムを、バリアントの注釈プロセスの一部として組み込み、同定された候補バリアントの潜在的な有害性について通知するために使用した。

10

【0338】

実施例3：B4GALT1 rs551564683 N352S 出現頻度はアーミッシュで濃縮されている

約4700名のアーミッシュの対象におけるエクソーム配列決定及び相関性解析により、第9染色体上のrs551564683は総コレステロール値と高い相関性があることが見出された（p = 1.3E-10）（図4参照）。RS551564683は、B4GALT1タンパク質の352位でセリンがアスパラギンに変化するミスセンスバリアントをコードしている。この領域における次の最も高いLDL関連バリアントはrs149557496であり、p値はわずか10<sup>-5</sup>であったが、このことはN352Sバリアントが最も可能性の高い原因バリアントであることを示唆している。特に図4を参照すると、エクソーム配列データにおいて、Asn352Ser B4GALT1とともに最も高いLDのバリアントは、HRCCT1のrs149557496であり、2.8Mbの距離、R<sup>2</sup>は0.78、アーミッシュにおけるLDLのP値が10<sup>-5</sup>であった。アーミッシュの全ゲノム配列データ（TOPMED）からは、この領域においてLDL-Cとより相関性の高いバリアントを同定することができなかった。

20

【0339】

30

さらなる解析により、B4GALT1 N352Sバリアントの出現頻度はアーミッシュ集団において1000倍超濃い頻度であることが明らかになった（図5参照）。データからは、4725のアーミッシュのコホートにおいて、rs551564683を含む対立遺伝子の548のヘテロ接合体保因者が同定され、13の保因者が対立遺伝子に関してホモ接合体であることが示された（図5参照）。これに対して、研究者が利用できる他の集団コホートの合計125,401名の集合データセットを解析し、この集合データセットでは79のヘテロ接合体と5つのホモ接合体のみが同定された。アーミッシュコホートの対立遺伝子の出現頻度は、集合データセットの約0.0025に対して、約0.06と推定された（図5参照）。アーミッシュにおけるこの対立遺伝子のより高い出現頻度は、遺伝的浮動が原因であり得ると考えられる。

40

【0340】

実施例4：B4GALT1 N352Sは、血清脂質の減少及びASTの増加と相関するB4GALT1 N352S変異と、血清脂質、冠動脈疾患（CAD）、及び肝形質を含む様々な表現型との相関性を評価した。参照対立遺伝子についてホモ接合性、代替対立遺伝子についてヘテロ接合性、及び代替対立遺伝子についてホモ接合性の個体を含むアーミッシュのコホートに基づいて、関連付けを行った。脂質及び肝臓の形質に対する遺伝子型の平均ならびにCADのリスクを測定し、対象の年齢及び年齢の二乗、対象の性別、及び研究（一定年数の期間にわたるいくつかの研究から表現型データを収集したため）の影響を除去することにより、効果の測定値を調整した。心臓脂肪の場合、遺伝子型の平均をBMIに対してさらに調整した。測定した表現型に対する変異の効果量を、95%の信頼

50

区間で測定した。形質及び結果を図 6、図 7、及び図 8 に示す。

【 0 3 4 1 】

図 6 に示すように、N 3 5 2 S 変異の存在は、一般的に、血清脂質、特に、強力な統計的有意性を達成した総コレステロール ( $p$  値  $1.3 \times 10^{-10}$ ) 及び LDL ( $p$  値  $1.8 \times 10^{-9}$ ) レベルの低下と相關していた。この変異に関してヘテロ接合性及びホモ接合性である個体は、LDL レベルについてそれぞれ  $17.3 \text{ mg/dL}$  及び  $31.2 \text{ mg/dL}$  の低下を示した。このバリエントには冠動脈石灰化の減少傾向が認められた。さらに、この変異の存在は、アスパラギン酸アミノトランスフェラーゼ (AST) レベルの増加と相關していた ( $p$  値  $6.0 \times 10^{-8}$ )。AST レベルの劣性モデルの  $p$  値は、 $9 \times 10^{-23}$  であると測定された。この変異は、アラニンアミノトランスフェラーゼ (ALT) レベル、アルカリホスファターゼレベル、または肝臓脂肪レベルの増加と相關しているように思われなかった。コレステロール、LDL、及び AST のレベルを図 7 にグラフで示す。図 7 では、参照対立遺伝子についてホモ接合 (TT)、代替対立遺伝子についてヘテロ接合 (CT)、及び代替対立遺伝子についてホモ接合 (CC) である対象について、コレステロール、LDL、及び AST のレベルを示す。表示の値は未調整である。対象の年齢及び年齢の二乗、性別、及び研究の調整に基づいて、値を再計算した (図 7 の下部に集計)。

【 0 3 4 2 】

脂質亜分画に対する N 3 5 2 S 変異の影響も評価した。これらの結果を図 8 に示す。参考対立遺伝子についてホモ接合性、代替対立遺伝子についてヘテロ接合性、及び代替対立遺伝子についてホモ接合性の個体を含むアーミッシュコホートに基づいて、関連付けを行った。図 8 の結果は、B 4 G A L T 1 N 3 5 2 S 変異が、試験したすべての脂質亜分画の減少と相關していることを示す。

【 0 3 4 3 】

実施例 5 : B 4 G A L T 1 N 3 5 2 S はフィブリノーゲンレベルの低下と相關する試料のサブセットにおいて B 4 G A L T 1 N 3 5 2 S 変異とフィブリノーゲンレベルとの相関性も評価した。実施例 4 で評価した血清脂質、CAD、及び肝臓形質について、フィブリノーゲンレベルとの関連付けを、代替対立遺伝子についてホモ接合性、参考対立遺伝子についてヘテロ接合性、及び代替対立遺伝子についてホモ接合性の個体を含むアーミッシュコホートに基づいて行った。フィブリノーゲンレベルの遺伝子型の平均を、クロピドグレルのレジメン下にない個体 (薬物未投与) 及びクロピドグレルのレジメン下の個体 (クロピドグレル投与) の 2 つの部分群で測定し、分析の一環として、各群の平均レベルを、対象の年齢及び年齢の二乗、対象の性別、研究の影響を除去することにより調整した。フィブリノーゲンレベルに対する変異の効果量を、95% の信頼区間で測定した。図 9 に示すように、N 3 5 2 S 変異の存在は、薬物未投与 ( $p$  値  $1.15 \times 10^{-3}$ ) 群及びクロピドグレル投与 ( $p$  値  $2.74 \times 10^{-5}$ ) 群それぞれのフィブリノーゲンレベルの低下と相關していた。薬剤未投与の部分群は、フィブリノーゲンのおよそ  $24 \text{ mg/dL}$  の低下を示した (図 9 参照)。クロピドグレル投与の部分群は、フィブリノーゲンのおよそ  $32.5 \text{ mg/dL}$  の低下を示した (図 9 参照)。

【 0 3 4 4 】

実施例 6 : さらなる B 4 G A L T 1 N 3 5 2 S 相関性

アーミッシュのコホート内で、B 4 G A L T 1 N 3 5 2 S 変異と、クレアチニンレベル、推算糸球体濾過量 (eGFR)、好塩基球レベル、及びヘマトクリット割合を含む他の形質との相関性の評価も行った。図 9 に示すように、バリエントは、クレアチニンレベルのわずかな増加と弱く相關していたが、eGFR、好塩基球レベル、またはヘマトクリット割合とは有意に相關していなかった。

【 0 3 4 5 】

実施例 7 : ゼブラフィッシュにおける b 4 g a l t 1 オルソログノックダウン

細胞ベースのアッセイのエビデンスと並行して、B 4 G A L T 1 p. A s n 3 5 2 S e r の LDL への影響を調べるためにゼブラフィッシュモデルを探求した。

10

20

30

40

50

## 【0346】

## ゼブラフィッシュ飼育、モルホリノ注射及び検証

野生型 (Tubingen) ゼブラフィッシュストックを使用して、モルホリノ注射用の胚を生成した。成魚は 27 ~ 29 で維持及び繁殖させ、胚は 28.5 で発育させた。すべての動物は、メリーランド大学実験動物委員会によって承認されたプロトコールに従って飼育及び維持した。モルホリノアンチセンスオリゴヌクレオチド (MO) は、b4gal t1 を標的とした以前に公開された MO (Machingo et al., Dev. Biol., 2006, 297, 471 - 482) に基づいて取得した (Gene Tools, Inc.)。1 ~ 2 細胞期に MO を注射し、野生型 b4gal t1 転写産物の qRT-PCR 定量により検証した。オフターゲット毒性を、p53 の 113 アイソフォームの qRT-PCR 定量化によって評価した (Robu et al., PLoS Genet., 2007, 3, e78)。mRNA レスキュー実験のために、ヒト B4GALT1 mRNA を、その遺伝子の野生型または N352S バリアントのオーブンリーディングフレーム (ORF) を有する pCS2+ プラスミドベクターから転写させた。mRNA を様々な濃度で MO と混合し、1 ~ 2 細胞期の胚に共注射した。各注射実験では、合計 200 ~ 400 個の胚に注射し、各実験を最低 3 回繰り返した。

10

## 【0347】

## ゼブラフィッシュにおける LDL 定量化

受精後 (dpf) 5 日の幼虫 100 匹を、実験ごとに、氷冷した 10 μM プチル化ヒドロキシトルエン 400 μl で均質化した。ホモジネートを、脂質抽出の準備として、0.45 μm Dura PVDF メンブレンフィルター (Millipore) で濾過した。HDL 及び LDL / VLDL コレステロールアッセイキット (Cell Biolabs, Inc.) を使用して、製造業者のプロトコールに従ってホモジネートを処理した。沈殿及び希釈後、SpectraMax Gemini EM プレートリーダー及び Soft Max Pro マイクロプレートデータ取得及び分析ソフトウェア (Molecular Devices) を使用した蛍光定量的分析により試料を分析した。

20

## 【0348】

ゼブラフィッシュオルソログ (b4gal t1) のゲノムノックアウトを、CRISP R / Cas9 を介したエクソン 2 のターゲティングを使用して生成した。ノックアウト動物における胚の致死性のマウス報告と一致して、注射した F0 動物は成体まで生存することができず、幼若期で一貫して死亡した。生存率の欠如を回避するために、以前に報告された、胚内に注射したスプライスプロッキングアンチセンスモルホリノオリゴヌクレオチド (MO) を使用したノックダウンアプローチ (Machingo et al., Dev. Biol., 2006, 297, 471 - 482) を採用した。MO の有効性を、2 つの異なる濃度で qRT-PCR によって検証し (図 10 参照)、オフターゲット毒性の可能性を排除した (図 11 参照)。LDL レベルの変化を定量化するために、8 ng の MO を注射し、注射した胚を受精後 (dpf) 5 日間まで培養し、その段階で以前に公開されたプロトコール (O'Hare et al., J. Lipid Res., 2014, 55, 2242 - 2253) に従って幼虫の総 LDL を分析した。対照幼虫に比べて、MO を注射した幼虫において LDL の有意な低下が観察され、このことは LDL 恒常性における b4gal t1 の役割と一致する (図 12 参照)。この結果は、エクソン 2 を標的とする第 2 のスプライスプロッキング MO を使用して確認され、2 ng の MO を注射すると LDL 濃度が低下した (データは示さず)。これらの観察の特異性を検証し、ゼブラフィッシュにおけるヒト B4GALT1 の機能を試験するために、ヒト遺伝子のオーブンリーディングフレーム (ORF) を保有する pCS2+ プラスミドからの in vitro 転写により、ヒト遺伝子をコードする完全長キャップ mRNA を生成させた。ノックダウン表現型をレスキューする野生型ヒト mRNA の能力を評価するために、b4gal t1 MO とともに胚に共注射し、未給餌幼虫の LDL を評価した。3 種類の濃度の mRNA (10 pg、25 pg、及び 50 pg) を 8 ng の MO と共に注射した。50 pg の B4GALT1 mRNA の共注射により、対照 MO のみを注射した幼虫と統計的に見分けがつかない

30

40

50

L D L レベル ( p 値 = 0 . 1 4 ) が得られ、このことは、ヒト m R N A がゼブラフィッシュ遺伝子のノックダウンの効果をレスキューすることができたことを示唆している ( 図 1 2 参照 ; 幼虫を b 4 g a l t 1 に対する M O で処置し、 M O を W T ヒト B 4 G A L T 1 m R N A と共に注射するか ( W T レスキュー ) 、または M O を、 A s n 3 5 2 S e r 変異をコードする B 4 G A L T 1 m R N A と共に注射した ( N 3 5 2 S レスキュー ) ) 。

#### 【 0 3 4 9 】

これらのデータは、ヒト B 4 G A L T 1 のバリアントの機能的解釈のためのこの系の使用をサポートするものであり、ヒト野生型 B 4 G A L T 1 m R N A が全身性 L D L レベルの調節に関してゼブラフィッシュにおいて機能的であることを示唆している。 B 4 G A L T 1 機能に対する p . A s n 3 5 2 S e r の影響をさらに調べた。部位特異的変異誘発を使用して ( O ' H a r e e t a l . , H e p a t o l o g y , 2 0 1 7 , 6 5 , 1 5 2 6 - 1 5 4 2 ) 、ヒト B 4 G A L T 1 O R F 構築物のコーディング配列に T から C への変異を導入し、完全長 m R N A を生成した。 B 4 G A L T 1 p . 3 5 2 S e r m R N A と M O の共注射により、 L D L 表現型のレスキュー能力が低下した。結果として得られた L D L 濃度は、野生型 m R N A と M O の共注射から得られたものより 1 5 % 低く、統計的に有意な効果であった ( 4 6 . 6  $\mu$  M に対して 3 9 . 9  $\mu$  M 、 p 値 = 0 . 0 2 ) 。しかしながら、このレベルの L D L はまた、 b 4 g a l t 1 M O 単独よりも統計的に大きく ( p 値 = 0 . 0 1 ) ( 図 1 2 参照 ) 、このことはミスセンスバリアントによって導入される機能の部分的な欠陥を示唆している。

#### 【 0 3 5 0 】

##### 実施例 8 : 標的化された遺伝子型判定

Q u a n t S t u d i o システム ( T h e r m o F i s h e r S c i e n t i f i c ) を使用した標的化された S N P 遺伝子型判定を、 3 , 2 3 6 名の O O A 対象に対して実施した。 1 4 個の S N P の L D 構造に基づいて、 7 つの S N P を遺伝子型判定用に選択し、 r s 5 5 1 5 6 4 6 8 3 の相関性のエビデンスは 4 . 1 E - 1 3 であり、一方、他の S N P については約 E - 1 0 であったことから ( 図 1 4 ) 、 r s 5 5 1 5 6 4 6 8 3 がこの領域における原因バリアントであることが確認された。

#### 【 0 3 5 1 】

実施例 9 : B 4 G A L T 1 N 3 5 2 S は、タンパク質安定性または細胞局在性を変化させずに酵素活性の低下を引き起こす

B 4 G A L T 1 の特性の調査を、ヒトエピトープタグ標識 F 1 a g - B 4 G A L T 1 3 5 2 A s n またはエピトープタグ標識 F 1 a g - B 4 G A L T 1 3 5 2 S e r を過剰発現している C O S - 7 及び H u h 7 細胞で行った ( 図 1 5 及び 1 6 ) 。図 1 5 に示すように、 B 4 G A L T 1 または F 1 a g 抗体を使用した F 1 a g - 3 5 2 A s n または F 1 a g - 3 5 2 S e r の共焦点顕微鏡画像は、同一の染色パターンを示す ( スケールバー = 1 0  $\mu$  m ) 。図 1 6 に示すように、 H u h 7 細胞の間接免疫蛍光による細胞内局在性は、内在的に発現する B 4 G A L T 1 とゴルジ体マーカーである T G N 5 6 の共局在を示した。ヒトエピトープタグ標識 F 1 a g - B 4 G A L T 1 3 5 2 A s n またはエピトープタグ標識 F 1 a g - B 4 G A L T 1 3 5 2 S e r を過剰発現させた場合でも、同様の共局在パターンが観察された ( 図 1 6 ) 。図 1 6 に示すように、ヒト肝癌 H u h 7 細胞において過剰発現させた内在性 B 4 G A L T 1 、 F 1 a g - 3 5 2 A s n 、及び F 1 a g - 3 5 2 s e r は、トランスゴルジネットワークマーカー T G N 4 6 と共に局在していた。トランスゴルジネットワークマーカー T G N 4 6 と関連する内在性 B 4 G A L T 1 、 F 1 a g - 3 5 2 A s n 、及び F 1 a g - 3 5 2 S e r の細胞内局在の共焦点顕微鏡画像をスケールバー = 1 0  $\mu$  m で示す。

#### 【 0 3 5 2 】

C O S - 7 細胞は、内在性 B 4 G A L T 1 の含有量が低いことが観察されたため ( 図 1 7 、パネル B ) 、この細胞株を使用して、タンパク質安定性及び / または定常状態レベルに対するミスセンス変異の影響、及びガラクトシルトランスフェラーゼ活性を評価した。結果からは、ミスセンス変異がタンパク質の安定性及び / または定常状態レベルに影響を

10

20

30

40

50

及ぼさないことが示された（ウエスタンプロットによる）（図17）。図17に、タンパク質の安定性及び／または定常状態レベルに対する352Serの影響を示す。パネルAは、COS7細胞内で発現させた352Asnまたは352Ser Flagタグタンパク質融合体のいずれかを遊離EGFPとともに発現しているCOS7細胞を示す。細胞溶解物を、市販の抗体を使用して、B4GALT1、バクチン、及びEGFPのウエスタンプロットによって分析した。4つの類似した実験のうちの1つを示す。パネルBは、RT-qPCR分析によって測定したB4GALT1遺伝子のmRNA発現レベルを示す。データは4つの実験の平均±標準誤差を表す。

#### 【0353】

352Serの触媒活性を測定するために、非形質移入COS-7細胞、及び発現ベクター単独を形質移入したか、野生型またはバリエント型B4GALT1のcDNAインサートを保有するCOS-7細胞の溶解物を、ガラクトシルトランスフェラーゼ活性について分析した。FLAGタグ標識タンパク質の発現に対して正規化した場合（図18、パネルA及びBのイムノプロッティング実験）、352Serの酵素活性は352Asnに比べておよそ50%低下した（図18、パネルC）。図18に、活性に対する352Ser変異の影響を示す。パネルA及びBは、COS7細胞内で発現させた352Asnまたは352Ser Flagタグタンパク質融合体のいずれかを発現しているCOS7細胞を示す。細胞溶解物をウサギ抗Flag IgGまたはウサギ前免疫対照IgGの存在下でインキュベートした。免疫沈降物を、市販の抗体を使用してB4GALT1またはFlagに対するウエスタンプロットにより分析した。4つの類似した実験のうちの1つを示す。パネルCは、市販キット（R&D）で測定した免疫沈降物のB4GALT1活性を示す。各データポイントは、免疫沈降物において回収された352Asnまたは352Serタンパク質の量とのB4GALT1特異的活性の計算した比率の平均を表す。ImageJソフトウェアを使用したデンシトメトリーにより、ウエスタンプロットECLのシグナルを定量化した。データは、4つの実験の平均±標準誤差を表す（\*、p<0.05、352Asn対352Ser）。

#### 【0354】

これらの実験は、このミスセンス変異がタンパク質発現のレベル及びその局在化に影響を及ぼさないが、酵素活性を低下させることを示している。

実施例10：先天性グリコシル化障害（CDG）試験のための炭水化物欠損トランスフェリン

3つの遺伝子型群（8つのマイナーなホモ接合体、8つのヘテロ接合体、8つのメジャーなホモ接合体）の24名の対象由来の血清試料0.1mlを使用してCDG試験を実施した。マイナーなホモ接合体のそれぞれを、親類または親縁係数に基づく近縁同性個体であるヘテロ接合体及びメジャーなホモ接合体と一致させた。年齢、及び保因者の状態もまた、APOB<sup>R3527Q</sup>におけるメジャーな脂質改変遺伝子の対立遺伝子と一致させた。

#### 【0355】

水で希釈した試料を、免疫親和性カラムを使用して二重洗浄した。溶出タンパク質のグリコシル化プロファイリングを、APOCIII及びトランスフェリンに固有の2つのスキヤン範囲で動作する質量分析計を使用して実施した。各タンパク質のグリコフォーム比を使用して、グリコシル化欠損を判定した。CDG試験は、Mayo病院のMayo医学研究所で実施した。

#### 【0356】

結果からは、24個の試料すべてが、正常レベルのモノオリゴ糖/ジオリゴ糖トランスフェリン比、a-オリゴ糖/ジオリゴ糖トランスフェリン比、APOCI<sup>III-1/APOCI<sup>III-2</sup>比、及びAPOCI<sup>III-0/APOCI<sup>III-2</sup>比を有していたことが示された。しかしながら、すべての野生型試料が正常レベルのトリシアロ/ジオリゴ糖トランスフェリン比を有していたが、すべてのヘテロ接合体のレベルは中間範囲にあり、すべてのマイナーなホモ接合体のレベルは異常であり、一致する野生型及びヘテロ接合体よりも有意に高かった（p=7.6E-10）（図19）。これらの結果は、このミスセ</sup></sup>

10

20

30

40

50

ンス変異が、B 4 G A L T 1 の酵素活性の低下の結果として、グリコシル化の欠損と相關していることを示している。

【0357】

実施例 11：血漿糖タンパク質の網羅的 N 結合グリカン分析

デシアリル化及び低ガラクトシル化がトランスクルタリノンのみに影響を及ぼしているのか、他の糖タンパク質に及ぶのかを判定するために、Regeneron の分析化学グループが網羅的 N - グリカン分析を実施した。レクチン濃縮糖タンパク質を、5 組のメジャー及びマイナーホモ接合体の血清から 2 連で抽出し、親水性相互作用クロマトグラフィーを使用して標識グリカンの網羅的 N 結合グリカン分離を実施し、蛍光によって検出し、質量分析 (HILIC - F L R - M S) によって分析した (図 20 及び表 5)。図 20 に示すように、B 4 G A L T 1 N 352S のマイナー (SS) 及びメジャー (NN) ホモ接合体の一致ペア由来の糖タンパク質の N - グリカン分析の代表的な HILIC - F L R - M S スペクトルを示す。結果からは、マイナーなホモ接合体が、1 つのガラクトースと 1 つのシアル酸だけを含む二分岐グリカン ( $p = 3.1 \text{ E-5}$ )、1 つのガラクトースを含む非シアル酸付加二分岐グリカン ( $p = 0.001$ )、及びガラクトースとシアル酸の両方が欠落したトランケート型二分岐グリカン ( $p = 0.005$ ) を含む低ガラクトシル化及び低シアル酸付加グリカンのレベルが有意に高いことが示された。他方、マイナーなホモ接合体は、2 つのガラクトースと 2 つのシアル酸を有する二分岐グリカンのレベルが有意に低い ( $p = 0.001$ ) (表 5)。マイナーなホモ接合体では、全体的なガラクトシル化 ( $p = 9.2 \text{ E-5}$ ) 及びシアル酸付加 ( $p = 0.001$ ) が有意に低かったが、一方、フコシル化レベルには差異がなかった ( $p = 0.5$ )。血清の CDT 及び網羅的 N - グリカン分析の両方が、マイナーなホモ接合体の炭水化物欠損糖タンパク質のレベルの有意な増加を示したが、このことは B 4 G A L T 1 N 352S がタンパク質グリコシル化の欠損をもたらすことを示している。

【0358】

【表 5】

表 5：マイナーなホモ接合体とメジャーなホモ接合体の間で有意に異なるグリカンのピーク面積割合の平均 (± s.d.)

グリカン	メジャーなホモ接合体	マイナーなホモ接合体	P値
G0F	0.58 ± 0.34	1.84 ± 0.48	0.005
G1	0.19 ± 0.12	0.91 ± 0.16	0.001
G1S1	0.63 ± 0.16	4.7 ± 0.38	3.1E-5
G2S2	39.3 ± 0.79	31.5 ± 1.8	0.001

【0359】

本開示は、上記で記載及び例示した実施形態に限定されるものではなく、添付の特許請求の範囲内で変更及び修正が可能である。本開示はまた、本明細書に挙げる任意のヘッダーの使用によるいかなる様式にも限定されるべきではない。

10

20

30

40

50

【図面】

【図 1】

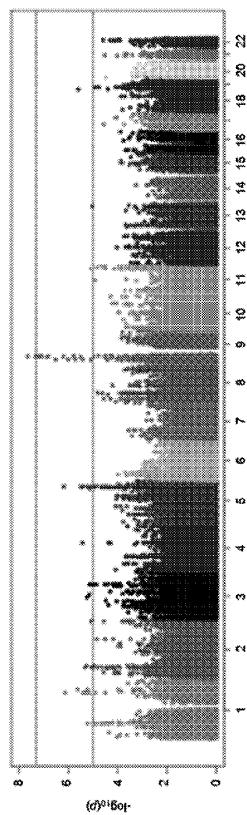
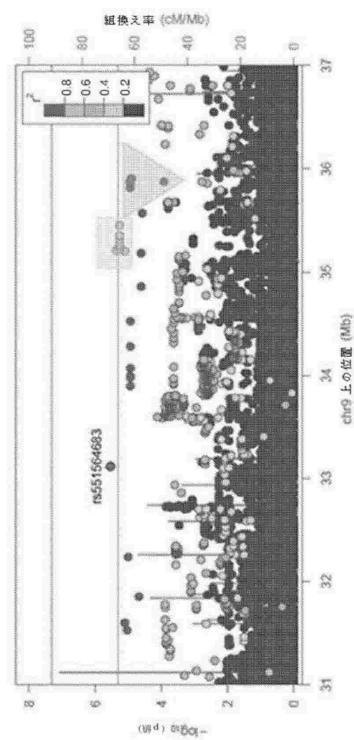


Figure 1

【図 2】



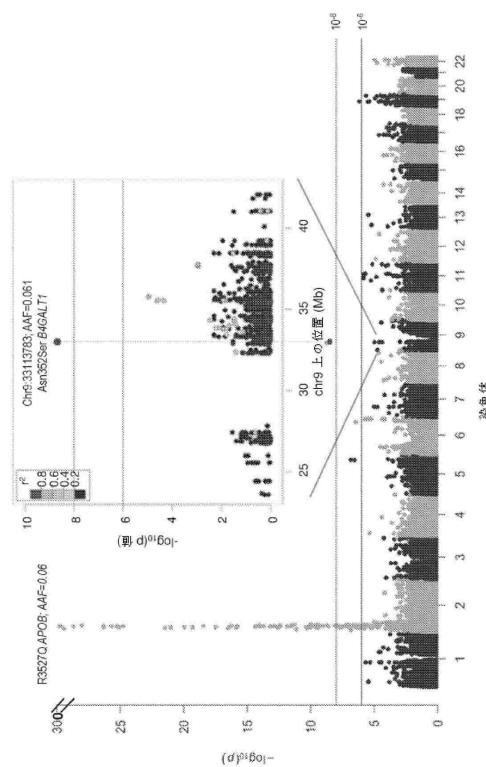
10

20

【図 3】

ノックアウト		コピー数	9315223876	9315994651	931853653	932231957	9551564683	933992787	933971380
A	115	1	1	1	1	1	1	1	1
B	7	1	1	1	1	1	1	0	0
C	6	0	0	0	0	0	1	1	1
ノックイン		コピー数	933991394	934066623	934277757	934520551	934830994	935178756	93570095
A	115	1	1	1	1	1	1	1	1
B	7	0	0	0	0	0	0	0	0
C	6	1	1	1	1	1	1	1	1

【図 4】

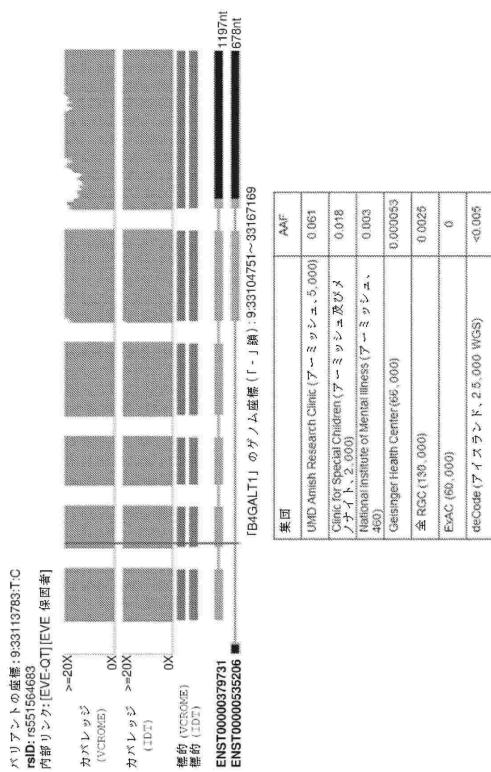


30

40

50

【図 5 - 1】



【図 5 - 2】

集団	集団サイズ	△テロ 保因者数	△モ 保因者数	AAF
BCH	580	0	0	0.00000
CNCD	96	0	0	0.00000
COL-CHUNG	4064	0	0	0.00000
COL-CHUNG_FBC	1193	0	0	0.00000
CSC	2084	65	5	0.017994
DHS	4654	0	0	0.00000
DRIFT-NIMH	460	3	0	0.002081
DUKE	8783	2	0	0.000114
ECO	858	0	0	0.00000
GHS	66216	7	0	0.000053
IOWA	58	0	0	0.00000
MADGC	2393	1	0	0.000209
MARIO	127	0	0	0.00000
NIAID	95	0	0	0.00000
SIMONS	96	0	0	0.00000
SINAI	96	0	0	0.00000
TAICHI	13963	0	0	0.00000
TSK	3407	0	0	0.00000
UCHC	1039	0	0	0.00000
UMD	4725	548	13	0.060741
UPENN	11451	1	0	0.000044
UTAH	3688	0	0	0.00000
合計	130126	627	18	0.002548

10

20

図 5 (続き)

【図 6 - 1】

表現型	遺伝子型平均*	遺伝子型平均 (未調整)	カウント数 (参照へテロ/代替)	効果 (95% CI)	P 値
CAD 関連影響					
冠動脈疾患 (log10)	0.30 / 0.201 / NA	0.298 / 0.20 / NA	204/42/0	-22.4 [-38.3, -9.52] %	0.0438
冠動脈石炭化 (log10)***	1.87 / 1.44 / 0	71.5 / 71.4 / NA	546/54/1	-19 %	0.1349
心臓肥厚 **	69.5 / 82.7 / NA		148/27/0	12.78 [1.05, 24.46] %	0.0339
肝形質					
AST (log10) (μU/L)	18.0 / 18.6 / 35.2	18.0 / 18.7 / 34.3	3910/3771/2	7.15 [4.72, 12.39] %	6.0E-05
ALT (log10) (μU/L)	17.2 / 16.9 / 26.3	17.2 / 16.9 / 19.7	3978/359/13	-0.23 [-4.50, 4.71] %	0.8920
アルカリホスファターゼ (log10) (μU/L)	53.8 / 52.9 / 62.9	53.8 / 53.1 / 61.8	3878/531/13	0.92 [-3.87, 2.69] %	0.5931
肝臓脂肪 (EBCT)に基づく	1.27 / 1.33 / NA	1.27 / 1.32 / NA	3665/36/0	0.05 [-0.005, 0.110] %	

【図 6 - 2】

表現型	遺伝子型平均*	遺伝子型平均 (未調整)	カウント数 (参照へテロ/代替)	効果 (95% CI)	P 値
CAD 関連影響					
冠動脈疾患 (log10)	0.30 / 0.201 / NA	0.298 / 0.20 / NA	204/42/0	-22.4 [-38.3, -9.52] %	0.0438
冠動脈石炭化 (log10)***	1.87 / 1.44 / 0	71.5 / 71.4 / NA	546/54/1	-19 %	0.1349
心臓肥厚 **	69.5 / 82.7 / NA		148/27/0	12.78 [1.05, 24.46] %	0.0339
肝形質					
AST (log10) (μU/L)	18.0 / 18.6 / 35.2	18.0 / 18.7 / 34.3	3910/3771/2	7.15 [4.72, 12.39] %	6.0E-05
ALT (log10) (μU/L)	17.2 / 16.9 / 26.3	17.2 / 16.9 / 19.7	3978/359/13	-0.23 [-4.50, 4.71] %	0.8920
アルカリホスファターゼ (log10) (μU/L)	53.8 / 52.9 / 62.9	53.8 / 53.1 / 61.8	3878/531/13	0.92 [-3.87, 2.69] %	0.5931
肝臓脂肪 (EBCT)に基づく	1.27 / 1.33 / NA	1.27 / 1.32 / NA	3665/36/0	0.05 [-0.005, 0.110] %	

30

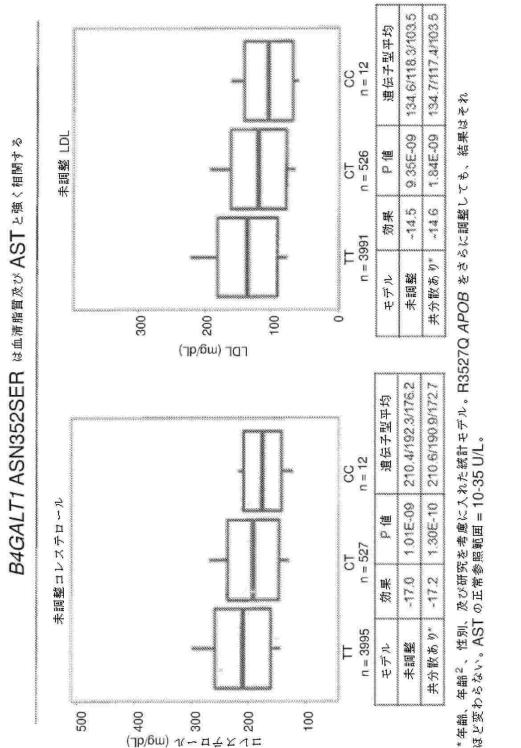
0.0774\* 遺伝子型平均は、年齢、年齢の二乗、性別及び研究の影響を除去して、臨床尺度から算出  
\*\* BMI にわたっても調整  
\*\*\* UMD において、年齢、性別、研究を共変数として用いたモデルで生成した結果

図 6 (続き)

50

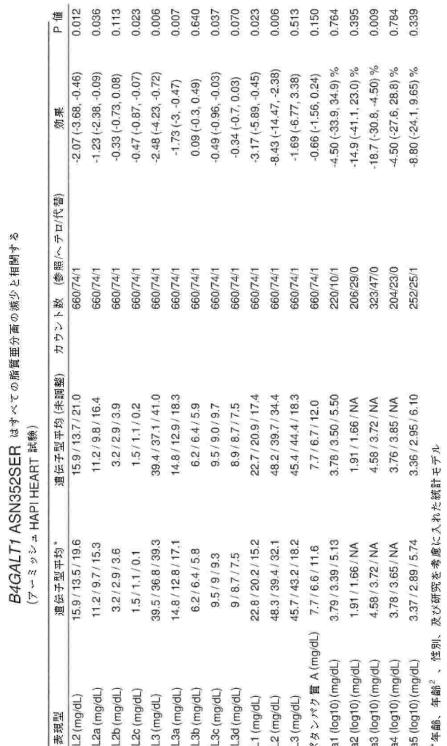
【 図 7 - 1 】

【図 7 - 2】



【図8】

【圖 9】



B4GA/T1 ASN352SEB はフィブリノゲンレベルの減少と相關する

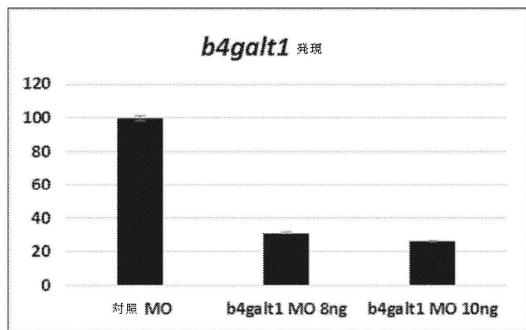
薬理型	遺伝子型平均		遺伝子型カント数 (参照ヘテロ型)	発現 (参照ヘテロ型)	P 値
	オルニチン (mg/dL)	アラニン (mg/dL)			
アラニン高	280.6 ± 255.8	280.5 ± 257.1	273.3	56.4 ± 43	23.97 (-36.35, 9.59)
アラニン低	276.2 ± 71.182	275.1 ± 97.0	54.6 ± 43	32.45 (-47.5, -7.41)	2.7E-05

節、年齢<sup>2</sup>、性別、及び研究参考書に記載された統計モデル

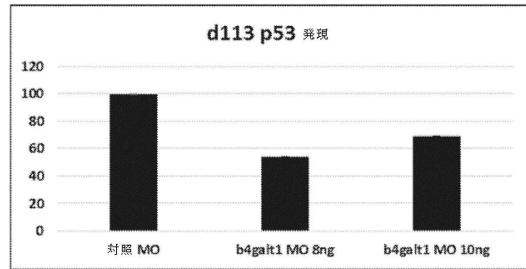
\* 年齢、年齢<sup>2</sup>、性別、及び研究を考慮に入れた統合モデル。R3527Q APOB をさらに調整しても、結果はそれ共分能であり。7.9%、6.02E-08、17.99/18.62/35.25

卷之四

【図 1 0】

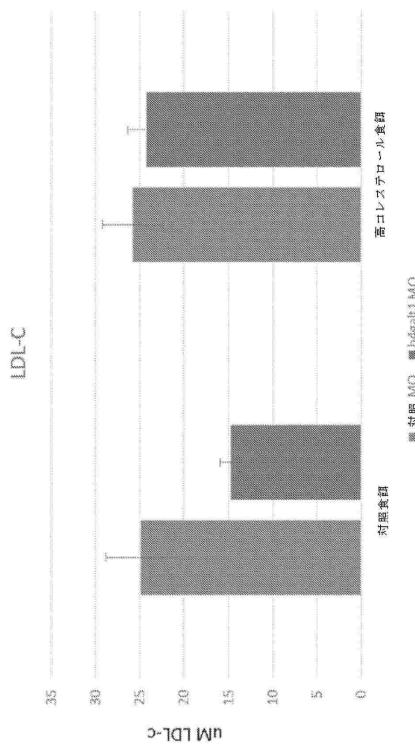


【図 1 1】

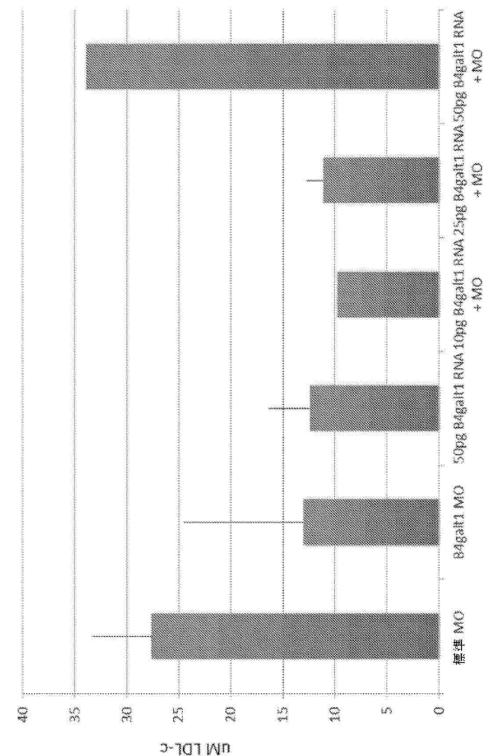


10

【図 1 2】



【図 1 3】



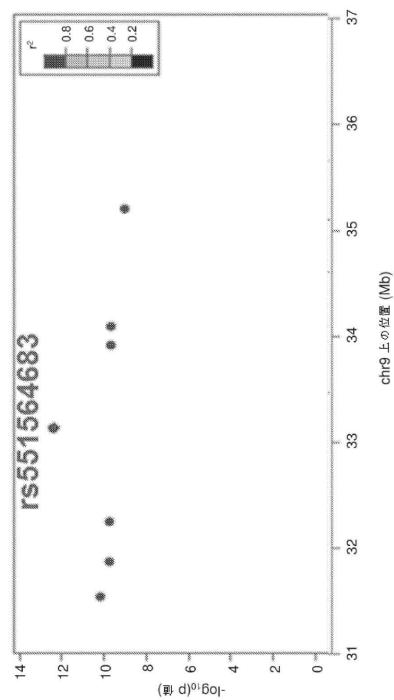
20

30

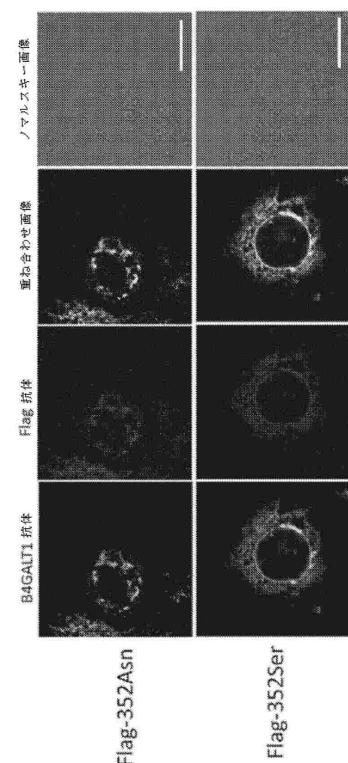
40

50

【図 1 4】



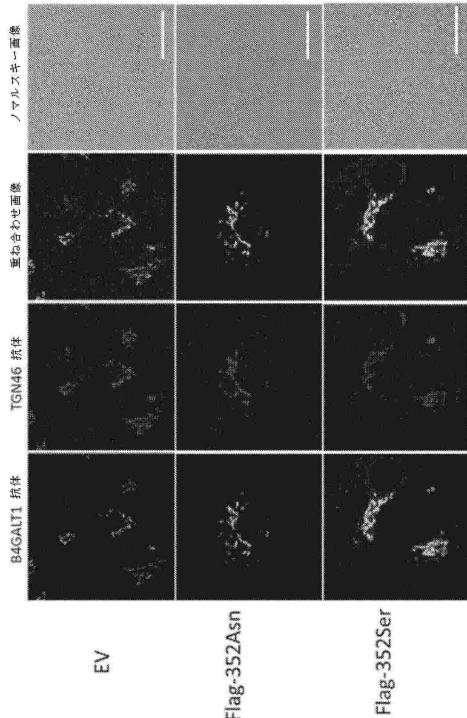
【図 1 5】



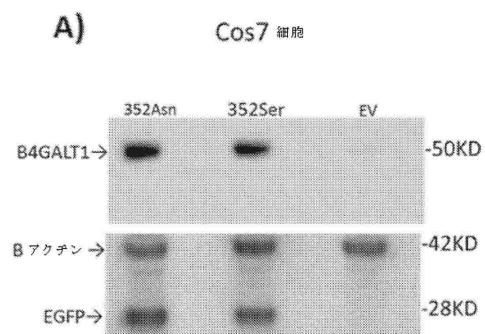
10

20

【図 1 6】



【図 1 7 - 1】



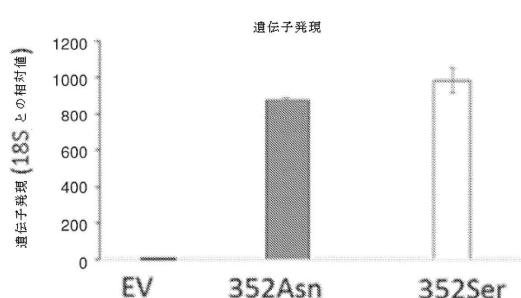
30

40

50

【図 17 - 2】

B)

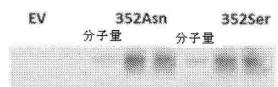


【図 18】

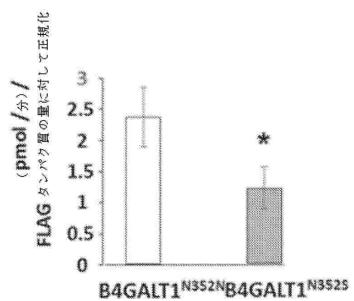
A) B4GALT1 抗体



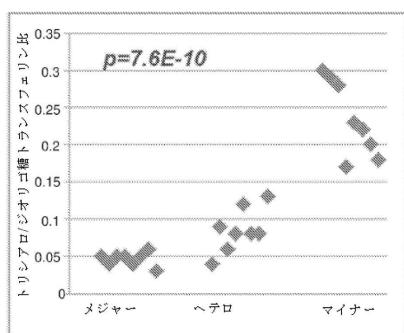
B) Flag 抗体



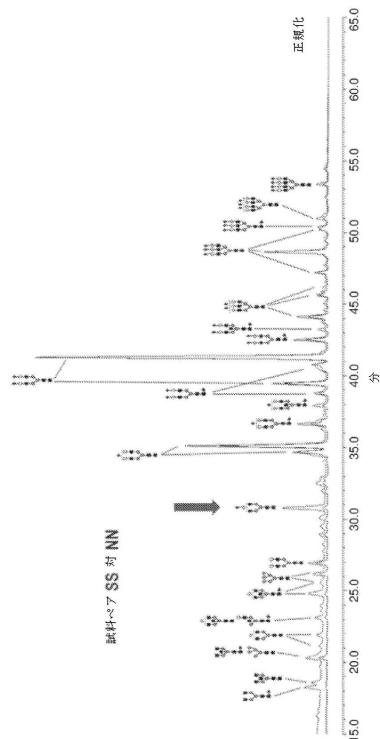
C) 活性



【図 19】



【図 20 - 1】



10

20

30

40

50

【図 2 0 - 2】

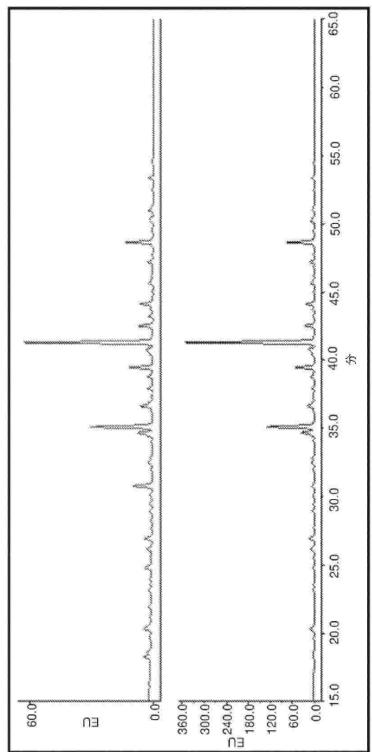


図 2 0 (続き)

10

20

【配列表】

0007661009000001.app

30

40

50

## フロントページの続き

## (51)国際特許分類

	F	I	
C 1 2 N	1/15 (2006.01)	C 1 2 N	1/15
C 1 2 N	1/19 (2006.01)	C 1 2 N	1/19
C 1 2 N	1/21 (2006.01)	C 1 2 N	1/21
C 1 2 N	5/10 (2006.01)	C 1 2 N	5/10
C 1 2 N	9/10 (2006.01)	C 1 2 N	9/10
C 1 2 N	15/09 (2006.01)	C 1 2 N	15/09
C 1 2 N	15/63 (2006.01)	C 1 2 N	15/09 110
C 1 2 Q	1/6874(2018.01)	C 1 2 N	15/63 Z
G 0 1 N	33/68 (2006.01)	C 1 2 Q	1/6874 Z
		G 0 1 N	33/68

(31)優先権主張番号 62/550,161

(32)優先日 平成29年8月25日(2017.8.25)

(33)優先権主張国・地域又は機関

米国(US)

(31)優先権主張番号 62/659,344

(32)優先日 平成30年4月18日(2018.4.18)

(33)優先権主張国・地域又は機関

米国(US)

弁理士 恩田 誠

(74)代理人 100068755

弁理士 恩田 博宣

(74)代理人 100142907

弁理士 本田 淳

(74)代理人 100152489

弁理士 中村 美樹

(72)発明者 モンタセール、メイ

アメリカ合衆国 21201 メリーランド州 ボルチモア ウエスト ファイエット ストリート 3  
05 アパートメント 1411

(72)発明者 ファン フート、クリストファー

アメリカ合衆国 10591 ニューヨーク州 タリータウン オールド ソー ミル リバー ロード  
777

(72)発明者 シュルディナー、アラン

アメリカ合衆国 10591 ニューヨーク州 タリータウン オールド ソー ミル リバー ロード  
777

(72)発明者 ガッタ、ジューシー デラ

アメリカ合衆国 10591 ニューヨーク州 タリータウン オールド ソー ミル リバー ロード  
777

(72)発明者 ヒーリー、マシュー

アメリカ合衆国 10591 ニューヨーク州 タリータウン オールド ソー ミル リバー ロード  
777

(72)発明者 プールネン、マリヤ

アメリカ合衆国 10591 ニューヨーク州 タリータウン オールド ソー ミル リバー ロード  
777

合議体

審判長 中村 浩

審判官 小金井 悟

審判官 荒木 英則

(56)参考文献 国際公開第2009/025645 (WO, A1)

国際公開第2009/102820 (WO, A2)

nature genetics, 2008年, Vol.40, pp.161-169  
(58)調査した分野 (Int.Cl. , DB名)

C 1 2 N 1 5 / 0 0 - 1 5 / 9 0

J S T P l u s ( J D r e a m I I I )

P u b M e d

C A P l u s / M E D L I N E / B I O S I S / R E G I S T R Y ( S T N )

S w i s s P r o t / G e n e S e q

G e n B a n k / E M B L / D D B J / G e n e S e q