

(19)日本国特許庁(JP)

(12)公表特許公報(A)

(11)公表番号

特表2023-543763

(P2023-543763A)

(43)公表日 令和5年10月18日(2023.10.18)

(51)国際特許分類		F I		テーマコード(参考)	
C 1 2 Q	1/6869(2018.01)	C 1 2 Q	1/6869	Z	2 G 0 4 5
A 6 1 P	1/16 (2006.01)	A 6 1 P	1/16		4 B 0 6 3
A 6 1 K	45/00 (2006.01)	A 6 1 K	45/00		4 C 0 8 4
A 6 1 K	38/02 (2006.01)	A 6 1 K	38/02		4 C 0 8 5
A 6 1 K	39/395(2006.01)	A 6 1 K	39/395	D	4 C 0 8 6
		審査請求	未請求	予備審査請求	未請求 (全74頁) 最終頁に続く

(21)出願番号	特願2023-518817(P2023-518817)	(71)出願人	500341551 ケンブリッジ エンタープライズ リミテッド
(86)(22)出願日	令和3年9月23日(2021.9.23)		イギリス国 シービー 2 1 ティーエヌ ケンブリッジシャー ケンブリッジ トリ ニティ レーン ジ オールド スクールズ
(85)翻訳文提出日	令和5年5月11日(2023.5.11)	(71)出願人	512010753 ゲノム・リサーチ・リミテッド
(86)国際出願番号	PCT/GB2021/052469		イギリス・グレーター・ロンドン・NW 1・2 B E・ロンドン・ユーストン・ロ ード・2 1 5・ギブス・ビルディング
(87)国際公開番号	WO2022/064198	(74)代理人	110000796 弁理士法人三枝国際特許事務所
(87)国際公開日	令和4年3月31日(2022.3.31)	(72)発明者	ホーア マシュー イギリス国 シービー 2 0 アールイー
(31)優先権主張番号	2015056.1		最終頁に続く
(32)優先日	令和2年9月23日(2020.9.23)		
(33)優先権主張国・地域又は機関	英国(GB)		
(81)指定国・地域	AP(BW,GH,GM,KE,LR,LS,MW,MZ,NA ,RW,SD,SL,ST,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,RU,TJ,TM),EP(AL,A T,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR ,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV,MC, 最終頁に続く		

(54)【発明の名称】 バイオマーカー

(57)【要約】

本発明は、対象における非アルコール性脂肪性肝疾患 (N A F L D) 又はアルコール関連脂肪性肝疾患 (A R L D) を診断及び/又は予後判定する方法であって、対象の1つ以上の肝細胞に選択的利点を付与するDNA、RNA及び/又はタンパク質における体細胞突然変異を検出することを含む、方法に関する。本発明はまた、治療剤による治療から利益を得る、N A F L D 若しくはA R L D に罹患している対象を同定するため、及び/又は疾患モニタリングの増加から利益を得る、N A F L D 若しくはA R L D に罹患している対象を同定するための方法を提供する。本発明はまた、N A F L D 又はA R L D の治療に有用な治療剤を提供する。

【特許請求の範囲】**【請求項 1】**

対象における非アルコール性脂肪性肝疾患（NAFLD）又はアルコール関連脂肪性肝疾患（ALD）を診断及び／又は予後判定するための方法であって、

a) 対象の1つ以上の肝細胞に由来するDNA、RNA及び／又はタンパク質を含む生物学的試料を提供することと、

b) 前記肝細胞に選択的利点を付与する前記DNA、RNA及び／又はタンパク質における体細胞突然変異を検出することであって、前記肝細胞に選択的利点を付与する体細胞突然変異の存在は、前記対象がNAFLD若しくはALDに罹患しているか、NAFLD若しくはALDを発症するリスクがあるか、より重症型の肝疾患を発症するリスクがあるか、及び／又は肝疾患に関連する疾患若しくは状態を発症するリスクがあることを示す、検出することと、を含む、方法。

10

【請求項 2】

ステップb)において検出される前記体細胞突然変異が、FOXO1、GPAM、CIDE B、ACVR2A、ALB及びTNRC6Bからなる群から選択される遺伝子内にある、請求項1に記載の方法。

【請求項 3】

ステップb)において検出される前記体細胞突然変異が、FOXO1、GPAM及びCIDE Bからなる群から選択される遺伝子内にある、請求項1又は2に記載の方法。

【請求項 4】

ステップb)で検出される前記体細胞突然変異が、前記FOXO1遺伝子内にある、請求項1～3のいずれか一項に記載の方法。

20

【請求項 5】

前記FOXO1遺伝子内の前記体細胞突然変異が、前記1つ以上の肝細胞によって発現されるFOXO1タンパク質のN末端14-3-3タンパク質結合モチーフ内のミスセンス及び／又はナンセンス突然変異をもたらす、請求項4に記載の方法。

【請求項 6】

前記FOXO1遺伝子内の前記体細胞突然変異が、前記1つ以上の肝細胞によって発現されるFOXO1タンパク質においてS22W若しくはR21Lアミノ酸置換又はS22ナンセンス突然変異をもたらす、請求項4又は5に記載の方法。

30

【請求項 7】

前記FOXO1遺伝子内の前記体細胞突然変異が、前記1つ以上の肝細胞によって発現されるFOXO1タンパク質においてS22Wアミノ酸置換をもたらす、請求項6に記載の方法。

【請求項 8】

ステップb)で検出される前記体細胞突然変異が、前記GPAM遺伝子内にある、請求項1～3のいずれか一項に記載の方法。

【請求項 9】

前記GPAM遺伝子内の前記体細胞突然変異が、前記1つ以上の肝細胞によって発現されるミトコンドリアグリセロール-3-リン酸アシルトランスフェラーゼ（GPAT）の機能を損なうか又は無効にする、請求項8に記載の方法。

40

【請求項 10】

ステップb)で検出される前記体細胞突然変異が、前記CIDE B遺伝子内にある、請求項1～3のいずれか一項に記載の方法。

【請求項 11】

前記CIDE B遺伝子内の前記体細胞突然変異が、前記1つ以上の肝細胞によって発現されるCIDE Bタンパクの機能を損なうか又は無効にする、請求項10に記載の方法。

【請求項 12】

ステップb)が、CLCN5遺伝子内の体細胞突然変異を検出するステップを更に含む、請求項1～11のいずれか一項に記載の方法。

50

【請求項 13】

前記生物学的試料が、循環遊離DNA (cfDNA) を含む、請求項 1 ~ 12 のいずれか一項に記載の方法。

【請求項 14】

前記生物学的試料が、肝生検試料に由来するDNA、RNA 及び / 又はタンパク質を含む、請求項 1 ~ 13 のいずれか一項に記載の方法。

【請求項 15】

前記生物学的試料が、1つ以上の肝細胞に由来するDNA、RNA 及び / 又はタンパク質を含む、請求項 1 ~ 14 のいずれか一項に記載の方法。

【請求項 16】

ステップ b) は、前記対象がNAFLD若しくはARLDに罹患しているか、NAFLD若しくはARLDを発症するリスクがあるか、並びに / 又は非アルコール性脂肪性肝炎 (NAFH)、肝線維症、肝硬変、がん (例えば、肝細胞がん (HCC)) 及び肝不全からなる群から選択されるより重症型の肝疾患を発症するリスクがあるかどうかを決定することを含む、請求項 1 ~ 15 のいずれか一項に記載の方法。

10

【請求項 17】

ステップ b) は、前記対象が、消化管がん、肥満、2型糖尿病、高血圧、脂質異常症 (例えば、高コレステロール血症) 及び心血管疾患からなる群から選択される肝疾患に関連する疾患又は状態を発症するリスクがあるかどうかを決定することを含む、請求項 1 ~ 16 のいずれか一項に記載の方法。

20

【請求項 18】

前記対象が、消化管がん、肥満、2型糖尿病、高血圧、脂質異常症 (例えば、高コレステロール血症) 及び / 又は心血管疾患にも罹患している対象である、請求項 1 ~ 16 のいずれか一項に記載の方法。

【請求項 19】

ステップ a) が、前記対象の2つ以上の肝細胞、例えば前記対象の約10 ~ 約50,000個の肝細胞に由来するDNA、RNA 及び / 又はタンパク質を含む生物学的試料を提供することを含む、請求項 1 ~ 18 のいずれか一項に記載の方法。

【請求項 20】

前記方法が、

i) 前記対象から得られた生物学的試料を提供することと、

ii) 前記生物学的試料中の前記FOXO1及び / 又はGPAM遺伝子における1つ以上の体細胞突然変異を検出することであって、1つ以上の体細胞突然変異の存在は、前記対象がNAFLDに罹患しているか、NAFLDを発症するリスクがあるか、より重症型の肝疾患を発症するリスクがあるか、及び / 又は肝疾患に関連する疾患若しくは状態を発症するリスクがあることを示す、検出することと、を含む、請求項 1 ~ 19 のいずれか一項に記載の方法。

30

【請求項 21】

疾患モニタリングの増加から利益を得る、NAFLD又はARLDに罹患している対象を同定するための方法であって、請求項 1 ~ 20 のいずれか一項に記載の方法ステップを含む、方法。

40

【請求項 22】

肝臓におけるFOXO1、CIDEB、ACVR2A、GPAM又はTNRC6Bタンパク質活性を阻害又は調節する治療剤による治療から利益を得る、NAFLD又はARLDに罹患している対象を同定するための方法であって、請求項 1 ~ 20 のいずれか一項に記載の方法ステップを含む、方法。

【請求項 23】

NAFLD又はARLDの治療に使用するための治療剤であって、前記使用が、

請求項 1 ~ 20 のいずれか一項に記載の方法を実行することと、

前記治療剤を前記対象に投与することと、を含み、前記治療剤が、FOXO1、CID

50

EB、ACVR2A、GPAM又はTNRC6Bタンパク活性を阻害又は調節するものである、治療剤。

【請求項24】

NAFLD又はARLDに罹患している対象を治療する方法であって、請求項1～20のいずれか一項に記載の方法を実行すること、

FOXO1、CIDE B、ACVR2A、GPAM又はTNRC6Bタンパク質活性を阻害又は調節する、ある用量の治療剤を前記対象に投与することと、を含む、方法。

【請求項25】

前記治療剤が、小分子、ペプチド（環状ペプチドを含む）、タンパク質（例えば、治療用抗体、抗体断片及び抗体模倣物）、核酸（例えば、アンチセンスヌクレオチド配列；三重らせん；siRNA又はmiRNA；及び生物学的に活性なタンパク質、ポリペプチド又はペプチドをコードするヌクレオチド配列を含むがこれらに限定されない、DNA及びRNAヌクレオチド）、合成又は天然無機分子、合成又は天然有機分子、及びCRISPRからなる群から選択される、請求項23に記載の使用又は請求項24に記載の方法。

10

【請求項26】

請求項1～20のいずれか一項に記載の方法に従って、対象におけるNAFLD又はARLDの診断又は予後診断に使用するためのインビトロ診断キットであって、FOXO1、CIDE B、ACVR2A、ALB、GPAM及びTNRC6Bからなる群から選択される遺伝子内の1つ以上の体細胞突然変異を検出し、任意選択的に前記CLCN5遺伝子若しくはNEAT1遺伝子内の1つ以上の体細胞突然変異を検出し、かつ/又はテロメア長を測定するための1つ以上の試薬を含む、インビトロ診断キット。

20

【請求項27】

対象におけるNAFLD又はARLDを診断又は予後判定するための方法であって、ある用量の診断プローブを前記対象に投与するステップと、

前記対象における前記診断プローブを検出するステップであって、前記診断プローブが、前記対象のFOXO1、CIDE B、ACVR2A、ALB、GPAM及びTNRC6Bからなる群から選択される遺伝子内の1つ以上の体細胞突然変異の存在及び/又は非存在を示し、前記FOXO1、CIDE B、ACVR2A、ALB、GPAM及びTNRC6Bからなる群から選択される遺伝子内の1つ以上の体細胞突然変異の存在は、前記対象がNAFLD若しくはARLDに罹患しているか、NAFLD若しくはARLDを発症するリスクがあるか、より重症型の肝疾患を発症するリスクがあるか、及び/又は消化管がんなどの肝疾患に関連する疾患若しくは状態を発症するリスクがあることを示す、ステップと、を含む、方法。

30

【発明の詳細な説明】

【技術分野】

【0001】

序論

本発明は、対象における非アルコール性脂肪性肝疾患（non-alcoholic fatty liver disease、NAFLD）又はアルコール関連脂肪性肝疾患（alcohol-related fatty liver disease、ARLD）を診断及び/又は予後判定する方法に関する。本発明はまた、治療剤による治療から利益を得る、NAFLD若しくはARLDに罹患している対象を同定するため、及び/又は疾患モニタリングの増加から利益を得る、NAFLD若しくはARLDに罹患している対象を同定するための方法にも関する。NAFLD又はARLDの治療に有用な治療剤も本明細書において提供される。

40

【背景技術】

【0002】

世界人口の四分の一が罹患していると推定される非アルコール性脂肪性肝疾患（NAFLD）は、非アルコール性脂肪性肝炎（non-alcoholic steatohepatitis、NASH）、肝線維症、肝硬変、肝細胞がん（hepatocellular carcinoma、HCC）、及び肝不全に進行し得る肝臓の炎症及び損傷の症候群である。NAFLD患者の約10%～30

50

%がNASHを発症し、NASH患者の約25%~40%が、不良な長期予後に関連する肝線維症及び肝硬変を発症し続けると推定されている(Dyson et al., Frontline Gastroenterology 2014; 5, 211(218))。研究により、NAFLDを有する患者は、肝臓がん及び消化管型がんを発症するリスクも高いことが示されている(Allens et al., J Hepatol. 2019; 71(6): 1229-1236)。

【0003】

NAFLDの初期段階は、肝臓脂肪症の形成を特徴とし、これは、疾患が進行し、最終的により深刻な肝臓損傷を引き起こすまで気付かれない可能性がある。したがって、NAFLDの初期段階は、通常の血液検査後の血液中の異常な肝酵素の偶発的な検出によってのみ検出されることが多い。NAFLDが疑われると、超音波などの画像化技術が、肝臓における肝臓脂肪症の存在を確認するための初期調査においてしばしば使用される。しかしながら、NAFLDを決定的に診断するためには、侵襲性肝生検を実施して、肝臓組織における肝臓脂肪症の存在を確認することがしばしば必要である。これは、患者の危害のリスクを伴い、したがって、特定の場合にのみ使用される。したがって、ほとんどの場合、患者がNAFLDを有するかどうか、及びそれが進行性であるかどうかの評価は、画像化及び血液検査の使用を通じて間接的である。

【0004】

疾患特異的バイオマーカー、例えば遺伝子及び/又はタンパク質レベルでの疾患特異的突然変異の同定により、様々な疾患及び状態、特にがんを確実に診断及び予後診断するための見込みが示されている。これに関して、インスリン抵抗性又は脂質合成に関連する遺伝子における遺伝的変異(すなわち、生殖系列多型)は、NAFLDなどの肝疾患を発症する遺伝的リスクに対して軽微な影響を有することが示唆されている(例えば、Valenti et al., Journal of Hepatology, 2009, 50, S265及びHakim et al., Hepatology, 2021, doi: 10.1002/hep.32038を参照されたい)。しかしながら、NAFLD及びARLDの発症は主に生活様式及び環境因子によって駆動されるので、遺伝的変異の検出は、患者の肝臓の状態への限られた洞察のみを提供する。肝硬変及び肝細胞がん(HCC)などの重症型の肝疾患に罹患している患者の肝臓における体細胞突然変異の存在も報告されている(例えば、Zhu et al., Cell. 2019; 177(3): 608-621, Kim et al., J Gastroenterol. 2019; 54(7): 628-640, Brunner et al., Nature, 2019, 574, 538-542, Torrecilla et al., J Hepatol. 2017; 67(6): 1222-1231, and Nault et al., Hepatology. 2014; 60(6): 1983-92を参照されたい)。

【0005】

かなりの割合のヒト集団の健康及び平均余命に対するNAFLDの影響にもかかわらず、現在、NAFLDを有する患者並びに肝硬変及びHCCなどのより重症型の肝疾患を発症する可能性が高い患者を同定する信頼できる診断又は予後診断検査はない。注目すべきことに、NAFLDの治療に利用可能な特定の治療もない。治療は、多くの場合、運動及び食事プログラムの使用、並びに/又は肥満、2型糖尿病、高血圧及び高コレステロールなどの関連危険因子を低減するための治療に限定される。

【0006】

アルコール関連脂肪性肝疾患(ARLD)の初期段階は、異なる病因を有するが、NAFLDと同様の病態生理を有する。NAFLDに関しては、初期段階のARLDを有する患者を同定するための信頼できる診断又は予後診断検査もなく、ARLDの治療に利用可能な特定の治療もない。

【0007】

したがって、NAFLD又はARLDの早期診断及びより重症型の肝疾患を発症するリスクがある対象の同定を可能にする、より信頼性の高い診断方法及び予後診断方法が依然

10

20

30

40

50

として必要とされている。また、N A F L D 又は A R L D の治療及び / 又は予防に有効な治療も必要とされている。

【発明の概要】

【0008】

本発明は、対象における非アルコール性脂肪性肝疾患 (N A F L D) 又はアルコール関連脂肪性肝疾患 (A R L D) を診断及び / 又は予後判定するための方法であって、

a) 対象の1つ以上の肝細胞に由来する D N A 、 R N A 及び / 又はタンパク質を含む生物学的試料を提供することと、

b) 肝細胞に選択的利点を付与する D N A 、 R N A 及び / 又はタンパク質における体細胞突然変異を検出することであって、肝細胞に選択的利点を付与する体細胞突然変異の存在は、対象が N A F L D 若しくは A R L D に罹患しているか、N A F L D 若しくは A R L D を発症するリスクがあるか、より重症型の肝疾患を発症するリスクがあるか、及び / 又は肝疾患に関連する疾患若しくは状態を発症するリスクがあることを示す、検出することと、を含む方法を提供する。

10

【0009】

本発明は、対象における非アルコール性脂肪性肝疾患 (N A F L D) を診断及び / 又は予後判定するための方法であって、

a) 対象から得られた生物学的試料を提供することと、

b) 生物学的試料中の F O X O 1 及び / 又は G P A M 遺伝子内の1つ以上の体細胞突然変異を検出することであって、1つ以上の体細胞突然変異の存在は、対象が N A F L D に罹患しているか、N A F L D を発症するリスクがあるか、より重症型の肝疾患を発症するリスクがあるか、及び / 又は肝疾患に関連する疾患若しくは状態を発症するリスクがあることを示す、検出することと、を含む方法を提供する。

20

【0010】

本発明はまた、F O X O 1 、 G P A M 、 C I D E B 、 A C V R 2 A 、 A L B 若しくは T N R C 6 B タンパク活性を阻害若しくは調節する治療剤による治療から利益を得る、N A F L D 又は A R L D に罹患している対象を同定するための方法、及び / 又は疾患モニタリングの増加から利益を得る、N A F L D 又は A R L D に罹患している対象を同定するための方法を提供する。また、本明細書に開示されるのは、N A F L D 又は A R L D の治療に使用するための治療剤であって、F O X O 1 、 G P A M 、 C I D E B 、 A C V R 2 A 、 A L B 又は T N R C 6 B タンパク活性を阻害又は調節するものである、治療剤、及び F O X O 1 、 G P A M 、 C I D E B 、 A C V R 2 A 、 A L B 又は T N R C 6 B タンパク活性を阻害又は調節する治療剤を対象に投与することによって、N A F L D 又は A R L D を治療又は予防する方法である。

30

【0011】

本明細書に記載されるように、本発明者らは、アルコール関連脂肪性肝疾患 (A R L D) 、非アルコール性脂肪性肝疾患 (N A F L D) 及び正常対照 (すなわち、健康な対象由来の肝臓試料) を含む 3 4 個の肝臓試料にわたって 1 5 9 0 個のゲノムからの体細胞突然変異を分析した。この分析から、本発明者らは、肝細胞に選択的利点を付与する体細胞突然変異 (がんにおけるドライバー突然変異と同種であると考えられる突然変異) の発生が、対象における N A F L D 及び A R L D の発症の中心であることを発見した。本発明者らは、対象の肝細胞におけるそのような体細胞突然変異の (直接的若しくは間接的な) 検出、定量及び / 又はモニタリングが、N A F L D 及び A R L D の診断及び予後診断検査に特に有用であることを見出した。特に、N A F L D 又は A R L D を有する患者において、本発明者らは、F O X O 1 、 G P A M 、 C I D E B 、 A C V R 2 A 、 A L B 及び T N R C 6 B 遺伝子における様々な突然変異を同定した。注目すべきことに、本発明者らは、F O X O 1 、 G P A M 、 C I D E B 、 A C V R 2 A 、 A L B 及び T N R C 6 B 遺伝子において同定された突然変異を N A F L D 又は A R L D のマーカーとして使用することにより、N A F L D 又は A R L D の初期診断及び予後診断を可能にすることができることを見出した。更に、本明細書に記載の F O X O 1 、 G P A M 、 C I D E B 、 A C V R 2 A 、 A

40

50

L B 及び T N R C 6 B 突然変異はまた、N A F L D 又は A R L D の治療及び / 又は予防のための新規分子標的を提供する。

【 0 0 1 2 】

本発明はまた、対象における N A F L D 又は A R L D の診断及び / 又は予後診断に使用するためのインビトロ診断キットであって、F O X O 1、G P A M、C I D E B、A C V R 2 A、A L B 及び T N R C 6 B からなる群から選択される 1 つ以上の遺伝子内の 1 つ以上の体細胞突然変異を検出し、任意選択的に C L C N 5 遺伝子、N E A T 1 遺伝子内の 1 つ以上の体細胞突然変異を検出し、かつ / 又はテロメア長を測定するための 1 つ以上の試薬を含む、インビトロ診断キットを提供する。

【 0 0 1 3 】

誤解を避けるために、本明細書で使用される場合、「アルコール関連肝疾患」及び「A R L D」という用語は、過剰なアルコール消費によって引き起こされる、対象における肝臓損傷を指す。N A F L D に関して、A R L D の初期段階は、典型的には、肝臓脂肪症の形成を特徴とする。過剰なアルコール消費によって引き起こされる肝臓脂肪症の形成は、アルコール性脂肪性肝疾患 (alcoholic fatty liver disease、A F L D) と呼ばれることもある。A R L D はまた、アルコール性肝炎及び肝硬変に関連している可能性がある。したがって、A F L D は、A R L D の亜属と考えられ得る。

【 0 0 1 4 】

また、誤解を避けるために、本出願全体を通して、A R L D という用語は、A F L D という用語を組み込むものと考えべきである。したがって、本発明はまた、対象における N A F L D 又は A F L D を診断及び / 又は予後判定する方法、治療剤による治療から利益を得る、N A F L D 又は A F L D に罹患している対象を同定するための方法、疾患モニタリングの増加から利益を得る N A F L D 又は A F L D に罹患している対象を同定するための方法、及び N A F L D 又は A F L D の治療に有用な治療剤に関する。

【 図面の簡単な説明 】

【 0 0 1 5 】

【 図 1 】慢性肝疾患において獲得された収束性 F O X O 1 突然変異を調査するために使用される階層的実験設計の概要を示す。正常対照、A R L D 及び N A F L D について 3 4 個の肝臓をサンプリングした (3 2 個の肝臓については 1 個の試料 / 肝臓、2 個の肝臓については 8 個の試料 / 肝臓)。これらの試料をレーザーキャプチャーマイクロダイセクションに供して、2 1 ~ 5 2 個のマイクロダイセクション / 試料を生成し、そのそれぞれを個別に全ゲノム配列決定した (全体で 1 5 9 0 個の全ゲノム)。

【 図 2 A 】罹患患者からのマイクロダイセクションによって分類された F O X O 1 における体細胞突然変異の分布を示す。円グラフは、各マイクロダイセクションにおける突然変異体対立遺伝子を報告する配列決定リードの割合を示す。ホットスポット S 2 2 残基は、1 4 - 3 - 3 核外輸送タンパク質による結合並びに A K T 1 及び A M P K によるリン酸化のための標準的な、高度に保存されたモチーフ内にある。

【 図 2 B 】 P D 3 7 2 3 9 における F O X O 1 突然変異に関する。P D 3 7 2 3 9 及び P D 3 7 9 1 8 はいずれも N A F L D 患者である。図 2 B 及び図 2 C の左パネルは系統樹を示し、より濃い線の枝は、独立して獲得された突然変異を示している。実線は、ネスティングが鳩の巣原理に従うことを示す。破線は、肝細胞が細胞の 1 0 0 % 未満であることを表すと仮定して、ネスティングが鳩の巣原理に従うことを示す。右パネルは、患者の肝生検のヘマトキシリン及びエオシン (haematoxylin and eosin、H & E) 染色光学顕微鏡写真上にマッピングされた系統樹からのクローンを示し、F O X O 1 突然変異体クローンは、系統樹上に示されたより濃い線の枝に一致するように網掛けした。

【 図 2 C 】 P D 3 7 9 1 8 における F O X O 1 突然変異に関する。P D 3 7 2 3 9 及び P D 3 7 9 1 8 はいずれも N A F L D 患者である。図 2 B 及び図 2 C の左パネルは系統樹を示し、より濃い線の枝は、独立して獲得された突然変異を示している。実線は、ネスティングが鳩の巣原理に従うことを示す。破線は、肝細胞が細胞の 1 0 0 % 未満であることを表すと仮定して、ネスティングが鳩の巣原理に従うことを示す。右パネルは、患者の肝生

10

20

30

40

50

検のヘマトキシリン及びエオシン (haematoxylin and eosin、H & E) 染色光学顕微鏡写真上にマッピングされた系統樹からのクローンを示し、FOXO1 突然変異体クローンは、系統樹上に示されたより濃い線の枝に一致するように網掛けした。

【図 2 D】NAFLD 患者である PD37907 からのマイクロダイセクションのうちの 1 つにおける第 13 染色体に影響を及ぼす染色体異常事象を示す。黒点は、染色体に沿った補正リード深度を表す。線及び弧は、構造バリエーションを表す。FOXO1 を切断する構造バリエーションが強調されており (図では、「FOXO1 を再配置する」と示されている)、第 1 のイントロン内の遺伝子を切断し、第 1 のコードエクソンを保存するが、残りのコードエクソンを欠失させると予測される。

【図 2 E】PD37239 の H & E 染色切片上に置かれた図 2 B のクローンマップを示す。図 2 E の左側には、FOXO1 突然変異を有する及び有しない代表的な試料からの生の配列決定データが示されており、H & E 切片上のそれらの物理的位置が矢印によって示されている。S22W 及び R21L 突然変異の位置を矢印で示す。H & E 切片の周りに配置された散布図は、試料対における突然変異のバリエーション対立遺伝子画分プロットを表す。x 軸及び y 軸タイトルのグレースケールシェードは、H & E 切片のクローンマップのグレースケールシェードと一致する。クローン的に関連する試料対において、突然変異のほとんどは、両方の試料によって共有され、ゼロではないバリエーション対立遺伝子画分を有する突然変異のクラウドとして明らかである。クローン的に関連のない試料において、突然変異は x 軸及び y 軸に沿っており、1 つの例外は FOXO1 突然変異であり、それが 2 つのクローンにおいて独立して獲得されることを示す。

【図 3 A】PD37234 についての系統樹及びクローンマップを示す。図の左パネルは系統樹を示し、より濃い線の枝は、独立して獲得された突然変異を示している。実線は、ネスティングが鳩の巣原理に従うことを示す。破線は、肝細胞が細胞の 100% 未満であることを表すと仮定して、ネスティングが鳩の巣原理に従うことを示す。右パネルは、肝臓の H & E 染色顕微鏡写真上にマッピングされた系統樹からのクローンを示し、FOXO1 - 突然変異体クローンは、系統樹上に示されたより濃い線の枝に一致するように網掛けした。

【図 3 B】PD37105 についての系統樹及びクローンマップを示す。図の左パネルは系統樹を示し、より濃い線の枝は、独立して獲得された突然変異を示している。実線は、ネスティングが鳩の巣原理に従うことを示す。破線は、肝細胞が細胞の 100% 未満であることを表すと仮定して、ネスティングが鳩の巣原理に従うことを示す。右パネルは、肝臓の H & E 染色顕微鏡写真上にマッピングされた系統樹からのクローンを示し、FOXO1 - 突然変異体クローンは、系統樹上に示されたより濃い線の枝に一致するように網掛けした。

【図 3 C】PD37245 についての系統樹及びクローンマップを示す。図の左パネルは系統樹を示し、より濃い線の枝は、独立して獲得された突然変異を示している。実線は、ネスティングが鳩の巣原理に従うことを示す。破線は、肝細胞が細胞の 100% 未満であることを表すと仮定して、ネスティングが鳩の巣原理に従うことを示す。右パネルは、肝臓の H & E 染色顕微鏡写真上にマッピングされた系統樹からのクローンを示し、FOXO1 - 突然変異体クローンは、系統樹上に示されたより濃い線の枝に一致するように網掛けした。

【図 4 A】慢性肝疾患における CIDE B 内の体細胞突然変異の分布を示す。アミノ酸残基は、野生型タンパク質配列の上に示される慢性肝疾患において観察された突然変異とともに、タイプごとに着色されている。

【図 4 B】NAFLD の患者である PD48637 から分析されたクイノー (Couinaud) セグメントのうちの 1 つにおける CIDE B 突然変異を示す。左パネルは系統樹を示し、より濃い線の枝は、独立して獲得された突然変異を示している。実線は、ネスティングが鳩の巣原理に従うことを示す。破線は、肝細胞が細胞の 100% 未満であることを表すと仮定して、ネスティングが鳩の巣原理に従うことを示す。右パネルは、肝臓の H & E 染色顕微鏡写真上にマッピングされた系統樹からのクローンを示し、CIDE B - 突然変

10

20

30

40

50

異体クローンは、系統樹上に示されたより濃い線の枝に一致するように網掛けした。1つのクローンは、C I D E Bにおいて、異なる対立遺伝子上に2つの独立した点突然変異を有した（複合ヘテロ接合性）。

【図4C】慢性肝疾患患者におけるC I D E B突然変異の更なる例を示す。C I D E B突然変異を有するPD48367のクイノーセグメントのうちの一つについて、系統樹及びクローンマップを示す。

【図5A】ゲノム位置に従うG P A Mにおける体細胞突然変異の分布を示す。円グラフは、各マイクロダイセクションにおける突然変異体対立遺伝子を報告する配列決定リードの割合を示す。代表的な種について、各突然変異残基±3アミノ酸の進化的保存を示す複数の配列アラインメントを示す。

【図5B】A R L D患者であるPD37110からのマイクロダイセクションにおけるG P A Mの上流の縦列重複を示す。G P A Mは無傷のままであるが、縦列重複は遺伝子の20kb上流から始まる。

【図5C】A R L D患者であるPD37231におけるG P A M突然変異を示す。左パネルは系統樹を示し、より濃い線の枝は、独立して獲得された突然変異を示している。実線は、ネスティングが鳩の巣原理に従うことを示す。破線は、肝細胞が細胞の100%未満であることを表すと仮定して、ネスティングが鳩の巣原理に従うことを示す。右パネルは、肝臓のH & E染色顕微鏡写真上にマッピングされた系統樹からのクローンを示し、G P A M - 突然変異体クローンは、系統樹上に示されたより濃い線の分岐に一致するように網掛けした。

【図5D】A R L D患者におけるG P A M突然変異の更なる例を示す。系統樹及びクローンマップは、PD37111のクイノーセグメントのうちの一つ（図5D）及びPD37232のクイノーセグメントのうちの一つ（図5E）について示される。

【図5E】A R L D患者におけるG P A M突然変異の更なる例を示す。系統樹及びクローンマップは、PD37111のクイノーセグメントのうちの一つ（図5D）及びPD37232のクイノーセグメントのうちの一つ（図5E）について示される。

【図6A】ゲノム位置に従うA C V R 2 Aにおける体細胞突然変異の分布を示す。円グラフは、各マイクロダイセクションにおける突然変異体対立遺伝子を報告する配列決定リードの割合を示す。

【図6B】A C V R 2 Aのコピー喪失を生じる構造バリエーションを示す異なる患者における2つのマイクロダイセクションを示す。黒点は、染色体に沿った補正リード深度を表す。線及び弧は、構造バリエーションを表す。

【図7】ゲノム位置に従うC L C N 5における体細胞突然変異の分布を示す。円グラフは、各マイクロダイセクションにおける突然変異体対立遺伝子を報告する配列決定リードの割合を示す。

【図8】ゲノム位置に従う、長鎖非コードR N A N E A T 1における体細胞突然変異の分布を示す。円グラフは、各マイクロダイセクションにおける突然変異体対立遺伝子を報告する配列決定リードの割合を示す。

【図9】ゲノム位置に従うT N R C 6 Bにおける体細胞突然変異の分布を示す。円グラフは、各マイクロダイセクションにおける突然変異体対立遺伝子を報告する配列決定リードの割合を示す。

【図10A】C末端e G F Pと融合したF O X O 1の示された野生型又は突然変異体構築物でトランスフェクトされたH e p G 2細胞の生細胞画像化を示す。細胞を、核（H o e c h s t 33342）及び細胞質（S P Y - 555 - アクチン）マーカーで対比染色した。生細胞画像化を、一晚の血清飢餓、次いで100nMインスリンでの刺激後に行った。

【図10B】生細胞画像化（野生型細胞、 $n > 6186$ 及びF O X O 1 S 2 2 W細胞、 $n \pm 7172$ /時点）中のl o g核 - 細胞質蛍光比（平均±S E M）として表されるe G F P局在化の定量を示す。

【図10C】4つの条件（野生型F O X O 1構築物（インスリンあり又はなし）；S 2 2

10

20

30

40

50

W F O X O 1 構築物 (インスリンあり又はなし)) にわたってそれぞれ 5 回反復して (横列) 測定された H e p G 2 細胞において測定された代謝産物 (縦列) の濃度のヒートマップを示す。図に示されているのは、突然変異体構築物と野生型構築物との間で多重仮説検定 ($q < 0.01$) の補正後に有意に異なった 43 個の代謝産物であり、ペントースリン酸経路及び解糖 / 糖新生経路からの中間体が太字で強調されている (すなわち、ヘキソース - リン酸、ジヒドロキシアセトン - リン酸、ペントース - リン酸、セドヘプツロース 7 - リン酸、グリセロール - 3 - リン酸及びグリセルアルデヒド - 3 - リン酸)。

【図 10D】C 末端 GFP と融合した F O X O 1 の示された野生型構築物又は突然変異体構築物によるトランスフェクション後の H e p G 2 細胞を示す。細胞を D A P I で対比染色して核を強調し、一晚の血清飢餓条件後 (左) 及び 100 nM インスリンへの 15 分間の曝露後 (右) に画像化した。

10

【図 10E】F O X O 1 - G F P の核 - 細胞質比によるスケールで擬似着色された H e p G 2 細胞のカバースリップ全体の広視野図を示す。血清飢餓 (左)、インスリン 100 nM に 15 分間曝露した後 (中)、又は 5% ウシ胎児血清 (foetal calf serum、F C S) に 15 分間曝露した後 (右)、の条件下で細胞を画像化した。

【図 10F】H C C 細胞株における野生型及び突然変異体 F O X O 1 - G F P 構築物についての核 - 細胞質比を示す。F O X O 1 - G F P の核 - 細胞質比によるスケールで擬似着色された H e p 3 B (図 11A) 及び P L C / P R F 5 (図 11B) 細胞の広視野図を示す。血清飢餓 (左)、15 分間ウシ胎児血清 (F C S) (右) の 15 分間インスリン 100 nM に曝露した後 (中)、の条件下で細胞を画像化した。

20

【図 10G】H C C 細胞株における野生型及び突然変異体 F O X O 1 - G F P 構築物についての核 - 細胞質比を示す。F O X O 1 - G F P の核 - 細胞質比によるスケールで擬似着色された H e p 3 B (図 11A) 及び P L C / P R F 5 (図 11B) 細胞の広視野図を示す。血清飢餓 (左)、15 分間ウシ胎児血清 (F C S) (右) の 15 分間インスリン 100 nM に曝露した後 (中)、の条件下で細胞を画像化した。

【図 10H】示されるような異所性 e G F P タグ付き野生型又は突然変異体 F O X O 1 構築物を発現し、ビヒクル又はインスリン (100 nM) で 15 分間処理された H e p G 2 細胞の免疫プロットを示す。細胞を、イムノブロットングによって示されたタンパク質について分析した。分子量マーカー (k D a) を示す。

【図 11A】遺伝子オントロジー (Gene Ontology、GO; <http://geneontology.org> (GO:0061621)) からの「標準的解糖」遺伝子セットにおける遺伝子についての遺伝子発現レベルのヒートマップを示す。x 軸上の遺伝子の順序は、有意性のレベル (及び変化の方向) によって決定され、y 軸上の試料の順序は、状態 (F O X O 1 状態及びインスリン状態) による。

30

【図 11B】リアクトーム (Reactome (R - H S A - 69278; <https://reactome.org/>)) からの「細胞周期、有糸分裂」遺伝子セットにおける遺伝子についての遺伝子発現レベルのヒートマップを示す。x 軸上の遺伝子の順序は、有意性のレベル (及び変化の方向) によって決定され、y 軸上の試料の順序は、状態 (F O X O 1 状態及びインスリン状態) による。

【図 11C】リアクトーム (9615017 - <https://reactome.org/>) の「酸化ストレス、代謝及びニューロン遺伝子の F O X O 媒介転写」遺伝子セット (図 11C); GO (0016042) の「脂質異化プロセス」遺伝子セット (図 11D); 及び GO (0006915) の「アポトーシス」遺伝子セット (図 11E) についての濃縮プロットを示す。それぞれにおいて、上のパネルは、遺伝子セットが F O X O 1 突然変異体構築物の存在下で最も上方制御されたものから最も下方制御されたものへとトラバースされるときに累積濃縮スコアを反映する。それぞれの下のパネルは、測定された全ての遺伝子にわたる遺伝子セットにおける各遺伝子の順位を示す。

40

【図 11D】リアクトーム (9615017 - <https://reactome.org/>) の「酸化ストレス、代謝及びニューロン遺伝子の F O X O 媒介転写」遺伝子セット (図 11C); GO (0016042) の「脂質異化プロセス」遺伝子セット (図 11

50

D) ; 及び GO (0 0 0 6 9 1 5) の「アポトーシス」遺伝子セット (図 1 1 E) についての濃縮プロットを示す。それぞれにおいて、上のパネルは、遺伝子セットが FOXO1 突然変異体構築物の存在下で最も上方制御されたものから最も下方制御されたものへとトランスラバースされる際の累積濃縮スコアを反映する。それぞれの下のパネルは、測定された全ての遺伝子にわたる遺伝子セットにおける各遺伝子の順位を示す。

【図 1 1 E】リアクトーム (9 6 1 5 0 1 7 - https : / / reactome . org /) の「酸化ストレス、代謝及びニューロン遺伝子の FOXO 媒介転写」遺伝子セット (図 1 1 C) ; GO (0 0 1 6 0 4 2) の「脂質異化プロセス」遺伝子セット (図 1 1 D) ; 及び GO (0 0 0 6 9 1 5) の「アポトーシス」遺伝子セット (図 1 1 E) についての濃縮プロットを示す。それぞれにおいて、上のパネルは、遺伝子セットが FOXO1 突然変異体構築物の存在下で最も上方制御されたものから最も下方制御されたものへとトランスラバースされる際の累積濃縮スコアを反映する。それぞれの下のパネルは、測定された全ての遺伝子にわたる遺伝子セットにおける各遺伝子の順位を示す。

【図 1 2 A】各患者において分析された試料から外挿された、ドライバー突然変異を保有する推定累積肝臓質量を示す積み重ね棒グラフを示す。計算は、各患者について 1 5 0 0 g の総肝臓質量を仮定する。バーは、同定された 6 つの反復的に突然変異した遺伝子の各々について斜線で示され、x 軸上の患者コードは疾患状態について着色されている。

【図 1 2 B】野生型クローンと比較した、4 つの最も頻繁に突然変異した遺伝子 (FOXO1、CIDE B、GPAM 及び ACVR2A) についての推定クローンサイズを示す。点は箱ひげ図上に重ねられ、ここで、中央値は太い黒線で示され、四分位数範囲は細い黒い箱で示される。ひげは、データの全範囲又は 2 5 / 7 5 パーセンタイル + 1 . 5 x 四分位数間範囲 (いずれか小さい方) を示す。

【図 1 2 C】特定の遺伝子内に突然変異を有するクローンを保有したか否かによる、コホートにおける患者の年齢の分布の散布図を示す。

【図 1 2 D】2 型糖尿病を有する患者又は有しない患者の割合の積み重ね棒グラフを、それらが各遺伝子においてドライバー突然変異を保有するかどうかによって示す。

【図 1 2 E】NAFLD 活動性スコア (NAFLD Activity Score、NAS) が各遺伝子にドライバー突然変異を保有したか否かによる NAFLD 活動性スコア (NAS) の分布の積み重ね棒グラフを示し、低いスコアは、低い程度の組織学的異常を示す。

【図 1 3 A】患者 (x 軸) ごとのテロメア長 (y 軸) の分布の散布図を示す。各点は、その患者内の各クローンにおける最も高い中央値のバリエーション対立遺伝子画分 (variant allele fraction、VAF) を有する構成マイクロダイセクションについてのゲノム配列決定データから推定された平均テロメア長を表す。図 1 3 A では、患者が x 軸上に並べられ、疾患状態ごとに示されている (ブロック = 正常 / 健常、空白 = ARLD、線 = NAFLD) 。各疾患実体内で、患者は年齢の昇順で並べられている。図 1 3 B では、点は箱ひげ図上に重ねられ、ここで、中央値は太い黒線で示され、四分位数範囲は細い黒い箱で示される。ひげは、データの全範囲又は 2 5 / 7 5 パーセンタイル + 1 . 5 x 四分位数間範囲 (いずれか小さい方) の印を示す。

【図 1 3 B】患者 (x 軸) ごとのテロメア長 (y 軸) の分布の散布図を示す。各点は、その患者内の各クローンにおける最も高い中央値のバリエーション対立遺伝子画分 (variant allele fraction、VAF) を有する構成マイクロダイセクションについてのゲノム配列決定データから推定された平均テロメア長を表す。図 1 3 A では、患者が x 軸上に並べられ、疾患状態ごとに示されている (ブロック = 正常 / 健常、空白 = ARLD、線 = NAFLD) 。各疾患実体内で、患者は年齢の昇順で並べられている。図 1 3 B では、点は箱ひげ図上に重ねられ、ここで、中央値は太い黒線で示され、四分位数範囲は細い黒い箱で示される。ひげは、データの全範囲又は 2 5 / 7 5 パーセンタイル + 1 . 5 x 四分位数間範囲 (いずれか小さい方) の印を示す。

【図 1 4 A】それぞれ、ARLD (図 1 4 A) 及び NAFLD (図 1 4 B) からの 2 つの代表的な系統樹上に階層化されたテロメア長を示す。枝は、マイクロダイセクションのテロメア長に従ってグレースケールで網掛けされており、最も高い中央値のバリエーション対立

10

20

30

40

50

遺伝子画分がその枝に割り当てられている。内部ノードは最大尤度を用いて推定され、スケールは各枝に沿って補間された。

【図14B】それぞれ、ARLD（図14A）及びNAFLD（図14B）からの2つの代表的な系統樹上に階層化されたテロメア長を示す。枝は、マイクロダイセクションのテロメア長に従ってグレースケールで網掛けされており、最も高い中央値のバリエーション対立遺伝子画分がその枝に割り当てられている。内部ノードは最大尤度を用いて推定され、スケールは各枝に沿って補間された。

【図14C】テロメア長ごとに網掛けされた系統樹の更なる例を示す。テロメア長は、正常肝臓（上）、ARLD（中）及びNAFLD（下）からの2つの代表的な系統樹上に階層化されている。枝は、試料のテロメア長に従ってグレースケールで網掛けされており、最も高いバリエーション対立遺伝子画分がその枝に割り当てられている。内部ノードは最大尤度を用いて推定され、スケールは各枝に沿って補間された。

【図15】テロメア長に対するクローンサイズ（ \log_{10} 当たり（ μm^2 ））、年齢（10歳当たり）及び疾患状態（NAFLD及びARLD対正常）の効果サイズの事後分布を示す。MCMC（Markov Chain Monte Carlo）サンプラーからの密度プロットを十進法で網掛けして示す。

【図16A】本発明者らによって注目された新しい突然変異シグネチャーの詳細を示す。

【図16B】同じ肝臓試料内の近くのクローン間の活動の変動性が示されている。

【図16C】5年の期間にわたる1人の患者からの3つの肝臓切除試料を示す。

【図16D】5年の期間にわたる1人の患者からの3つの肝臓切除試料を示す。

【図16E】5年の期間にわたる1人の患者からの3つの肝臓切除試料を示す。

【図16F】5年の期間にわたる1人の患者からの3つの肝臓切除試料を示す。

【図16G】正常肝細胞、ARLD罹患細胞、NAFLD罹患細胞の試料における、及び8個全ての解剖学的セグメントがサンプリングされた2人のNAFLD患者におけるシグネチャーの分布が示されている。

【図16H】正常肝細胞、ARLD罹患細胞、NAFLD罹患細胞の試料における、及び8個全ての解剖学的セグメントがサンプリングされた2人のNAFLD患者におけるシグネチャーの分布が示されている。

【図16I】正常肝細胞、ARLD罹患細胞、NAFLD罹患細胞の試料における、及び8個全ての解剖学的セグメントがサンプリングされた2人のNAFLD患者におけるシグネチャーの分布が示されている。

【図17】シグネチャー抽出のために使用される濃度パラメータ設定を含む全体的なHDPノード構造を示す。

【発明を実施するための形態】

【0016】

本明細書に記載されるように、本発明者らは、対象における非アルコール性脂肪性肝疾患（NAFLD）又はアルコール関連脂肪性肝疾患（ARLD）を診断及び/又は予後判定するための方法であって、対象における1つ以上の肝細胞に選択的利点を付与する体細胞突然変異の発生を検出することを含む、方法を開発した。したがって、本発明は、対象におけるNAFLD又はAFLDを診断又は予後判定するための方法であって、

a) 対象の1つ以上の肝細胞に由来するDNA、RNA及び/又はタンパク質を含む生物学的試料を提供するステップと、

b) 肝細胞に選択的利点を付与するDNA、RNA及び/又はタンパク質における体細胞突然変異を検出するステップであって、体細胞突然変異の存在は、対象がNAFLD若しくはARLDに罹患しているか、NAFLD若しくはARLDを発症するリスクがあるか、より重症型の肝疾患を発症するリスクがあるか、及び/又は肝疾患に関連する疾患若しくは状態を発症するリスクがあることを示す、検出するステップと、を含む方法を提供する。特定の実施形態では、肝細胞に選択的利点を付与する体細胞突然変異の検出は、対象がNAFLD若しくはARLDに罹患しているか、NAFLD若しくはARLDを発症するリスクがあるか、及び/又は非アルコール性脂肪性肝炎（NASH）、肝線維症、肝

10

20

30

40

50

硬変、がん（例えば、肝細胞がん（HCC））、若しくは肝不全を発症するリスクがあることを示す。例示的な実施形態では、本明細書に記載のNAFLD又はARLDの診断又は予後判定方法は、NAFLD又はARLDのインビトロ診断又は予後判定方法である。

【0017】

誤解を避けるために、本明細書で使用される場合、「予後判定」という用語は、対象におけるNAFLD若しくはARLDの可能性のある経過及び転帰、並びに/又は対象がNAFLD若しくはARLDから回復する可能性を推定/予測するプロセスを指す。例えば、本発明の方法を使用することによって決定されるように、肝疾患が特定の治療に应答して退行していない対象は、予後不良であると見なされ得る。

【0018】

a) 対象の1つ以上の肝細胞に由来するDNA、RNA及び/又はタンパク質を含む生物学的試料を提供すること：

典型的には、1つ以上の肝細胞に由来するDNA、RNA及び/又はタンパク質を含む生物学的試料は、対象から得られた血液試料、尿試料又は組織試料から得られる。血液試料は、対象から得られた血清、血漿又は全血試料であってもよい。組織試料は、生検試料、特に対象から得られた肝生検試料であってもよい。肝生検試料は、肝細胞（hepatocyte、HC）、肝星細胞（hepatic stellate cell、HSC）、クッパー細胞（Kupffer cell、KC）又は肝類洞壁内皮細胞（liver sinusoidal endothelial cell、LSEC）などの肝細胞を含む。好ましくは、対象から得られた肝生検試料は、肝細胞を含む。

【0019】

典型的には、生物学的試料は、少なくとも10個、少なくとも100個、又は少なくとも1000個、少なくとも10,000個、少なくとも 10^4 個、少なくとも 10^5 個、又は少なくとも 10^6 個の肝細胞に由来するDNA、RNA及び/又はタンパク質を含む。例えば、約10～約10⁶個の肝細胞、約100～約 10^5 個、約500～約 10^4 個、約10～約50,000個、又は約1000～約50,000個（例えば、約1000～約10,000個、約1000～約20,000個、約1000～約30,000個、若しくは約1000～約40,000個の肝細胞）である。例示的な実施形態では、生物学的試料は、対象の約10～約50,000個の肝細胞（例えば、約10～約50,000個の肝細胞）に由来するDNA、RNA及び/又はタンパク質を含む。

【0020】

試料が1つ以上の肝細胞に由来するDNA及び/又はRNAを含む特定の実施形態では、試料は、ゲノムDNA（genomic DNA、gDNA）及び/又はメッセンジャーRNA（messenger RNA、mRNA）を含む。例えば、試料は、対象から得られた血液試料又は生検試料中に存在する細胞から抽出されたDNA又はRNAを含み得る。

【0021】

ある特定の実施形態では、DNA、RNA及び/又はタンパク質を含む生物学的試料は、対象からの肝生検試料に由来する。対象から肝生検を得るための方法は、当技術分野において既知であり、臨床従事者の日常的な能力の範囲内である。肝生検試料は、例えば、経皮的、経頸静脈的又は腹腔鏡下肝生検法によって対象から得てもよい。

【0022】

例示的な実施形態では、DNA、RNA及び/又はタンパク質を含む生物学的試料は、対象から得られた肝生検試料のマイクロダイセクションに由来する。例えば、肝生検試料を対象から得て、その後、約1～約100（例えば、10、20、30、40、50、60、70、80、90又は100）の別個のマイクロダイセクション試料に切断してもよい。典型的には、肝生検試料は、約10～約60個のマイクロダイセクション試料に切断される。肝生検試料から得られた各別個のマイクロダイセクション試料は、約10～約50,000個の肝細胞（10～約50,000個の肝細胞）を含み得る。典型的には、各別個のマイクロダイセクション試料は、約100～約500個の肝細胞を含む。

【0023】

10

20

30

40

50

肝細胞などの細胞からDNA及び/又はRNAを抽出するための方法は、当技術分野において既知である。例えば、DNA及びRNAは、細胞を溶解するための試薬を使用して細胞又は組織試料から単離することができ、その後、カラムベースの手法及び/又はビーズベースの手法を行って、DNA及び/又はRNAを精製することができる。細胞又は組織試料からDNA又はRNAを単離するために使用され得る市販キットの例としては、QIAamp（登録商標）DNA Kit、EZ1（登録商標）DNA Tissue Kit、Oligotex（登録商標）Direct mRNA Kit、及びArcturus（登録商標）PicoPure（登録商標）DNA Extraction Kitが挙げられるが、これらに限定されない。

【0024】

単離されたDNA又はRNAは、分析前に増幅され得る。増幅の前に、単離されたRNAは、逆転写酵素を使用する逆転写によってDNAに変換され得る。DNA又はRNAは、当該分野で既知の技術を使用して、例えば、ポリメラーゼ連鎖反応（polymerase chain reaction、PCR）ベースの方法又は逆転写ポリメラーゼ連鎖反応（reverse transcription polymerase chain reaction、RT-PCR）ベースの方法を使用することによって増幅され得る。単離されたDNA又はRNAはまた、使用される配列決定技術に好適なDNAライブラリー又はRNAライブラリーを構築するために使用され得る。例えば、トランスポザーゼベースの方法を用いてDNAライブラリーを構築してもよい。

【0025】

試料から単離されたDNA又はmRNAは、定量的PCR（Quantitative PCR、qPCR）及び定量的逆転写PCR（Quantitative reverse transcription PCR、RT-qPCR）などの当技術分野で既知の方法を使用して定量してもよい。したがって、ある特定の実施形態では、本方法は、対象から得られた生物学的試料から得られたDNA又はRNAの量を増幅及び/又は定量するステップを更に含む得る。

【0026】

生物学的試料がDNAを含む実施形態では、DNAは循環遊離DNA（circulating free DNA、cfDNA）であり得る。すなわち、本発明の方法のステップa）において提供される生物学的試料は、対象の1つ以上の肝細胞に由来するcfDNAを含み得る。本明細書で使用される場合、「循環遊離DNA」（cfDNA）という用語は、血漿中に放出されており、血流中又は尿中を自由に循環していることが見出されるDNA断片を指す。cfDNAは、一般に、小さな断片（70～200bp）からなる二本鎖DNAである。したがって、生物学的試料は、cfDNAを検出及び/又は単離することができる、当技術分野で既知の任意の好適な試料であってもよい。例えば、試料は、好適には、cfDNAを含む血液試料、血漿試料、又は尿試料であり得る。生物学的試料中の肝臓由来cfDNAを同定するための方法は当技術分野で既知であり、そのような方法の例は、例えば、Jiang et al., J Hepatology, 2019, 71(2), 409-421, Moss et al., Nat Commun, 2018, 9, 5068, 及びPunia et al., BMC Gastroenterol, 2021, 21, 149-159に記載されている。

【0027】

生物学的試料がcfDNAを含む実施形態では、本方法は、試料からcfDNAを単離することを更に含む得る。cfDNAは、当技術分野で既知の様々な技法を使用して試料から単離することができる。例えば、cfDNAは、カラムベースの手法及び/又はビーズベースの手法によって単離することができる。いくつかの実施形態では、cfDNAは、例えば、QIAamp（登録商標）循環核酸キットなどの市販のキットを使用して、カラムベースの手法によって単離される。いくつかの実施形態では、cfDNAは、ビーズベースの手法、例えば、Maxwell RSC cfDNA Plasma Kit（Promega）などの市販のキットを使用する自動cfDNA抽出システムによって単離される。

【0028】

10

20

30

40

50

単離された c f D N A は、分析前に増幅され得る。したがって、本方法は、単離された c f D N A の増幅を更に含み得る。c f D N A を増幅するのに好適な技術は当技術分野で既知であり、クローニング、ポリメラーゼ連鎖反応 (P C R)、特異的対立遺伝子のポリメラーゼ連鎖反応 (polymerase chain reaction of specific alleles、P A S A)、ポリメラーゼ鎖ライゲーション、ネステッドポリメラーゼ連鎖反応などが挙げられるが、これらに限定されない。

【 0 0 2 9 】

生物学的試料がタンパク質を含む実施形態では、突然変異タンパク質のレベルは、当技術分野で既知の任意の好適な技術を使用して測定することができる。本方法は、試料からタンパク質を単離することと、次いで、存在する突然変異タンパク質の量を評価すること
10
と、を含み得る。使用される方法及び必要とされる精度のレベルに依存して、精製ステップが実施され得る。いくつかの状況において、試料中の細胞の単純な溶解が十分であり得る。存在する突然変異タンパク質のレベルは、酵素結合免疫吸着アッセイ (enzyme-linked immunosorbent assay、E L I S A)、ウェスタンブロット分析及び質量分析から選択される 1 つ以上の技術を使用して評価され得る。

【 0 0 3 0 】

本明細書で使用される場合、「対象」とは、ヒトなどの哺乳動物を含む動物を指す。好ましくは、患者はヒト対象である。実施形態では、対象は、N A F L D 若しくは A R L D を有することが知られているか若しくはその疑いがあるか、及び / 又は N A F L D 若しくは A R L D を発症するリスクを有することが知られているか若しくはその疑いがある。あ
20
る特定の実施形態では、対象は、N A F L D 若しくは A R L D を発症するリスクを増加させること、より重症型の肝疾患を発症するリスクを増加させること、及び / 又は肝疾患に関連する疾患若しくは状態を発症するリスクを増加させることが知られている疾患又は状態を有する対象である。例えば、対象は、消化管がん、肥満、2 型糖尿病、高血圧、脂質異常症 (例えば、高コレステロール血症) 及び / 又は心血管疾患に罹患していることが知られているか、又はその疑いがある対象であってもよい。

【 0 0 3 1 】

b) 肝細胞クローンに選択的利点を付与する D N A、R N A 及び / 又はタンパク質における体細胞突然変異を検出すること :

体細胞突然変異は、個体の生涯の間に獲得される遺伝子の D N A 配列の永久的な変化である。すなわち、体細胞突然変異は個体の生殖系列 D N A には存在せず、したがって生殖系列多型のように親から遺伝されない。体細胞突然変異は、ヌクレオチドの置換、欠失又は細胞の D N A への挿入を作り出す各細胞分裂で起こる D N A 複製の非忠実度に起因して自発的に起こり得る。体細胞突然変異は、紫外線放射、化学物質曝露又はウイルス感染などの環境因子によっても引き起こされ得る。
30

【 0 0 3 2 】

遺伝子における体細胞突然変異は、突然変異タンパク質を産生するために細胞によって発現され得、遺伝子における 1 つ以上のヌクレオチド置換 (すなわち、「単一ヌクレオチドバリエーション (single-nucleotide variant) 」又は「 S N V 」である各ヌクレオチド置換) は、体細胞非突然変異核酸配列によってコードされるアミノ酸と比較して、コード
40
される異なるアミノ酸をもたらす。したがって、正常な非罹患細胞 (例えば、健康な肝細胞) におけるタンパク質 / ペプチドと比較して、タンパク質 / ペプチドにおいて異なるアミノ酸をもたらす可能性がある。ヌクレオチドの挿入及び / 又は欠失は、リーディングフレームエラー (すなわち、「インデル突然変異」又は「フレームシフト突然変異」) をもたらす。その結果として、タンパク質レベルで新しいアミノ酸配列をもたらす (すなわち、ヌクレオチドの挿入又は欠失が D N A のリーディングフレームを変化させ、その結果、突然変異後に、正常細胞 (例えば、健康な肝細胞) と比較して、D N A によってコードされるアミノ酸の大部分又は全てを変化させる) 可能性がある。追加的又は代替的に、挿入及び / 又は欠失は、終止コドンの導入をもたらす。その結果として、タンパク質レベルで短縮型タンパク質をもたらす (ナンセンス突然変異) 可能性がある。ヌクレオチド置換は
50

、コドンをお個々に変更し、タンパク質レベルでアミノ酸の置換及び／又は終止コドンの導入をもたらし、その結果として、タンパク質レベルで短縮型タンパク質をもたらす可能性がある。

【0033】

本発明の方法のステップb)において検出される体細胞突然変異は、対象の肝細胞に選択的利点を付与するものである。本明細書で使用される場合、「選択的利点」という用語は、同じ1つ以上の突然変異を有しない細胞よりも良好に細胞が生存及び／又は再生することを可能にする1つ以上の突然変異(例えば、体細胞突然変異は、対象の肝細胞に増殖利点を付与するものである)を介して所与の細胞に付与される利点を指す。例えば、肝細胞に選択的利点を付与する体細胞突然変異は、細胞内の脂質の蓄積によって引き起こされる毒性条件下で肝細胞が生存することを可能にするものであり得る。肝細胞に選択的利点を付与する突然変異は、肝臓の所与の微小環境における肝細胞の正の選択をもたらし、同じ1つ以上の突然変異を有しない肝細胞よりも良好に生存及び／又は再生することができる肝細胞を含む優性肝細胞クローンをもたらす可能性がある。したがって、例えば、ある特定の実施形態では、本発明の方法のステップb)において検出される体細胞突然変異は、肝臓における脂質蓄積に関連する毒性条件下で肝細胞が生存することを可能にすることによって、肝細胞に選択的利点を付与するものである。更に、又はあるいは、肝細胞に選択的利点を付与する体細胞突然変異は、肝臓における脂質蓄積に関連するもの、例えば、肝臓における脂質蓄積の増加及び／又は肝臓における脂質代謝の減少に関連するものであり得る。

10

20

【0034】

誤解を避けるために、本明細書で使用される場合、「ドライパー突然変異」という用語はまた、対象の細胞(例えば、肝細胞)に選択的利点を付与する変異を指す。

【0035】

肝細胞などの細胞に選択的利点を付与する体細胞突然変異を同定するための方法は、当技術分野で既知である。Martincorena, I. et al., Cell 2017, 171, 1029-1041.e21によって記載されたdN/DScv法及びReinbay, E. et al., Nature, 2020, 578, 102-111によって記載されたNBRアルゴリズムの使用を含む例示的な方法が本明細書に記載されている。

30

【0036】

本発明者らは、対象の肝臓に由来する試料中の複数の肝細胞において選択的利点を付与する独立して獲得された体細胞突然変異の検出が、対象におけるNAFLD又はARLDを診断及び／又は予後診断するための特に信頼できる方法を提供することを見出した。したがって、ある特定の実施形態では、本発明の方法のステップa)は、少なくとも10個、少なくとも100個、又は少なくとも1000個、少なくとも 10^5 個、又は少なくとも 10^6 個の肝細胞に由来するDNA、RNA及び／又はタンパク質を含む生物学的試料を提供することを含む。例えば、ステップb)は、少なくとも100個、少なくとも300個、少なくとも400個、少なくとも500個、少なくとも1000個の肝細胞を含む生物学的試料中の1つ以上の体細胞突然変異を検出することを含み得る。そのような実施形態では、ステップb)は、少なくとも10個、少なくとも100個、又は少なくとも1000個、少なくとも 10^5 個、又は少なくとも 10^6 個の肝細胞に由来する当該生物学的試料中の1つ以上の体細胞突然変異を検出することを含み、1つ以上の肝細胞に選択的利点を付与する1つ以上の体細胞突然変異の存在は、対象がNAFLD若しくはARLDに罹患しているか、NAFLD及び／若しくはARLDを発症するリスクがあるか、より重症型の肝疾患を発症するリスクがあるか、及び／又は肝疾患に関連する疾患若しくは状態を発症するリスクがあることを示す。

40

【0037】

本発明者らは、約1000~約50,000個の肝細胞(例えば、肝細胞)に由来するDNA、RNA及び／又はタンパク質を含む生物学的試料の分析が、対象の肝臓の遺伝的

50

ランドスケープ、並びにノ又は肝臓におけるNAFLD若しくはARLDの進行及びノ若しくは重症度の特に信頼できるスナップショットを提供することを見出した。したがって、ある特定の実施形態では、本発明の方法のステップa)は、対象の約1000~約50,000個の肝細胞(例えば、約1000~約50,000個の肝細胞)に由来するDNA、RNA及びノ又はタンパク質を含む生物学的試料を提供することを含む。

【0038】

更に、又はあるいは、本発明の方法のステップb)は、対象から得られた1つ以上の異なる生物学的試料を使用して繰り返され得る。例えば、1つ以上の肝生検試料又はそのマイクロダイセクション試料、1つ以上の血液試料及びノ又は1つ以上の尿試料である。例えば、ステップb)は、対象の肝臓における異なる位置からの肝生検から得られた対象の肝細胞に由来するDNA、RNA及びノ又はタンパク質を含む1つ以上の異なる生物学的試料(例えば、肝臓の異なる葉から得られた肝生検試料)を用いて繰り返されてもよい。

10

【0039】

更に、又はあるいは、本発明の方法のステップb)は、対象由来の1つの肝生検試料から得られた1つ以上の異なる生物学的試料を使用して繰り返されてもよい。例えば、ステップb)は、対象由来の肝生検試料を2つ以上の別個のマイクロダイセクション試料にマイクロダイセクションすることによって得られた1つ以上の生物学的試料を用いて繰り返されてもよい。

【0040】

1つ以上の異なる肝生検、その1つ以上の異なるマイクロダイセクション試料、1つ以上の異なる血液試料、及びノ又は1つ以上の異なる尿試料などの1つ以上の異なる生物学的試料の分析は、対象の肝臓の遺伝的ランドスケープ、並びにノ又は肝臓全体にわたるNAFLD若しくはARLDの進行及びノ若しくは重症度へのより広い洞察を提供する。したがって、ある特定の実施形態では、本発明の方法は、対象の肝細胞に由来するDNA、RNA及びノ又はタンパク質を含む1つ以上の異なる生物学的試料を提供することと、対象がNAFLD若しくはARLDに罹患しているか、NAFLD及びノ若しくはARLDを発症するリスクがあるか、より重症型の肝疾患を発症するリスクがあるか、及びノ又は肝疾患に関連する疾患若しくは状態を発症するリスクがあるかを決定するために、1つ以上の異なる生物学的試料を用いてステップb)を繰り返すことと、を含む。

20

【0041】

ある特定の実施形態では、ステップb)において検出される、対象の肝細胞に選択的利点を付与する体細胞突然変異は、FOXO1、GPAM、CLDEB、ACVR2A、ALB及びTNRC6Bからなる群から選択される遺伝子内にある。

30

【0042】

したがって、ある特定の実施形態では、本発明の方法のステップb)は、対象の肝細胞に選択的利点を付与するFOXO1遺伝子内の1つ以上の体細胞突然変異を検出することを含む。FOXO1は、インスリンシグナル伝達の下流の主要な転写因子をコードする。絶食状態では、インスリンなしで、FOXO1タンパク質は肝細胞の核において活性であり、糖新生、解糖及び脂肪分解経路における遺伝子の発現を上方制御する。インスリンがその受容体に結合すると、AKTはPI3Kを介して活性化される。続いて、AKTは核内のFOXO1タンパク質をリン酸化し、FOXO1タンパク質の24位のスレオニンは3つの既知のAKTリン酸化標的のうちの1つである。

40

【0043】

NAFLD又はARLDに罹患している対象において、本発明者らは、FOXO1タンパク質のインスリン媒介核外輸送を損なうFOXO1遺伝子内の体細胞突然変異を同定した。FOXO1機能の障害は、肝内脂肪が増加した対象において非常に優勢であるインスリン感受性の低下に寄与する。FOXO1遺伝子におけるこれらの突然変異のうちの1つ以上の存在は、対象がNAFLD若しくはARLDに罹患していること、又は対象のNAFLD若しくはARLDを発症するリスクが増加しているか、より重症型の肝疾患を発症するリスクが増加しているか、及びノ又は肝疾患に関連する疾患若しくは状態を発症する

50

リスクが増加していることを示し得る。したがって、ある特定の実施形態では、本方法は、FOXO1タンパク質のインスリン媒介核外輸送を損なうFOXO1タンパク質のアミノ酸突然変異をもたらすFOXO1遺伝子内の1つ以上の体細胞突然変異の検出を含む。

【0044】

典型的には、FOXO1遺伝子において検出される体細胞突然変異のうちの少なくとも1つは、FOXO1タンパク質のN末端14-3-3タンパク質結合モチーフをコードするFOXO1遺伝子の領域(すなわち、FOXO1タンパク質のアミノ酸20~28)中にある。本発明者らは、FOXO1タンパク質のN末端14-3-3タンパク質結合モチーフ内の体細胞突然変異が、対象におけるNAFLD又はARLDを特に示すことを見出した。したがって、好ましい実施形態では、本発明の方法のステップb)は、FOXO1タンパク質のN末端14-3-3タンパク質結合モチーフ内のアミノ酸突然変異をもたらすFOXO1遺伝子内の1つ以上の体細胞突然変異を検出することを含む。

10

【0045】

好ましい実施形態では、本発明の方法のステップb)は、FOXO1タンパク質においてS22Wアミノ酸置換、R21Lアミノ酸置換及び/又はS22ナンセンス突然変異(本明細書ではS22*と呼ぶ)をもたらす1つ以上の体細胞突然変異を検出することを含む。本発明者らは、FOXO1タンパク質におけるS22Wアミノ酸置換、R21Lアミノ酸置換又はS22ナンセンス突然変異をもたらすFOXO1遺伝子における体細胞突然変異が、対象におけるNAFLD又はARLDを特に示すことを見出した。したがって、特に好ましい実施形態では、本発明の方法のステップb)は、FOXO1タンパク質においてS22Wアミノ酸置換をもたらすFOXO1遺伝子内の1つ以上の体細胞突然変異の検出を含む。

20

【0046】

本発明者らはまた、GPAM遺伝子内の体細胞突然変異は、対象におけるNAFLD若しくはARLDの指標、対象がNAFLD若しくはARLDを発症するリスクの増加の更なる指標、及び/又は対象がより重症型の肝疾患若しくは関連疾患若しくは状態を発症するリスクの増加の指標として使用され得ることを見出した。GPAM遺伝子は、グリセロール-3-リン酸アシルトランスフェラーゼ1、ミトコンドリアタンパク質(本明細書ではGPAMタンパク質又はGPATタンパク質と呼ぶ)をコードする。GPAMタンパク質は、グリセロール-3-リン酸による長鎖アシル-CoAのエステル化を触媒する酵素である。典型的には、GPAM遺伝子において検出される1つ以上の体細胞突然変異は、GPAMタンパク質の機能障害又は機能無効、特にグリセロール-3-リン酸による長鎖アシル-CoAのエステル化を触媒するGPAMタンパク質の能力の障害又は無効をもたらす。本発明者らは、GPAMタンパク質の機能障害又は機能無効につながるGPAM遺伝子における体細胞突然変異が、NAFLD又はARLDの診断及び/又は予後のための特に有用なマーカーであることを見出した。

30

【0047】

したがって、ある特定の実施形態では、本発明の方法のステップb)は、対象の肝細胞に選択的利点を付与するGPAM遺伝子内の1つ以上の体細胞突然変異を検出することを含む。ある特定の好ましい実施形態では、GPAM遺伝子において検出される1つ以上の体細胞突然変異は、GPAMタンパク質の機能を損なうか又は無効にする。本発明の方法のステップb)において検出され得るGPAM遺伝子内の体細胞突然変異の例としては、GPAMタンパク質のG790W、E730K、A619P、R519G、H486R、L332S、F313L、Y292C、G273V、L225M及び/又はQ210R置換、GPAMタンパク質のL322フレームシフト突然変異又はR118(すなわち、R118*)ナンセンス突然変異をもたらす突然変異が挙げられる。

40

【0048】

本発明者らはまた、CIDEB遺伝子内の体細胞突然変異が、対象におけるNAFLD若しくはARLDの指標、対象がNAFLD若しくはARLDを発症するリスクの増加の指標、及び/又は対象がより重症型の肝疾患若しくは関連する疾患若しくは状態を発症す

50

るリスクの増加の更なる指標として使用され得ることを見出した。CIDE Bは、肝細胞において活性なCIDEファミリーの主要なメンバーであり、ノックアウトマウスモデルは、食事性脂肪性肝炎に対する抵抗性及びインスリン感受性の増加を示す(Liet al., Diabetes, 2007, 56, 2523-2532)。CIDEタンパク質は、異なる液滴上のCIDEタンパク質間のホモ二量体の形成によって媒介される細胞内脂質液滴の融合を調節する(Barneda et al. Elife, 2015, 4, 1-24、及びSun et al. Nat. Commun. 2013, 4, 15949)。CIDEタンパク質のホモ二量体化は、一方の脂肪液滴からのCIDEタンパク質上の正に荷電した残基と他方の負に荷電した残基との間の静電接触を介して起こる。

【0049】

本発明者らは、NAFLD又はARLDに罹患している患者のCIDE B遺伝子におけるナンセンス、終止欠失突然変異及びミスセンス突然変異を同定した。特に、本発明者らは、ミスセンス突然変異がCIDEタンパク質のホモ二量体化に関与する2つのドメインに主に位置し、それらの多くが荷電残基を中性残基に切り替える(R45W、R45Q、K140N、R144P)か、又は電荷を逆転させる(D42H、K62E、E78K)かのいずれかであることを見出した。以前のインビトロ突然変異誘発研究は、荷電残基での置換又はCIDE Bタンパクの短縮を引き起こす突然変異が、肝細胞における脂肪液滴の増殖を妨害することを示している(Barneda et al. Elife, 2015, 4, 1-24、及びSun et al. Nat. Commun. 2013, 4, 15949)。CIDE Bノックアウトマウスはまた、高脂肪食によって引き起こされる脂肪性肝炎に対して抵抗性であることが見出されており(Liet al. Diabetes, 2007, 56, 2523-2532)、したがって、CIDE B突然変異を不活性化させることが、代謝肝疾患において肝細胞に選択的利点をどのように付与し得るかについてのインビボでの証拠を提供する。

【0050】

したがって、ある特定の実施形態では、本発明の方法のステップb)は、対象の肝細胞に選択的利点を付与するCIDE B遺伝子内の1つ以上の体細胞突然変異を検出することを含む。ある特定の好ましい実施形態では、CIDE B遺伝子において検出される1つ以上の体細胞突然変異は、対象の肝細胞によって発現されるCIDE Bタンパクの機能を損なうか又は無効にする。例えば、CIDE B遺伝子内に存在する体細胞突然変異は、例えば、CIDE Bタンパク質中の荷電した残基の中性残基による置換、CIDE Bタンパク質中の中性残基の荷電した残基による置換、CIDE Bタンパク質中の負に荷電した残基の正に荷電した残基による置換、及び/又はCIDE Bタンパク質中の正に荷電した残基の負に荷電した残基による置換によって、CIDE Bタンパク質の電荷の変化をもたらすものであってもよい。例示的な実施形態では、CIDE B遺伝子における1つ以上の体細胞突然変異は、CIDE Bタンパクのホモ二量体化を損なうか又は無効にし、したがって、肝細胞内の脂肪液滴の融合及び増殖を防止する。本発明の方法のステップb)において検出され得るCIDE B遺伝子における体細胞突然変異の例としては、CIDE BタンパクのL4H、L7Q、F23S、W28R、D42H、R45Q、R45W、K62E、E72K、R122Q、I131V、A123P、K140N、R144P、L150Q、N151D、D165E及び/若しくはQ167E置換、又はCIDE BタンパクのW28(すなわちW28*)、W181(すなわちW181*)及び/若しくは220Y(すなわち*220Y)ナンセンス突然変異をもたらす突然変異が挙げられる。

【0051】

本発明者らはまた、ACVR2A遺伝子内の体細胞突然変異の検出が、対象におけるNAFLD若しくはARLDの指標、対象がNAFLD若しくはARLDを発症するリスクの増加の指標、及び/又は対象がより重症型の肝疾患若しくは関連する疾患若しくは状態を発症するリスクの増加の指標として使用され得ることを見出した。ACVR2A遺伝子は、TGF-スーパーファミリーにおけるアクチビン-Aの受容体をコードする。本発明者らは、NAFLD又はARLDに罹患している患者において、ACVR2Aにおける

10

20

30

40

50

13個のミスセンス突然変異、2個のナンセンス及び1個のスプライス部位インデル ($q = 7 \times 10^{-9}$)、並びに構造変異による4個の大規模欠失を同定した。

【0052】

したがって、ある特定の実施形態では、本発明の方法のステップb)は、対象の肝細胞に選択的利点を付与するACVR2A遺伝子内の1つ以上の体細胞突然変異を検出することを含む。

【0053】

本発明の方法のステップb)において検出され得るACVR2A遺伝子における体細胞突然変異の例としては、ACVR2Aタンパク質のC30R、L31P、C85Y、C105S、Y236C、M241V、V283G、T295A、H320Y、1375V、S424Y及びQ481P置換、並びにACVR2Aタンパク質のR41(すなわち、R41*)又はE234(すなわち、E234*)ナンセンス突然変異をもたらす突然変異が挙げられる。

10

【0054】

本発明者らはまた、ALB遺伝子内の体細胞突然変異の検出が、対象におけるNAFLD若しくはARLDの指標、対象がNAFLD若しくはARLDを発症するリスクの増加の指標、及び/又は対象がより重症型の肝疾患若しくは関連する疾患若しくは状態を発症するリスクの増加の指標として使用され得ることを見出した。ALB遺伝子は、プレプロアルブミンとして肝臓で合成される血清アルブミンをコードする。新生形態のプレプロアルブミンはプロセシングされ、血清アルブミンとして循環中に放出される。

20

【0055】

したがって、ある特定の実施形態では、本発明の方法のステップb)は、対象の肝細胞に選択的利点を付与するALB遺伝子内の1つ以上の体細胞突然変異を検出することを含む。ある特定の好ましい実施形態では、ALB遺伝子内の体細胞突然変異は、例えば、対象から得られた血液試料中に存在する血清アルブミンを特徴付けることによって、タンパク質レベルで検出してもよい。例えば、血液試料中の突然変異血清アルブミンの存在及び/又はレベルは、例えば、酵素結合免疫吸着アッセイ(ELISA)、ウェスタンブロット分析及び質量分析によって決定され得る。

【0056】

本発明者らはまた、TNRC6B遺伝子内の体細胞突然変異が、対象におけるNAFLD若しくはARLDの更なる指標、対象がNAFLD若しくはARLDを発症するリスクの増加の更なる指標、及び/又は対象がより重症型の肝疾患若しくは関連する疾患若しくは状態を発症するリスクの増加の更なる指標として使用され得ることを見出した。TNRC6Bは、マイクロRNAプロセシングに關与するタンパクをコードする(Meister et al., Curr Biol, 2005, 6; 15(23): 2149-55)。本発明者らは、NAFLD又はARLDに罹患している患者におけるTNRC6B遺伝子内の体細胞突然変異を同定した。特に、本発明者らは、NAFLD又はARLDに罹患している患者のTNRC6B遺伝子において、3つのナンセンス、2つの必須スプライス部位及び1つの大きなインフレーム欠失、並びに3つのミスセンス突然変異を見出した。

30

40

【0057】

したがって、ある特定の実施形態では、本発明の方法のステップb)は、対象の肝細胞に選択的利点を付与するTNRC6B遺伝子内の1つ以上の体細胞突然変異を検出することを含む。ある特定の好ましい実施形態では、TNRC6B遺伝子内の1つ以上の体細胞突然変異は、TNRC6Bタンパク質の機能を損なうか又は無効にする。

【0058】

本発明の方法のステップb)において検出され得るTNRC6B遺伝子における体細胞突然変異の例としては、TNRC6Bタンパク質のG1374C、T1535S及び/若しくはM1814V置換、TNRC6Bタンパク質のG536(すなわちG536*)、W1399(すなわちW1399*)若しくはQ1700(すなわちQ1700*)ナン

50

センス突然変異、並びに / 又は T N R C 6 B タンパク質の残基 1 6 3 ~ 1 8 0 のインフレーム欠失をもたらす突然変異が挙げられる。

【 0 0 5 9 】

ある特定の例示的な実施形態では、本発明の方法のステップ b) は、 F O X O 1、 G P A M 及び / 又は C I D E B 遺伝子内の 1 つ以上の体細胞突然変異を検出することを含む。すなわち、少なくとも 1 つの体細胞突然変異が F O X O 1 遺伝子において検出され、少なくとも 1 つの体細胞突然変異が G P A M 遺伝子において検出され、かつ / 又は少なくとも 1 つの体細胞突然変異が C I D E B 遺伝子において検出される。例えば、本方法は、 F O X O 1 タンパク質の 1 4 - 3 - 3 タンパク質結合モチーフをコードする F O X O 1 遺伝子の D N A の領域における F O X O 1 遺伝子中の 1 つ以上の体細胞突然変異の検出、機能障害又は機能無効を示す G P A M タンパク質をコードする G P A M 遺伝子中の 1 つ以上の突然変異の検出、及び / 又は機能障害又は機能無効を示す C I D E B タンパク質をコードする C I D E B 遺伝子中の 1 つ以上の突然変異の検出を含み得る。

10

【 0 0 6 0 】

対象の 1 つ以上の肝細胞に選択的利点を付与する体細胞突然変異を検出するステップ (例えば、 F O X O 1、 G P A M、 C I D E B、 A C V R 2 A、 A L B 及び T N R C 6 B からなる群から選択される遺伝子における 1 つ以上の体細胞突然変異の検出) は、本明細書に記載されるように、対象から得られた生物学的試料から得られた D N A を配列決定することによって達成され得る。更に、又はあるいは、 1 つ以上の体細胞突然変異は、 1 つ以上の体細胞突然変異を含む遺伝子から転写される生物学的試料中の m R N A を配列決定及び / 又は定量することによって検出され得る。例えば、 1 つ以上の体細胞突然変異は、 1 つ以上の体細胞突然変異を含む F O X O 1、 G P A M、 C I D E B、 A C V R 2 A、 A L B 又は T N R C 6 B 遺伝子から転写された生物学的試料中の m R N A を配列決定及び / 又は定量することによって検出され得る。

20

【 0 0 6 1 】

種々の D N A 及び R N A 配列決定手順が当該分野で既知であり、そして本明細書中に開示される方法を実施するために使用され得る。例えば、 S a n g e r 配列決定、 P o l o n y 配列決定、 4 5 4 パイロシーケンシング、コンビナトリアルプローブアンカー合成、 S O L i D 配列決定、 I o n T o r r e n t 半導体配列決定、 D N A ナノボール配列決定、 H e l i s c o p e 単一分子配列決定、単一分子リアルタイム (S i n g l e m o l e c u l e r e a l t i m e、 S M R T) 配列決定、 N a n o p o r e D N A 配列決定、 M i c r o f l u i d i c S a n g e r 配列決定及び I l l u m i n a 色素配列決定である。

30

【 0 0 6 2 】

更に、又はあるいは、 1 つ以上の体細胞突然変異は、 1 つ以上の体細胞突然変異を含む遺伝子によってコードされる突然変異アミノ酸を含むタンパク (例えば、 F O X O 1、 G P A M、 C I D E B、 A C V R 2 A、 A L B 又は T N R C 6 B タンパク) を検出することによって検出及び / 又は定量され得る。試料中の変異タンパク質の存在及び / 又はレベルは、当該分野で既知の任意の好適な技術を使用して測定され得る。本方法は、試料からタンパク質を単離することと、次いで、存在する突然変異タンパク質の量を評価することと、を含み得る。使用される方法及び必要とされる精度のレベルに依存して、精製ステップが実施され得る。いくつかの状況において、試料中の細胞の単純な溶解が十分であり得る。存在する突然変異タンパク質のレベルは、酵素結合免疫吸着アッセイ (E L I S A)、ウェスタンブロット分析及び質量分析から選択される 1 つ以上の技術を使用して評価され得る。

40

【 0 0 6 3 】

誤解を避けるために、本明細書で使用される場合、「重症型の肝疾患」という用語は、 N A F L D 又は A R L D の結果としての進行性肝線維症又は肝臓関連機能不全、例えば、非アルコール性脂肪性肝炎 (N A S H)、肝線維症、肝硬変、肝臓がん、及び肝不全を指す。具体的なタイプの肝臓がんとしては、肝細胞がん (H C C)、肝内胆管がん、脈管肉腫及び血管肉腫が挙げられる。

50

【0064】

誤解を避けるために、本明細書で使用される場合、「関連疾患又は状態」という用語は、肝疾患に関連する疾患及び状態を指す。このような疾患及び状態としては、食道、胃、胆管、胆嚢及び関連構造、膵臓、小腸、大腸、直腸及び肛門の消化管がんが挙げられる。肝疾患に関連する疾患及び状態の他の例としては、肥満、2型糖尿病、高血圧、脂質異常症（例えば、高コレステロール血症）及び心血管疾患が挙げられる。したがって、ある特定の実施形態では、本発明の方法のステップb)は、対象がNAFLD若しくはARLDに罹患しているか、NAFLD若しくはARLDを発症するリスクがあるか、及び/又は非アルコール性脂肪性肝炎（NAFLD）、肝線維症、肝硬変、がん（例えば、肝細胞がん（HCC））及び肝不全からなる群から選択されるより重症型の肝疾患を発症するリスクがあるかどうかを決定することを含む。特定の実施形態では、本発明の方法のステップb)は、対象が、消化管がん、肥満、2型糖尿病、高血圧、脂質異常症（例えば、高コレステロール血症）及び心血管疾患からなる群から選択される肝疾患に関連する疾患又は状態を発症するリスクがあるかどうかを決定することを含む。

10

【0065】

ある特定の実施形態では、本発明の方法のステップb)は、ステップa)において提供される生物学的試料中の1つ以上の体細胞突然変異のレベルを定量することを含む。1つ以上の体細胞突然変異のレベルを定量するのに適した技術は、当技術分野で既知であり、例えば、qPCR、RT-qPCR、DNA/RNAマイクロアレイの使用、及びDNA配列決定データ、特に次世代配列決定（例えば、Illumina色素配列決定及びIon Torrent半導体配列決定）データの分析が挙げられる。

20

【0066】

本明細書で使用される場合、「1つ以上の体細胞突然変異のレベル」及び「体細胞突然変異のレベル」という用語は、所与の生物学的試料中の1つ以上の体細胞突然変異のコピー数、所与の生物学的試料中の1つ以上の体細胞突然変異を含むDNA分子の濃度、所与の生物学的試料中の1つ以上の体細胞突然変異を含むmRNA分子の濃度、又は所与の生物学的試料中の1つ以上の体細胞突然変異によってコードされる1つ以上の突然変異アミノ酸を含むタンパク質の濃度を指す。

【0067】

典型的には、所与の生物学的試料中の1つ以上の体細胞突然変異のレベル（すなわち、突然変異のコピー数、突然変異を有するDNAの濃度、突然変異を有するmRNAの濃度及び/又は突然変異を有するタンパク質の濃度によって評価される）は、健康な対象から得られた参照生物学的試料（すなわち、NAFLD又はARLDを有しない対象から得られた肝臓試料）、NAFLD又はARLDを有することが知られている対象から得られた参照生物学的試料、及び/又は同じ対象から得られた参照生物学的試料中の1つ以上の体細胞突然変異のレベルと比較される。

30

【0068】

参照生物学的試料が健康な対象由来である実施形態では、対象由来の所与の生物学的試料中の1つ以上の体細胞突然変異のレベルが参照生物学的試料中のレベルよりも高い場合、これは、対象がNAFLD若しくはARLDを有するか、NAFLD若しくはARLDを発症するリスクが増加しているか、より重症型の肝疾患を発症するリスクが増加しているか、及び/又は関連疾患若しくは状態を発症するリスクが増加していることを示す。対象由来の所与の生物学的試料中の各1つ以上の体細胞突然変異のレベルが、健康な対象から得られた参照生物学的試料中のレベルと同等であるか、又はそれより低い場合、これは、対象がNAFLD若しくはARLDを有していないか、NAFLD若しくはARLDを発症するリスクが増加していないか、より重症型の肝疾患を発症するリスクが増加していないか、及び/又は肝疾患に関連する疾患若しくは状態を発症するリスクが増加していないことを示す。

40

【0069】

参照生物学的試料がNAFLD又はARLDを有することが知られている対象由来であ

50

る実施形態では、対象由来の所与の生物学的試料中の1つ以上の体細胞突然変異のレベルが参照生物学的試料中のレベルと同等であるか、又はそれより高い場合、これは、対象がNAFLD若しくはARLDを有するか、NAFLD若しくはARLDを発症するリスクが増加しているか、より重症型の肝疾患を発症するリスクが増加しているか、及び/又は関連疾患若しくは状態を発症するリスクが増加していることを示す。対象由来の所与の生物学的試料中の各1つ以上の体細胞突然変異のレベルが、参照生物学的試料中のレベルよりも低い場合、これは、対象がNAFLD若しくはARLDを有しないか、NAFLD若しくはARLDを発症するリスクが増加していないか、より重症型の肝疾患を発症するリスクが増加していないか、及び/又は関連する疾患若しくは状態を発症するリスクが増加していないことを示し得る。

10

【0070】

ある特定の実施形態では、参照生物学的試料は、所与の生物学的試料と同じ対象に由来する。典型的には、同じ対象由来の参照生物学的試料は、対象の身体若しくは肝臓の異なる領域から得られるか、及び/又は所与の生物学的試料とは異なる時点で対象から得られる。例えば、参照生物学的試料は、対象から得られた所与の生物学的試料よりも1~6ヶ月前、6~12ヶ月前、1~2年前又は2~3年前に対象から得ることができる。

【0071】

参照生物学的試料が同じ対象から得られるが、所与の生物学的試料よりも早い時点で得られる実施形態では、所与の生物学的試料における1つ以上の体細胞突然変異のレベルが参照生物学的試料におけるレベルよりも高い場合、これは、参照生物学的試料が対象から得られたより早い時点と比較して、対象のNAFLD若しくはARLDを発症するリスクが増加しているか、より重症型の肝疾患を発症するリスクが増加しているか、及び/又は関連する疾患若しくは状態を発症するリスクが増加していることを示し得る。あるいは、所与の生物学的試料における1つ以上の体細胞突然変異のレベルが参照生物学的試料におけるレベルよりも高い場合、これは、対象がNAFLD若しくはARLDを発症しているか、より重症型の肝疾患を発症しているか、及び/又は関連する疾患若しくは状態を発症していることを示し得る。

20

【0072】

対象由来の所与の生物学的試料中の各1つ以上の体細胞突然変異のレベルが参照生物学的試料中のレベルと同等である場合、これは、対象がNAFLD若しくはARLDを発症していないか、より重症型のNAFLD若しくはARLDを発症していないか、より重症型の肝疾患を発症していないか、及び/又は関連する疾患若しくは状態を発症していないことを示し得る。対象由来の所与の生物学的試料中の各1つ以上の体細胞突然変異のレベルが参照生物学的試料中のレベルよりも低い場合、これは、参照生物学的試料が対象から得られた初期の時点と比較して、対象がNAFLD若しくはARLDを発症するリスクが低下しているか、より重症型の肝疾患を発症するリスクが低下しているか、及び/又は関連疾患若しくは状態を発症するリスクが低下していることを示し得る。又は、対象由来の所与の生物学的試料中の各1つ以上の体細胞突然変異のレベルが参照生物学的試料中のレベルよりも低い場合、これは、対象が寛解期にあるNAFLD若しくはARLDを有するか、寛解期にあるより重症型の肝疾患を有するか、及び/又は寛解期にある関連疾患若しくは状態を有することを示し得る。

30

40

【0073】

ある特定の実施形態では、本発明の方法のステップb)は、対象から得られた所与の生物学的試料中に含有される異なる肝細胞クローンに由来する1つ以上の体細胞突然変異を検出及び/又は定量することを含み得る。本明細書で使用される場合、「肝細胞クローン」という用語は、共通の祖先を共有する同一肝細胞の群を指す。ある特定の実施形態では、本発明の方法のステップb)における1つ以上の体細胞突然変異(例えば、FOXO1、GPAM、CIDE B、ACVR2A、ALB及びTNRC6Bからなる群から選択される遺伝子内の1つ以上の体細胞突然変異)の検出は、対象から得られた異なる生物学的試料に含まれる異なる肝細胞クローンに由来する1つ以上の体細胞突然変異を検出及び/

50

又は定量することを含み得る。例えば、異なる肝細胞クローンは、肝臓の異なる領域から得られた異なる生物学的試料及び／又は異なる時点で対象から得られた異なる生物学的試料中に含まれ得る。それぞれの肝細胞クローンの肝細胞に選択的利点を付与する1つ以上（例えば、2、3、4、5、6つ以上）の異なる肝細胞クローンにおける1つ以上の同じ体細胞突然変異の検出は、対象におけるNAFLD又はARLDの存在の特に強力な指標である。したがって、ある特定の実施形態では、本発明の方法のステップb)は、対象から得られた生物学的試料に含まれる異なる肝細胞クローンに由来するDNA、RNA及び／又はタンパク質における1つ以上の体細胞突然変異を検出及び／又は定量することを含み、2つ以上の肝細胞クローンにおける当該肝細胞クローンの肝細胞に選択的利点を与える同じ体細胞突然変異の存在は、対象がNAFLD若しくはARLDに罹患しているか、NAFLD若しくはARLDを発症するリスクがあるか、より重症型の肝疾患を発症するリスクがあるか、及び／又は肝疾患に関連する疾患若しくは状態を発症するリスクがあることを示す。

10

【0074】

したがって、対象から得られた肝細胞及び／又は肝細胞クローンに選択的利点を付与する1つ以上の体細胞突然変異のレベルの検出及び／又は定量は、好適な疾患モニタリングプログラム、好適な治療を決定する際に臨床医を導くために、又は対象が寛解していることを臨床医に知らせるために使用することができる。

【0075】

生物学的試料が肝生検試料由来のDNA、RNA及び／又はタンパク質を含む実施形態では、肝細胞に選択的利点を付与する1つ以上の体細胞突然変異を有する肝細胞を含む肝生検試料又はそのマイクロダイセクションのサイズ及び／又は割合を測定し得る。肝細胞に選択的利点を付与する1つ以上の体細胞突然変異を有する肝細胞を含む肝生検試料又はそのマイクロダイセクションのサイズ及び／又は割合は、対象におけるNAFLD又はARLDの重症度及び／又は進行への更なる洞察を提供する。例えば、肝細胞に選択的利点を付与する1つ以上の体細胞突然変異を有する肝細胞を 0.5 cm^3 （すなわち、肝生検試料の50%以上の割合）を含む 1 cm^3 のサイズを有する肝生検試料は、対象が、その肝臓の有意な割合においてNAFLD又はARLDを有する（例えば、対象は、その肝臓の少なくとも5%、10%、20%、30%、40%、50%、又はそれ以上において脂肪症を有し得る）ことを示すと考えられ得る。したがって、ある特定の実施形態では、本発明の方法は、肝細胞に選択的利点を付与する1つ以上の体細胞突然変異を有する肝細胞を含む肝生検試料又はそのマイクロダイセクションのサイズ及び／又は割合を測定するステップを更に含む。したがって、肝細胞に選択的利点を付与する1つ以上の体細胞突然変異を有する肝細胞を含む肝生検試料又はそのマイクロダイセクション試料のサイズ及び／又は割合を使用して、好適な疾患モニタリングプログラム、好適な治療を決定する際に臨床医を導くか、又は対象が寛解していることを臨床医に知らせることができる。

20

30

【0076】

誤解を避けるために、本発明の方法のステップb)は、以下のうちのいずれか1つ以上における体細胞突然変異の検出及び／又は定量を含み得る：

- 遺伝子を含むDNA（例えば、cfDNA）分子、
- 遺伝子の一部分を含むDNA（例えばcfDNA）分子、
- 遺伝子の一部分に対応する配列を含むRNA分子、
- 遺伝子に対応する配列によってコードされるタンパク質、及び／又は
- 遺伝子の一部分に対応する配列によってコードされるタンパク質。

40

【0077】

本発明の方法は、対象におけるNAFLD又はARLDの診断又は予後診断の信頼性を改善するための、1つ以上の追加のマーカーの検出及び／又は定量を更に含んでもよい。例えば、本明細書に記載されるように、本発明者らは、NAFLD又はARLDを有する対象の肝細胞における新しい突然変異シグネチャーに注目した。特に、本発明者らは、肝疾患が進行するにつれて新しい突然変異シグネチャーの強度が増加することに注目した（

50

図16D～図16Fを参照されたい)。したがって、ある特定の実施形態では、本発明の方法は、対象の1つ以上の肝細胞に由来するDNA、RNA及び/又はタンパク質における突然変異シグネチャーの同定、及び任意選択的なモニタリングを更に含み得る。

【0078】

誤解を避けるために、本明細書で使用される場合、「突然変異シグネチャー」という用語は、NAFLD又はARLDを有することが既知であるか又はその疑いがある対象において観察される体細胞突然変異のセットを意味する。肝細胞などの細胞における突然変異シグネチャーを同定するための方法は、当技術分野において既知である。SigProfilerMatrixGeneratorソフトウェアの使用を含む例示的な方法が本明細書に記載される(Bergstrom, E. N. et al. BMC Genomics 20, 1-12 (2019))。ある特定の実施形態では、本発明の方法はまた、対象から得られた生物学的試料中のCLCN5遺伝子内の1つ以上の体細胞突然変異を検出するステップを含んでもよい。本発明者らは、CLCN5遺伝子における体細胞突然変異が、NAFLD又はARLDに罹患している患者において起こることを見出した。CLCN5遺伝子はクロライドチャンネルをコードし、これはX連鎖腎石灰化症を引き起こすが、生殖系列において突然変異した場合に既知の肝臓表現型を引き起こさない。CLCN5遺伝子内の体細胞突然変異は、対象におけるNAFLD又はARLDの更なる指標、対象がNAFLD又はARLDを発症するリスクの増加の更なる指標、及び/又は対象がより重症型の肝疾患又は関連疾患若しくは状態を発症するリスクの増加の更なる指標であり得る。特定の例示的な実施形態では、CLCN5遺伝子において検出される1つ以上の体細胞突然変異は、変異を含有する肝細胞に選択的利点を付与するものである。

10

20

【0079】

本発明の方法において検出され得るCLCN5遺伝子における体細胞突然変異の例としては、CLCN5タンパク質のG330D、S371G及び/若しくはG570R置換をもたらす突然変異、又はCLCN5タンパク質のC171(すなわちC171*)、W192(すなわちW192*)、E676(すなわちE676*)若しくはG791(すなわちG791*)ナンセンス突然変異が挙げられる。

【0080】

ある特定の実施形態では、本発明の方法はまた、対象から得られた生物学的試料中の1つ以上の非コード体細胞突然変異を検出するステップを含んでもよい。本発明者らは、NAFLD又はARLDに罹患している患者のNEAT1遺伝子における非コード突然変異を検出した。したがって、ある特定の実施形態では、本発明の方法は、NEAT1遺伝子における1つ以上の非コード体細胞突然変異を検出することを更に含む。ある特定の実施形態では、NEAT1遺伝子において検出される1つ以上の非コード体細胞突然変異は、突然変異を含む肝細胞に選択的利点を付与するものである。

30

【0081】

本発明の方法において検出され得るNEAT1遺伝子における非コード体細胞突然変異の例としては、GRCh37d5ヒト参照ゲノムによるそれらのゲノム位置を参照することによって提供される、ヒト染色体11上の以下の突然変異が挙げられる：65190784 G>A；65193333 C>T；65194195 del19；65197674 C>T；65197688 C>G；65198610 del47；65201066 del AATT(図8では「delAATT」と称される)；65201763 G>T；65202053 G>A；65202131 delT；65202783 T>A；65203823 C>T；65203921 C>A；65204731 C>T；65205445 T>C；65205770 T>A；65207563 delA；65210139 del11；及び65210711 C>T(図8も参照されたい)。

40

【0082】

特定の実施形態では、本発明の方法はまた、生物学的試料中のDNA分子のテロメア長を決定するステップと、生物学的試料中のDNA分子のテロメア長を、正常な肝細胞から

50

得られたDNA分子のテロメア長と比較するステップと、を含み得る。以前の報告は、NAFLD又はARLDなどの慢性肝疾患がテロメア短縮に関連することを示唆している(Wiemann et al., FASEB J, 2002, 16, 935-42)。すなわち、罹患肝細胞において、テロメア長は、非罹患肝細胞におけるテロメア長よりも短いことが多い。したがって、対象から得られた生物学的試料中のDNA分子のテロメア長は、対象におけるNAFLD若しくはARLDの更なる指標、対象がNAFLD若しくはAFLDを発症するリスクの増加の更なる指標、対象がより重症型の肝疾患を発症するリスクの増加の更なる指標、及び/又は対象に関連する疾患若しくは状態を発症するリスクの増加の更なる指標として使用され得る。したがって、ある特定の実施形態では、本発明の方法は、

10

c) 生物学的試料中のDNA分子のテロメア長を決定するステップと、

d) 生物学的試料中のDNA分子のテロメア長を、正常な肝細胞(例えば、NAFLD若しくはARLDに罹患していない対象から、又は健康な対象の肝臓の領域から得られた肝細胞)から得られたDNA分子の平均テロメア長と比較するステップと、を更に含む得る。典型的には、DNA分子は、同じ対象由来の正常な肝細胞から、又はNAFLD若しくはARLDに罹患していない年齢が一致した対象由来の正常な肝細胞から得られる。ある特定の例示的な実施形態では、本方法は、

c) 生物学的試料中のDNA分子の平均テロメア長を決定するステップと、

d) 生物学的試料中のDNA分子の平均テロメア長を、正常な肝細胞から得られたDNA分子の平均テロメア長と比較するステップと、を更に含む得る。本明細書で使用される場合、「平均テロメア長」という用語は、所与の試料において測定されたテロメアの平均長である。平均テロメア長は、所与の試料についてのテロメア長の平均値、中央値又は最頻値であり得る。典型的には、平均テロメア長は、所与の試料についてのテロメア長の平均値である。あるいは、平均テロメア長は、所定の試料中のクローンについてのテロメア長の平均値、中央値又は最頻値であり得る。

20

【0083】

対象におけるNAFLD又はARLDを診断又は予後診断する場合、追加のリスク因子も考慮され得る。更なる危険因子としては、典型的には、年齢、メタボリックシンドロームの存在、性別、民族、食事因子、喫煙状態(例えば、現在の喫煙者又は以前の喫煙者)、閉塞性睡眠時無呼吸、肥満、2型糖尿病、高血圧、脂質異常症(例えば、高コレステロール血症)及び心血管疾患が挙げられる。体細胞突然変異に加えて遺伝因子(例えば、FOXO1、GPAM、CIDEB、ACVR2A、ALB及びTNRC6B、並びにCLCN5及びNEAT1遺伝子内の体細胞突然変異)もまた、本発明の方法において考慮され得る。例えば、パタチン様ホスホリパーゼドメイン含有3(PNPLA3)遺伝子における生殖系列多型は、NAFLDに対する遺伝的リスクを増加させることが知られており(Sookoian et al. J Lipid Res 2009; 50: 2111-6)、したがって考慮されてもよい。更なる例として、FOXO1及びGPAMにおける生殖系列多型などの、インスリン抵抗性又は脂質合成に関連する遺伝子における生殖系列多型もまた考慮されてもよい。NAFLD又はARLDに関連していると推測される生殖系列多型の例は、Valenti et al., Journal of Hepatology, 2009, 50, S265、及びHakim et al., Hepatology, 2021, doi: 10.1002/hep.32038に記載されている。

30

40

【0084】

例示的な実施形態では、本発明の方法は、対象の1つ以上の肝細胞に選択された利点を付与する1つ以上の体細胞突然変異の存在及び/又はレベル(例えば、コピー数)に基づいて、更に、対象の肝細胞の1つ以上の更なる遺伝的及び/又は物理的特徴、例えば、以下の特徴:

対象の肝細胞における突然変異シグネチャーの存在及び/又は発生;

CLCN5遺伝子における体細胞突然変異の存在及び/又はレベル(例えば、コピー数);

50

非コード体細胞突然変異、例えば、NEAT1遺伝子における非コード体細胞突然変異の存在及び/又はレベル(例えば、コピー数)；

肝生検試料中の肝細胞に選択的利点を付与する1つ以上の体細胞突然変異を保有する、DNA、RNA及び/又はタンパク質が由来する肝生検試料のサイズ及び/又は割合；

対象由来の1つ以上の肝細胞に由来するDNAのテロメア長；及び/又は

年齢、メタボリックシンドロームの存在、性別、民族、食事要因、喫煙状態(例えば、現在の喫煙者又は以前の喫煙者)、閉塞性睡眠時無呼吸、肥満、2型糖尿病、高血圧、脂質異常症(例えば、高コレステロール血症)及び心血管疾患などの対象の更なる危険因子、のうちの1つ以上に基づいて、スコアを計算することを含むステップc)を更に含む得る。

10

【0085】

対象の1つ以上の肝細胞における体細胞突然変異の存在及び/又はレベル(例えば、コピー数)、並びに上述の遺伝的及び/又は物理的特徴のうちの1つ以上に基づくスコアの計算は、好適な疾患モニタリングプログラム、好適な治療を決定する際に臨床医に更なるガイダンスを提供するか、又は対象が寛解していることを臨床医に知らせ得る。

【0086】

対象におけるNAFLD又はARLD進行をモニタリングする方法：

本発明はまた、疾患モニタリングの増加から利益を得る、NAFLD又はARLDに罹患している対象を同定するための方法であって、上記のステップa)及びb)を含む、方法を提供する。対象のモニタリングの増加は、循環肝酵素レベルをモニタリングするためのコンピュータ断層撮影(computerized tomography、CT)スキャン及び/又は血液検査などのNAFLD又はARLDに適したモニタリング技術を使用して、定期的な間隔(例えば、毎年1回、6ヶ月に1回又は3ヶ月に1回)で対象をモニタリングすることを含み得る。

20

【0087】

本明細書に記載される診断及び予後診断方法はまた、対象におけるNAFLD又はARLDの発症又は進行をモニタリングするために使用され得る。例えば、対象におけるNAFLD又はARLDのモニタリングは、本明細書に記載の方法のステップa)及びb)を数日、数週間、数ヶ月又は数年の間隔で繰り返すことによって達成され得る。本明細書に記載の診断方法及び予後診断方法は、NAFLD若しくはARLDを発症するリスクのある対象のモニタリング、より重症型の肝疾患を発症するリスクのある対象のモニタリング、肝疾患に関連する疾患若しくは状態を発症するリスクのある対象のモニタリング、及び/又はNAFLD若しくはARLDの治療を受けている対象におけるNAFLD若しくはARLDのモニタリングに特に適している。例示的な実施形態では、本明細書に記載の診断方法又は予後診断方法のステップa)及びb)は、毎週1回、2週間ごとに1回、3週間ごとに1回、4週間ごとに1回、毎月1回、2ヶ月ごとに1回、3ヶ月ごとに1回、又は6ヶ月ごとに1回実施される。特定の例示的な実施形態では、ステップa)及びb)は、2年に1回、毎年1回、毎年2回、又は毎年3回の間隔で繰り返される。例えば、ステップa)及びb)は、2年に1回、毎年1回、毎年2回、又は毎年3回、1~5年、1~10年、1~20年、1~30年、若しくは1~40年間にわたって繰り返され得る。

30

40

【0088】

NAFLD又はARLDを診断又は予後判定するためのインビボ方法：

本発明はまた、対象におけるNAFLD又はARLDを診断又は予後判定するための方法であって、

a)ある用量の診断プローブを対象に投与するステップと、

b)対象において診断プローブを検出するステップであって、当該診断プローブが、対象の肝臓におけるFOXO1、GPAM、CIDE B、ACVR2A、ALB及びTNRC6Bからなる群から選択される遺伝子内の体細胞突然変異の存在及び/又は非存在を示し、FOXO1、GPAM、CIDE B、ACVR2A、ALB又はTNRC6B遺伝子内の体細胞突然変異の存在は、対象がNAFLD若しくはARLDに罹患しているか、N

50

AFLD若しくはARLDを発症するリスクがあるか、より重症型の肝疾患を発症するリスクがあるか、及び/又は消化管がんなどの肝疾患に関連する疾患若しくは状態を発症するリスクがあることを示す、ステップと、を含む方法を提供する。

【0089】

本発明はまた、対象におけるNAFLD又はARLDを診断又は予後判定するための方法であって、

- 対象は、診断プローブが投与された対象であり、

- 当該方法は、対象において診断プローブを検出するステップを含み、診断プローブは、対象の肝臓におけるFOXO1、GPAM、CIDE B、ACVR2A、ALB及びTNRC6Bからなる群から選択される遺伝子内の1つ以上の体細胞突然変異の存在又は非存在を示すものであり、

10

1つ以上の体細胞突然変異の存在は、対象がNAFLD若しくはARLDに罹患しているか、NAFLD若しくはARLDを発症するリスクがあるか、より重症型の肝疾患を発症するリスクがあるか、及び/又は消化管がんなどの肝疾患に関連する疾患若しくは状態を発症するリスクがあることを示す、方法を提供する。

【0090】

本発明はまた、対象におけるNALFLD又はARLDの診断又は予後判定において使用するための診断プローブであって、当該使用が、

a) ある用量の診断プローブを対象に投与することと、

b) 対象において診断プローブを検出することであって、当該診断プローブが、対象の肝臓におけるFOXO1、GPAM、CIDE B、ACVR2A、ALB及びTNRC6Bからなる群から選択される遺伝子内の1つ以上の体細胞突然変異の存在及び/又は非存在を示し、1つ以上の体細胞突然変異の存在は、対象がNAFLD若しくはARLDに罹患しているか、NAFLD若しくはARLDを発症するリスクがあるか、より重症型の肝疾患を発症するリスクがあるか、及び/又は消化管がんなどの肝疾患に関連する疾患若しくは状態を発症するリスクがあることを示す、検出することと、を含む、方法を提供する。

20

【0091】

本明細書に記載のNALFLD又はARLDを診断又は予後判定するためのインビボ方法における使用に適した診断プローブには、1つ以上の体細胞突然変異を含むFOXO1、GPAM、CIDE B、ACVR2A、ALB若しくはTNRC6B遺伝子に特異的に結合するか、又は1つ以上の体細胞突然変異を含むそれぞれの遺伝子によってコードされるFOXO1、GPAM、CIDE B、ACVR2A、ALB若しくはTNRC6Bタンパクに特異的に結合する、小分子、ペプチド(環状ペプチドを含む)、タンパク、核酸(例えば、アンチセンスヌクレオチド配列; 三重らせん; siRNA又はmiRNA; 及び生物学的に活性なタンパク、ポリペプチド又はペプチドをコードするヌクレオチド配列を含むがこれらに限定されない、DNA及びRNAヌクレオチド)、合成又は天然無機分子、及び合成又は天然有機分子が含まれるが、これらに限定されない。このような予後診断プローブは、対象におけるNALFLDの画像化及び/又は診断若しくは予後判定に適した「標識」を含み得る。好適な標識としては、例えば、放射性同位体、放射性核種、同位体、陽電子放射体、ガンマ放射体、蛍光基、発光基、発色基、ビオチン(例えば、ストレプトアビジン複合体化と組み合わせる)又は光親和性基が挙げられる。選択される標識のタイプは、所望の検出方法に依存する。ある特定の実施形態では、NAFLDの診断又は予後判定に使用するための診断プローブは、蛍光インサイチュハイブリダイゼーション(fluorescence in situ hybridization、FISH)に適した蛍光標識を含む。

30

40

【0092】

誤解を避けるために、本明細書に記載のインビボ方法は、FOXO1、GPAM、CIDE B、ACVR2A、ALB及びTNRC6Bからなる群から選択される遺伝子内の1つ以上の体細胞突然変異の検出を可能にし、当該少なくとも1つ以上の体細胞突然変異が対象の肝細胞に選択的利点を付与する診断プローブの使用を含み得る。

50

【0093】

治療に適したNAFLD又はARLD患者の選択：

本発明はまた、FOXO1、GPAM、CIDE B、ACVR2A、又はTNRC6Bタンパク活性を阻害又は調節する治療剤による治療から利益を得る、NAFLD又はARLDに罹患している対象を同定するための方法を提供する。ある特定の実施形態では、本方法は、

a 1) 対象から得られた生物学的試料を提供するステップと、

b 1) 生物学的試料中のFOXO1、GPAM、CIDE B、ACVR2A、ALB、及びTNRC6Bからなる群から選択される遺伝子内の1つ以上の体細胞突然変異を検出するステップであって、1つ以上の体細胞突然変異の存在は、対象が、FOXO1、GPAM、CIDE B、ACVR2A、又はTNRC6Bタンパク活性を阻害又は調節する治療剤による治療から利益を得る対象であることを示す、ステップと、を含む。

10

【0094】

ある特定の実施形態では、本発明は、FOXO1又はGPAM活性を阻害又は調節する治療剤による治療から利益を得る、NAFLD又はARLDに罹患している対象を同定するための方法に関する。実施形態では、本方法は、

a 1) 対象から得られた生物学的試料を提供するステップと、

b 1) 生物学的試料中のFOXO1及び/又はGPAM遺伝子内の1つ以上の体細胞突然変異を検出するステップであって、1つ以上の体細胞突然変異の存在は、対象が、FOXO1又はGPAM活性を阻害又は調節する治療剤による治療から利益を得る対象であることを示す、ステップと、を含む。

20

【0095】

FOXO1、GPAM、CIDE B、ACVR2A、又はTNRC6Bタンパク質の活性を阻害及び/又は調節するために使用され得る好適な治療剤としては、小分子、ペプチド(環状ペプチドを含む)、タンパク質、核酸(例えば、アンチセンスヌクレオチド配列；三重らせん；siRNA又はmiRNA；及び生物学的に活性なタンパク質、ポリペプチド又はペプチドをコードするヌクレオチド配列を含むがこれらに限定されない、DNA及びRNAヌクレオチド)、合成又は天然無機分子、合成又は天然有機分子、及びCRISPR(すなわち、クラスタリングされた規則的な間隔の短い回文配列リピート)が挙げられるがこれらに限定されない。例えば、治療剤は、FOXO1、GPAM、CIDE B、ACVR2A、又はTNRC6Bタンパク質に結合し、その活性を阻害する有機小分子化合物であってもよく、又は例えば、治療剤は、1つ以上の体細胞突然変異を含むFOXO1、GPAM、CIDE B、ACVR2A、又はTNRC6Bタンパク質の発現をノックダウンするmiRNA又はsiRNAであってもよい。

30

【0096】

FOXO1タンパクの活性を阻害及び/又は調節する好適な治療剤の具体例としては、AS1842856とも呼ばれる5-アミノ-7-(シクロヘキシルアミノ)-1-エチル-6-フルオロ-4-オキソ-1,4-ジヒドロキノリン-3-カルボン酸(参照により本明細書に組み込まれる欧州特許出願公開第1650192(A1)号及びMolecular Pharmacology, 2010 78(5)961-970を参照されたい)、及び2-[2-(メチルアミノ)ピリミジン-4-イル]-1,5,6,7-テトラヒドロピロロ[3,2-c]ピリジン-4-オン(すなわち、Langlet et al., 2017の化合物8)などのLanglet et al. (参照により本明細書に組み込まれるCell, 2017, 171, 824-835)に開示された化合物が挙げられる。

40

【0097】

本発明における使用に適した治療剤の更なる例は、GPAMタンパクの小分子阻害剤であるFSG67である(参照により本明細書に組み込まれる、Wydysh et al., J Med Chem. 2009, 28; 52(10): 3317-27及びKuhajda et al., Am J Physiol Regul Integr Com

50

p Physiol. 2011; 301(1): R116-30)。

【0098】

本発明はまた、NAFLD又はARLDの治療において使用するための治療剤であって、当該使用が、

a 2) 対象から得られた生物学的試料を提供するステップと、

b 2) 生物学的試料中のFOXO1、GPAM、CIDE B、ACVR2A、ALB及びTNRC6Bからなる群から選択される遺伝子内の1つ以上の体細胞突然変異を検出するステップであって、1つ以上の体細胞突然変異の存在は、対象が、FOXO1、GPAM、CIDE B、ACVR2A、ALB又はTNRC6Bタンパク質活性を阻害又は調節する治療剤による治療から利益を得る対象であることを示す、ステップと、

c 2) FOXO1、GPAM、CIDE B、ACVR2A、又はTNRC6Bタンパク質活性を阻害又は調節する治療剤の1以上の用量を対象に投与するステップと、を含む、治療剤を提供する。

【0099】

ある特定の実施形態では、本発明は、NAFLDの治療において使用するための治療剤であって、当該使用が、

a 2) 対象から得られた生物学的試料を提供するステップと、

b 2) 生物学的試料中のFOXO1及び/又はGPAM遺伝子内の1つ以上の体細胞突然変異を検出するステップであって、1つ以上の体細胞突然変異の存在は、対象が、FOXO1又はGPAMタンパク質活性を阻害又は調節する治療剤による治療から利益を得る対象であることを示す、ステップと、

c 2) FOXO1又はGPAM活性を阻害又は調節する治療剤の1以上の用量を対象に投与するステップと、を含む、治療剤を提供する。

【0100】

誤解を避けるために、本明細書に記載されるステップa 1)、b 1)、a 2)及びb 2)は、独立して、本明細書の上記に記載される本発明によるNAFLD(又はARLD)を診断又は予後判定するための方法のステップa)及びb)について本明細書に記載される特徴及びステップのうちいずれかを含み得る。また、誤解を避けるために、本明細書に記載される方法のステップb 2)は、対象の1つ以上の肝細胞に選択的利点を付与するFOXO1、GPAM、CIDE B、ACVR2A、ALB及びTNRC6Bからなる群から選択される遺伝子内の1つ以上の体細胞突然変異を検出することを含み得る。

【0101】

誤解を避けるために、本明細書で使用される場合、「体細胞突然変異」という用語は、「1つ以上の体細胞突然変異」という用語を含むことが理解されるべきである。例えば、「体細胞突然変異」という用語は、1つ、2つ、3つ、4つ、5つ、6つ、又はそれ以上の体細胞突然変異を指し得る。本明細書における「FOXO1、GPAM、CIDE B、ACVR2A、ALB及びTNRC6Bからなる群から選択される遺伝子」への任意の言及は、FOXO1、GPAM、CIDE B、ACVR2A、ALB及びTNRC6B遺伝子のうちの1つ以上の選択を含むことも理解されるべきである。すなわち、本発明の方法は、例えば、FOXO1遺伝子内の1つ以上の体細胞突然変異、GPAM遺伝子内の1つ以上の体細胞突然変異、CIDE B遺伝子内の1つ以上の体細胞突然変異、FOXO1遺伝子内の1つ以上の体細胞突然変異、ACVR2A遺伝子内の1つ以上の体細胞突然変異、ALB遺伝子内の1つ以上の体細胞突然変異、及び/又はTNRC6B遺伝子内の1つ以上の体細胞突然変異の検出を含み得る。

【0102】

キット：

本発明はまた、対象におけるNAFLD又はARLD(又はAFLD)の診断又は予後診断に使用するためのインビトロ診断キットであって、FOXO1、GPAM、CIDE B、ACVR2A、ALB及びTNRC6Bからなる群から選択される遺伝子内の1つ以上の体細胞突然変異を検出するための1つ以上の試薬を含み、任意選択的にCLCN5遺

伝子及び/又はNEAT1遺伝子内の1つ以上の体細胞突然変異を検出するための1つ以上の試薬を含む、インビトロ診断キットを提供する。

【0103】

ある特定の実施形態では、本発明は、対象におけるNAFLDの診断又は予後診断に使用するためのインビトロ診断キットであって、FOXO1及び/又はGPAM遺伝子内の1つ以上の体細胞突然変異を検出するための1つ以上の試薬を含み、任意選択的にCLCN5遺伝子及び/又はACVR2A遺伝子内の1つ以上の体細胞突然変異を検出するための1つ以上の試薬を含む、インビトロ診断キットに関する。

【0104】

任意選択的に、キットは、テロメア長を測定するのに有用な1つ以上の試薬を含み得る 10

【0105】

「キット」という用語は、本発明の1つ以上のバイオマーカーを特異的に検出するための少なくとも1つの試薬、例えばプローブ又は小分子を含む任意の製品（例えばパッケージ又は容器）を指す。キットは、本発明の方法を実行するためのユニットとして販売促進、流通、又は販売され得る。

【0106】

キットは、FOXO1、GPAM、CIDE B、ACVR2A、ALB及びTNRC6Bからなる群から選択される遺伝子内の1つ以上の体細胞突然変異を検出及び/又は定量するのに必要な1つ以上の試薬を含んでもよい。本発明のキットはまた、CLCN5遺伝子及び/又はACVR2A遺伝子内の1つ以上の体細胞突然変異を検出及び/又は定量するための1つ以上の試薬を任意選択的に含んでもよい。試薬の特定の例としては、FOXO1、GPAM、CIDE B、ACVR2A、ALB又はTNRC6B遺伝子配列の一部とハイブリダイズするオリゴヌクレオチドが挙げられる。このようなオリゴヌクレオチドの設計は、当業者の能力の範囲内である。好適には、そのようなオリゴヌクレオチドは、1つ以上の体細胞突然変異を含むFOXO1、GPAM、CIDE B、ACVR2A、ALB又はTNRC6B遺伝子配列を増幅、定量又は配列決定するために使用され得る。 20

【0107】

ある特定の実施形態では、キットは、1つ以上の突然変異アミノ酸を含むFOXO1、GPAM、CIDE B、ACVR2A、ALB又はTNRC6Bタンパクに結合する結合タンパクを含んでもよい。好適な結合タンパク質には、1つ以上の突然変異アミノ酸を含むFOXO1、GPAM、CIDE B、ACVR2A、ALB又はTNRC6Bタンパク質に特異的に結合するポリクロナール又はモノクロナール抗体が含まれる。好適には、そのような結合タンパク質は、生物学的試料中の1以上の突然変異を含むFOXO1、GPAM、CIDE B、ACVR2A、ALB又はTNRC6Bタンパク質のレベルを定量するための方法において使用され得る。 30

【0108】

ある特定の実施形態では、キットは、1つ以上の参照標準、例えば、核酸、ペプチド、ポリペプチド又はタンパク質を更に含み得る。このような参照標準は、例えば、1つ以上の体細胞突然変異を含まないFOXO1、GPAM、CIDE B、ACVR2A、ALB及び/若しくはTNRC6B遺伝子又はタンパク質、例えば、「野生型FOXO1」及び「野生型GPAM」遺伝子又はタンパク質、並びに例えば、野生型CIDE B、ACVR2A、ALB若しくはTNRC6Bタンパク質又は遺伝子に対応する核酸、ペプチド、ポリペプチド又はタンパク質であってもよい。キットは、一般的な分子タグ（例えば、緑色蛍光タンパク質及びベータ-ガラクトシダーゼ）を含んでもよい。キット中の試薬は、個々の容器中に、又は単一容器中の2つ以上の試薬の混合物として提供されてもよい。更に、キット内の成分の使用を記載する説明資料を含めることができる。説明資料はまた、結果の解釈、例えば、肯定的な所見を構成するために特定の分析物のどのレベルをとるべきかについてのガイダンスを提供し得る。 40

【0109】

キットはまた、キットが設計される特定の用途を容易にするための追加の成分を含んでもよい。例えば、キットは、標識を検出する手段（例えば、酵素標識のための酵素基質、蛍光標識を検出するためのフィルタセット、ヒツジ抗マウスHRPなどの適切な二次標識など）及び対照に必要な試薬（例えば、対照生物学的試料又は標準）を更に含んでもよい。キットは更に、ヘルスケア環境において開示される本発明の方法において使用するために必要なグレードの緩衝液及び他の試薬を含み得る。非限定的な例としては、担体タンパク質又は界面活性剤などの非特異的結合を減少させる薬剤が挙げられる。

【0110】

等価物：

本発明は、本明細書において広く包括的に記載されている。当業者は、本明細書に記載される全てのパラメータ、寸法、材料及び構成が例示的であることを意味し、実際のパラメータ、寸法、材料及び/又は構成が、本発明の教示が使用される特定の1つ若しくは複数の用途に依存することを容易に理解するであろう。当業者は、本明細書に記載の本発明の特定の実施形態に対する多くの等価物を、日常的な実験のみを使用して認識するか、又は確認することができるであろう。したがって、前述の実施形態は、例としてのみ提示されており、添付の特許請求の範囲及びその均等物の範囲内では、本発明は、具体的に記載及び特許請求される以外の方法で実施され得ることを理解されたい。本発明は、本明細書に記載の各個々の特徴、システム、物品、材料、キット、及び/又は方法を対象とする。加えて、そのような特徴、システム、物品、材料、キット及び/又は方法が相互に矛盾しない場合、2つ以上のそのような特徴、システム、物品、材料、キット及び/又は方法の任意の組み合わせが、本発明の範囲内に含まれる。

【0111】

更に、一般的な開示内に含まれるより狭い種群及び亜群の各々も、本発明の一部を形成する。これは、除外されるものが本明細書に具体的に記載されているかどうかにかかわらず、属から任意の主題を除外する条件又は否定的限定（negative limitation）下で、本発明の一般的な説明を含む。

【0112】

参照による組み込み：

論文、特許及び特許出願、並びに本明細書で言及又は引用される他の文書及び電子利用可能な情報の内容は、各個々の刊行物が参照により組み込まれることが具体的かつ個別に示された場合と同じ程度まで、参照によりそれらの全体が本明細書に組み込まれる。出願人は、そのような論文、特許、特許出願、又は他の紙（physical）文書及び電子文書からの任意及び全ての材料及び情報を本出願に物理的に組み込む権利を保有する。

【0113】

ここで、以下の実施例を参照することにより、本発明を非限定的に説明する。

【実施例】

【0114】

材料及び方法

肝臓試料：

全ての生物学的試料は、Local Research Ethics Committee (16/N1/0196) によって承認された手順に従って、Addenbrooke's Hospital (Cambridge, UK) からインフォームドコンセントを得て収集した。試料を液体窒素中で急速凍結し、Human Research Tissue Bank of the Cambridge University Hospitals NHS Foundation Trustにおいて - 80 で保存した。

【0115】

結腸直腸がん転移の肝臓切除を受けている患者から正常な肝臓組織を得た。転移から離れた標本を得て、組織病理学的検査で腫瘍がないことを確認した。1人の患者（PD36718）は、同側の肝臓への術前門脈塞栓術を受けたが、切除の前にネオアジュバント化

学療法を受けた者はいなかった。バックグラウンド疾患肝組織は、肝細胞がん（HCC）のための肝臓切除、又はHCC若しくは肝不全のための肝移植のいずれかを受けている、非アルコール性脂肪性肝疾患（NAFLD）又はアルコール関連肝疾患（ALD）を有する対象から得た。経動脈化学塞栓術による単回治療を受けた1人の患者（PD37118）を除いて、術前局所領域療法を受けた者はいなかった。PNPLA3（rs738409）多型遺伝子型は、全ゲノム配列決定データに由来した。

【0116】

全ての肝臓標本を、レーザーキャプチャーマイクロダイセクションに使用された新鮮凍結ブロックから離れたホルマリン固定パラフィン包埋（formalin-fixed paraffin-embedded、FFPE）試料について、Kleinerシステム（Kleiner et al., Hepatology, 2005; 41(6): 1313-1321）に従って採点した。NAFLD用に関与されたKleinerスコアは、脂肪症、小葉炎症及び肝細胞バルーンングの存在を評価して、累積NAFLD活動性スコア（NAFLD activity score、NAS）を生成する。これを、ALDについて検証されたスコアリングシステムの非存在下で、正常試料及びALD試料に適用して、疾患のステージ及びグレードに関して研究の全ての試料間の比較可能性を可能にした。細胞性又は結節性異形成の存在又は非存在を、臨床FFPE試料において全体的に、並びに特に、レーザーキャプチャーマイクロダイセクション（laser capture microsection、LCM）及び配列決定に使用した新鮮凍結ブロックにおいて評価した。

【0117】

試料調製：

肝生検組織切片の調製、レーザーキャプチャーマイクロダイセクション（LCM）及びその後の細胞溶解、DNA抽出、及び全ゲノム配列決定（whole genome sequencing、WGS）に使用したプロトコルは、Brunner et al.（Nature, 2019, 574, 538-542）によって以前に記載されたものに従った。簡潔に述べると、20µm厚の組織切片（Leicaクリオームを用いて調製）を70%エタノールで固定し、Leica Microsystems LMD 7000を使用するその後のLCM生成のためにヘマトキシリン及びエオシンで染色した。次いで、Arcturus（登録商標）PicoPure（登録商標）DNA Extraction Kit（Thermo Fisher Scientific）を製造業者の指示に従って使用して、マイクロダイセクションした試料を溶解した。Illumina配列決定のためのDNAライブラリーを、ペアエンドWGSへの提出のために低入力量のDNAのために最適化されたプロトコルを使用して調製した。得られたリードを、BWA-MEMアルゴリズム（Li and Durbin, Bioinformatics, 2010, 26, 589-595）を使用してGRCh37d5ヒト参照ゲノムにマッピングした。

【0118】

この研究で使用されたデータセットは、5人の正常肝臓対照、アルコール関連肝疾患（ALD）を有する10人及びNAFLDを有する17人を含む、32個の肝臓試料からの1202個のゲノムを含んだ。これらの試料のうち9つは、同期性HCC及び根底にある肝硬変を有する患者からのものであった。5年の期間にわたってNAFLDを有する1人の患者からの3つの肝切除試料（試料PD37918b、PD37915b及びPD37910b）を含む、更なる8つの試料が、根底に肝硬変を有しないHCCを有していた。患者の臨床的及び組織学的特徴は、根底にある疾患プロセスについて予想された分布を示した。

【0119】

レーザーキャプチャーマイクロダイセクションを使用して、肝臓試料からの全ゲノム配列決定のために100~500個の肝細胞の連続群を単離した。マイクロダイセクションを31xの平均深度まで配列決定した。体細胞置換及び構造バリエーションを、Brunner et al.（2019）によって記載された方法に従ってコールしたが、インデルコールの特徴付け及び精度を改善するために勾配ブースティング回帰木を使用した。体細

胞突然変異の完全なカタログを用いて、系統樹構造、クローンサイズ、ドライバー突然変異、テロメア長及び突然変異シグネチャーを推測した。

【0120】

SNVコーリング：

この研究で使用されたSNVコーリングワークフローのいくつかのステップは、Brunner et al (2019) によって以前に記載された。基本的なSNV同定は、一塩基置換 (single base substitution、SBS) バリエーションをコールするために、期待値最大化 (CaVEMan) アルゴリズム (Jones, D. et al., Curr Protoc Bioinformatics 56, 15.10.1-15.10.18 (2016)) を通してCancer Variantsを使用し、患者ごとのバルク生検を適合正常対照とした。8つの解剖学的肝臓セグメントから生検された2人のNAFLD患者について、突然変異をコールするために、甲状腺濾胞LCM試料を不適合対照として使用した。二次十字形DNA構造の不正確なプロセッシングから生じる重複リード及びLCMライブラリー調製特異的人為的変異体を、特注のプロセッシング後フィルタリングで除去した。後者のフィルタリングステップは、少なくとも2つの配列決定されたDNA断片を支持する全ての変異体を考慮するように構成された。現在の研究では、Brunner et al (2019) に記載されているエントロピーメトリックベースのバリエーションフィルタリングステップを、Yoshida et al. (Yoshida, K. et al., Nature 578, 266-272 (2020)) によって記載されているようなベータ二項ベースのフィルタリング手法で置き換えたが、これは、真正の体細胞突然変異が典型的には過剰分散している (すなわち、各患者に属するゲノムのセット中の限られた数のゲノムのみが存在する) 一方で、系統的アーチファクト又は生殖系列バリエーションは一般に分散不足であり、同じ患者の生検からのマイクロダイセクションに由来する全てではないが多くのゲノムにわたって観察可能であるという原理に基づいている。この研究において、各SNVについての突然変異保有及び総リードの数は、患者特異的に全てのマイクロダイセクションにわたってコールされたSNV当たりの各塩基 (A、C、G、T) についての生の対立遺伝子カウントを列挙することによって計算され、ここで、0.1の分散推定値を有する突然変異は、真の体細胞バリエーションである可能性が高いと考えられた。有効性を保証するために、ゲノムブラウザを使用して最終SNVコールのサブセットの手動検査を実行した。更なるチェックは、組織学画像によって捕捉されるような空間的に近接するマイクロダイセクションが、(すなわち、同じ組織切片上のx-y空間に関して同じ近傍内、同じ肝硬変結節内、又は異なるz平面からの組織切片上の重複するx-y位置に) 共通の突然変異を共有することをチェックすることを含んでいた。

【0121】

SNVクラスターを同定するために階層ディリクレ過程を使用する：

ノンパラメトリックベイズ階層ディリクレ過程 (hierarchical Dirichlet process、HDP) を実施して、各患者生検について複数のマイクロダイセクションにわたってコールされた類似のバリエーション対立遺伝子画分 (VAF) を有するSNVをクラスタリングした。クラスタリングアルゴリズムの完全な数学的及び実施の詳細は、Brunner et al (2019) に記載されている。このN次元ディリクレ過程 (N-dimensional Dirichlet process、NDP) クラスタリング手法は、10,000回のバーンイン反復と、それに続くクラスタリングに使用された5,000回の事後ギブスサンプリング反復とを用いて実行された。このクラスのアルゴリズムは、見つけるクラスターの数に任意に予め指定する必要がないので、SNVクラスターの同定のために選択された。代わりに、各サンプリング反復において、前の反復において存在しなかった新しいクラスターに突然変異が割り当てられることになる定義された確率が存在する。他方、クラスターはまた、全てのメンバー突然変異が他のクラスターに割り当てられる場合、将来の反復において除去することもできる。このようにして、SNVクラスター数は、サンプリングチェーン全体にわたって変化することが許容される。多数のクラスターからなる過度に複雑な解決策 (これは、情報価値のないクラスターを作製する可能性を増加させる) を避けるた

めに、患者当たり100個のSNVクラスターという上限を課した。ラベル・スイッチングRパッケージ(Papastamoulis, P., Journal of Statistical Software 69, (2015))から補正されたECRアルゴリズムのマルチスレッドバージョンを、迅速なラベルスイッチング補正のために使用した。最低50個のユニークな突然変異を含むSNVクラスターのみを下流分析のために保持した。このアルゴリズムへの入力は、マイクロダイセクションごとの各コールされたバリアントのカバレッジ及びカウントからなる患者ごとのデータ表を含んでいた。

【0122】

32人の患者のうち11人(34.3%)に及ぶ少数のクラスター(966人のうち13人、1.34%)は、分割不足であることが見出され、いかなる突然変異も共有しないマイクロダイセクションからなっていた。このような問題のあるマイクロダイセクションを分離して、更なる独立したクラスターを形成した。対応する組織学画像の詳細な検査によって、患者特異的な配列決定された材料の各セットについて、空間的に隣接又は近接するマイクロダイセクションのクラスタリングを検証した。

10

【0123】

クラスタリングアルゴリズムは、設計によって試料にわたるVAFの類似性に基づいて各突然変異を1つのクラスターに割り当てるので、FOXO1ホットスポット突然変異は、患者生検のいくつかのマイクロダイセクションにおいて起こり得る複数の独立した獲得事象の可能性を説明するために、最小数のクラスターに手動で再割り当てされた。具体的には、儉約的ホットスポット再割り当ては、ホットスポットを保有するマイクロダイセクション(及びそれらがメンバーであるSNVクラスター)を最初に同定することを必要とした。次に、最も系統発生的に親のクラスターの最小サブセットが同定され(系統樹の推論のための方法を参照のこと)、ここで、各々は、ホットスポットについての支持の少なくとも1つの生のリードを有する最大数のメンバーマイクロダイセクションを有し、その結果、全てのホットスポット保有マイクロダイセクションを説明した。重要なことに、いくつかのホットスポット保有マイクロダイセクションは、複数のクラスターのメンバーであったが、ホットスポットは、このような場合に最大数のホットスポット含有切断を有するクラスターにのみ割り当てられた。少数の症例について、候補ホットスポット保有クラスターのいくつかのメンバー切断は、低いカバレッジに起因してホットスポット突然変異のいかなる証拠も有しなかった。これらの場合、組織学画像をチェックして、問題の切断の空間的位置決めが、同じクラスター内のホットスポット保有切断に対して近位であるか遠位であるかを決定した。これらの候補切断がホットスポットを保有する可能性の更なる支持として、真正ホットスポット含有切断との突然変異の共有を評価した。

20

30

【0124】

系統樹の推論：

Nik-Zainal et al. (Nik-Zainal, S. et al. The life history of 21 breast cancers. Cell 149, 994-1007 (2012))を適用して、NDPアルゴリズムによって同定された患者ごとのSNVクラスター間の系統発生的クローン関係を推測した。ここで、各クラスターは系統樹の枝として表される。所与のクラスターは、突然変異のクラスターを保有する細胞の割合が、同じマイクロダイセクション内の突然変異の別のクラスターを含有する細胞の割合と比較して、全てのメンバーのマイクロダイセクションにおいてより低く、それらのそれぞれの突然変異細胞分画(cell fraction、CF)の合計もまた100%超である場合、別のクラスター内にネスティングされている(すなわち、サブクローン関係)強い証拠を有すると見なされる。そうでなければ、ペアワイズ突然変異体CFの合計が100%以下である場合、ネスティングの弱い証拠のみが存在する。いくつかのマイクロダイセクションのみが、別のものと比較して、所与のSNVクラスターのより低いCFを有する場合、クラスターは、独立しており、かつ互いにネスティングされていないと解釈される。本研究では、突然変異体CFが0.05超であるクラスターのみを分析し、一方、各SNVクラスターのCFは、各マイクロダイセクションについてのVAFの中央

40

50

値の2倍として計算し、これは二倍性を仮定する。

【0125】

同じクラスターの他のメンバー切断といかなる突然変異も共有しないマイクロダイセクションを含有する問題のあるSNVクラスターを、新しい独立したクラスターに分割し、鳩の巣原理を使用して同じ患者生検内の全ての他のクラスターとの系統発生的関連性について個々に再評価した。

【0126】

選択下での突然変異の同定：

任意のコード突然変異が非腫瘍性慢性肝疾患組織において選択下にあるかどうかを決定するために、遺伝子、タンパク質ドメイン、コドン及びホットスポットのレベルについて dN/DS_{cv}法 (Martincorena, I. et al., Cell 171, 1029 - 1041, e21 (2017)) を使用して、同義突然変異獲得率から予測される数と比較してより多い数の非同義突然変異を有する遺伝子を同定した。この分析のために、全てのSNVクラスターのセットにわたる固有の突然変異を入力として使用し、一方、0.05未満の多重仮説検定について補正されたq値を有する突然変異を選択下にあると見なした。選択下にあり得る非コード突然変異の同定のために、NBRアルゴリズムを使用した (Rheinbay, E. et al., Nature, 578, 102 - 111 (2020))。

【0127】

インデルコーリング：

以前に記載された Brunner et al (2019) のように、cgpPindel を用いてインデルコーリングを実行した (Raine, K. M. et al., Curr Protoc Bioinformatics 52, 15.7.1 - 15.7.12 (2015))。ナイーブベイズアルゴリズムを使用して、各コールされたインデルを、NDPアルゴリズムを使用して同定されたSNVクラスターに割り当てた。SNVコーリング中に行ったように、ベータ二項過剰分散フィルタを、各患者生検から作製された一連のマイクロダイセクションにわたる各コールされたインデルの生カウントに適用して、アーチファクトを更にフィルタリング除去し、ここで、0.1以上の過剰分散値及び0.025以上のVAFを有するバリエーションは、真である可能性が高いと考えられた。

【0128】

真正物のインデルの特徴のより良い理解を得るために、患者当たりの近接したマイクロダイセクション間で共有されるインデルのグラウンドトゥルスセットを、研究コホートに対応する組織学画像のセットの詳細なレビューを通して最初に同定した。次に、ホモポリマーの数、インデルが反復HG19領域で発生したかどうか、ベータ二項フィルタからの分散推定値、及び様々なcgpPindel VCFフィールド (すなわち、曖昧なリード及び未知のリードを除外するVAF、QUAL - アンカーリードのマッピング品質の合計を示す、カバレッジ、coverage_{alt} : バリエーションサポートリードの数、PindelからのSIスコア、LEN - 塩基対におけるバリエーション長、REP : 反復の数、NB - Tum : 標的試料において同様のインデル事象を示すマイナス鎖に対して一次ライナーによってマッピングされたリードの数、ND : 標的試料におけるインデルを伴う又は伴わない一次ライナーからのマイナス鎖マッピングリードの数、NP - Tum : 標的試料におけるマイナス鎖に対してPindelによってマッピングされたリードの数、NR - Tum : NP - Tum及びND - Tumの固有の和集合、NU - Tum : NP - Tum及びNB - Tumの固有の和集合、PB - Tum : 標的試料における類似のインデル事象を示すプラス鎖に対して一次ライナーによってマッピングされたリードの数、PD - Tum : 標的試料におけるインデルを伴う又は伴わない一次ライナーからのプラス鎖マッピングリードの数、PP - Tum : 標的試料におけるプラス鎖に対してPindelによってマッピングされたリードの数、PR - Tum : PP - Tum及びPD - Tumの固有の和集合、PU - Tum : PP - Tum及びPB - Tumの固有の和集合、AMB - Tum : 標的試料において同じ特異性を有する両方の対立遺伝子上にマッピングしてい

10

20

30

40

50

る曖昧なリード、UNK - Tum : バリエント位置にミスマッチを含み、最初のヒットとして参照にアラインしない未知のリード、VT : 挿入又は欠失のいずれかのバリエント型)を含むインデルのいくつかの特徴を、統計モデリングのために考慮した。

【0129】

真のセットにおける可能性が高い実際のインデルの典型的な属性を解明するために、これらのパラメータを、RにおけるXGBoostアルゴリズムへの入力として使用して(Chen, T. & Guestrin, C. XGBoost: A Scalable Tree Boosting System, dl.acm.org, 785-794 (ACM, 2016). doi: 10.1145/2939672.2939785)を使用して、一連の1,000個の勾配ブースティング回帰木(分類及び回帰木(classification and regression tree, CART)としても知られる)を構築した。最小グリッド探索ベースの手法を使用して手動ハイパーパラメータ調整を実行して、学習率() 0.003、リーフノードパーティショニングのための最小損失低減() 20、回帰木を学習するための訓練データにおけるインデルの60%及び共変量の80%のサブサンプリング率、並びに最大木深度3に到達した。更に、モデル学習を実行しつつ、バイナリロジスティック目的関数を最適化し、L2正則化係数() 及びバイアスを0.5に設定し、一つ抜き交差検証(leave-one-out-cross-validation)を可能にし、かつモデル訓練メトリックを設定して、受信者動作曲線下面積(area under the receiver operator curve, AUROC)を評価した。他の全てのハイパーパラメータは、ソフトウェアドキュメンテーションに従ってデフォルト値又は推奨値に設定した。モデル訓練プロセスの終わりに、バイナリアンサンブル分類器が構築され、そこでは、各々が低い予測力を有する多数の浅い回帰木からのスコアが組み合わされて、高い予測力を与える。いわゆるアンサンブルブースティングフレームワークでは、回帰木のセットは、ロジスティック損失関数が最小化されるように従属変数に直接適合するように訓練された最初の木(t = 0)を用いて加法的に訓練され(この場合、インデルが患者固有の真理セットの特徴に類似する特徴を有するか否かである)、一方、後続の木は、前の反復(t - 1)まで学習された全体モデルの残差に適合するように訓練される。次いで、各インデルについて、回帰木基本学習器からの一連のスコアを重み付けして組み合わせて、所与のインデルが真のバリエント又は人為的バリエントである可能性が高いかどうかの確率を表す最終スコアに到達した。0.7超のモデルスコアを使用して、ありそうな実際のインデルを同定した。回帰木アンサンブルモデルの予測精度は、モデル予測されたインデルとグラウンドトゥールズインデルとのセット間の比較から得られるAUROCを評価することによって、患者ごとに評価した。

【0130】

ブースティング回帰フォレストモデルでは、N個の木(インデルの予測特徴をモデルスコアにマッピングする基本学習器関数 $f_t(x_i)$)は、予測因子及び従属変数対を含む訓練データセットから学習され

【0131】

【数1】

$\{(x_i, y_i)\}_{i=1}^n$ 、

ここで、各木は、葉ノードスコアのベクトルと、インデル特徴値を葉ノードに割り当てるマッピング関数としてのその構造とを記憶し、一方、全体モデルスコア(予測

【0132】

【数2】

\hat{y}_i

)は、各木からのスコアの合計として計算される。

【 0 1 3 3 】

【 数 3 】

$$\hat{y}_i^{(N)} = \sum_{t=1}^N f_t(x_i) = \hat{y}_i^{(t-1)} + \epsilon f_t(x_i) \quad [1]$$

【 0 1 3 4 】

ここで ϵ は学習率であり、これは、訓練データへの過剰適合から保護するために、全体モデルスコアに対する各木の寄与をスケールする。

【 0 1 3 5 】

ブースティング回帰フォレストを構築するために、逐次木構造学習の加法プロセスが採用され、ここで $f_t(x_i)$ は、以下の目的関数を最小化する各訓練反復において学習される。

【 0 1 3 6 】

【 数 4 】

$$\text{obj}^{(t)} = \sum_{i=1}^n \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t) \quad [2]$$

$$g_i = \partial_{\hat{y}_i^{(t-1)}} \mathcal{L} = \frac{(1-y_i)e^{\hat{y}_i^{(t-1)}}}{1+e^{\hat{y}_i^{(t-1)}}} - \frac{y_i e^{-\hat{y}_i^{(t-1)}}}{1+e^{-\hat{y}_i^{(t-1)}}} \quad [3]$$

$$h_i = \partial_{\hat{y}_i^{(t-1)}}^2 \mathcal{L} = (1-y_i) \left[\frac{e^{\hat{y}_i^{(t-1)}}}{1+e^{\hat{y}_i^{(t-1)}}} - \frac{e^{2\hat{y}_i^{(t-1)}}}{(1+e^{\hat{y}_i^{(t-1)}})^2} \right] + y_i \left[\frac{e^{-\hat{y}_i^{(t-1)}}}{1+e^{-\hat{y}_i^{(t-1)}}} - \frac{e^{-2\hat{y}_i^{(t-1)}}}{(1+e^{-\hat{y}_i^{(t-1)}})^2} \right] \quad [4]$$

$$\mathcal{L} = y_i \ln(1 + e^{-\hat{y}_i^{(t-1)}}) + (1 - y_i) \ln(1 + e^{\hat{y}_i^{(t-1)}}) \quad [5]$$

$$\Omega(f_t) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2 \quad [6]$$

ここで、

$\text{Obj}^{(t)}$ は、回帰木 t について最適化される目的関数であり、

$f_t(x_i)$ は、特徴値 x_i のスコアを評価する回帰木 t を示し、

g_i は、前の反復からのモデルスコアに関する損失関数の一次導関数であり

【 0 1 3 7 】

【 数 5 】

$$\hat{y}_t^{(t-1)}$$

h_i は、前の反復からのモデルスコアに関する損失関数の二次導関数であり

【 0 1 3 8 】

【 数 6 】

$$\hat{y}_t^{(t-1)}$$

\mathcal{L} は、モデルが訓練データにどれだけよく適合するかを記述するロジスティック損失関数であり、

$\Omega(f_t)$ は、 T 個の葉と葉ノード j に割り当てられたインデルのスコアのベクトル j の L_2 ノルムの関数として定義される回帰木 t の複雑さを記述する正則化項である。

【 0 1 3 9 】

あり得る実際のインデル又は人為的なインデルの二項分類が目的であったので、この研究ではロジスティック損失関数を使用した。ロジスティック損失を近似するために、二次テイラー展開を使用する。したがって、各訓練反復は、 g_i 及び h_i の計算から開始する。

【 0 1 4 0 】

【 数 7 】

$$I_j = \{i | q(x_i) = j\}$$

10

を葉ノード j に割り当てられたインデルのセットとすると、

【 0 1 4 1 】

【 数 8 】

$$f_t(x_i) = \omega_j$$

を使用して、ノード j にインデル i を割り当てることとそのインデルについて j のスコアをもたらすことを示すことができる。これは、目的関数が何らかの代数操作後に以下の形式をとることを可能にする。

【 0 1 4 2 】

【 数 9 】

$$\text{obj}^{(t)} = \sum_{j=1}^T \left[G_j \omega_j + \frac{1}{2} (H_j + \lambda) \omega_j^2 \right] + \gamma T \quad [7]$$

20

$$G_j = \sum_{i \in I_j} g_i \quad [8]$$

$$H_j = \sum_{i \in I_j} h_i \quad [9]$$

$\text{obj}^{(t)} = 0$ を設定し、 j に関して一次導関数をとることにより、最適な葉ノード重みについて解くことを可能にする

【 0 1 4 3 】

【 数 1 0 】

$$\omega_j^*:$$

【 0 1 4 4 】

【 数 1 1 】

$$\omega_j^* = -\frac{G_j}{H_j + \lambda} \quad [10]$$

40

【 0 1 4 5 】

[1 0] を [7] に代入すると、所与の木構造の最適性の尺度として目的関数値 obj^* が与えられ、ここで、より低い値がより好ましい。

【 0 1 4 6 】

【 数 1 2 】

$$\text{obj}^* = -\frac{1}{2} \sum_{j=1}^T \frac{G_j^2}{H_j + \lambda} + \gamma T \quad [11]$$

50

【 0 1 4 7 】

任意の所与の木について無数の可能なノード構成が存在することが多いので、最適値を有する木構造を見つけるために、全ての可能な木構造について obj^* を計算することは実際的ではない。代わりに、貪欲法を採用して、深度 0 の最小木から開始して、一度に一分岐ずつ木を成長させた。次いで、成長中の回帰木 t 内の各葉ノードについて、以下の利得メトリックを計算し、これは、各新しい分割が導入された後の目的関数値の変化を評価する。このようにして、利得が最大になるように木を成長させた。

【 0 1 4 8 】

【 数 1 3 】

$$\text{利得} = \frac{1}{2} \left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda} \right] - \gamma \quad [12] \quad 10$$

【 0 1 4 9 】

ここで

【 0 1 5 0 】

【 数 1 4 】

$$\frac{G_L^2}{H_L + \lambda} \text{ 及び } \frac{G_R^2}{H_R + \lambda}$$

は、新たに追加された左及び右の子ノード上のスコアをそれぞれ示し、

20

【 0 1 5 1 】

【 数 1 5 】

$$\frac{(G_L + G_R)^2}{H_L + H_R + \lambda}$$

は、新しい分割が考慮された親ノード上のスコアを示す。

これらの 3 つの括弧内の項はともに、訓練損失低減を表す。最後に γ は、新しい分割を導入するときに木構造に複雑さを追加するコストを考慮する正則化項を表す。更に γ はまた、所与の回帰木の所与の葉ノードを分割するために必要な損失低減の最小量としても機能し、ここで、より大きい値は、新たな分岐点の追加がより少ない木をもたらすことになる。これは、以下の条件がより頻繁に真になるためである。

30

【 0 1 5 2 】

【 数 1 6 】

$$\frac{1}{2} \left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda} \right] < \gamma$$

これは、好ましくない負の利得をもたらすことになり、そのような場合には、分岐が最適な木構造をもたらすことになるので、発生すべきでないことを示す。このようにして、木の各葉ノードにおける計算された利得は、損失低減を最大化するためにどの新しい分岐が生成され得るかを決定するために使用できる。所与の葉ノードにおける特定の特徵に対する最適な分割を同定するために、グループ化の各対（すなわち、ある閾値よりも小さい又は大きい特徴値を有するインデル）に対する利得の合計を計算するために、インデルをそれらの特徴値に従ってグループ化するための閾値を系統的に探索することができる。これを行うことによって、最大損失低減、したがって最大利得をもたらす最適閾値を同定することが可能である。この手法を使用して、予め指定された木深度に到達するまで、モデルによって考慮される全てのインデル特徴について最適な分割が反復的に同定され、その時点で、次の訓練反復において別の木構造の学習に進む前に、[1] において定義されるような全体モデルに追加された最適な木構造 $f_t(x_t)$ が同定された。アルゴリズムは、所定数の木が学習されたときにのみ終了する。

40

【 0 1 5 3 】

50

グラウンドトゥールースセットが、カバレッジが低い遺伝子座に真に存在する近接して共有されるインデル、又は特定のマイクロダイセクションに対してプライベートであるインデルを捕捉しないことは注目に値する。このカテゴリーにおけるインデルは、グラウンドトゥールースバリエーションに類似する属性を有することが多いことが見出され、したがって、高いモデルスコアを有し、これは、ゲノムブラウザを使用する手動レビューを通して実際的である可能性が高いことが検証された。

【0154】

各モデル特徴の重要性は、シャープレイ (Shapley) 値を使用して評価された。シャープレイ値は、ゲーム理論に根ざした概念であり、複数の寄与者から所与の結果への寄与を、証明可能なほど公平に決定するための方法を提供する (Shapley, L. S. 10
.17. A Value for n -Person Games. Contributions to the Theory of Games (AM-28), Volume II 307-318 (Princeton University Press, 2016). doi: 10.1515/9781400881970-018/html 及び Lundberg, S. M. & Lee, S.-L. A Unified Approach to Interpreting Model Predictions. 47
65-4774 (2017) を参照されたい)。数値は、評価されている特徴を除いて、特徴グループ化の全ての可能な順列と一緒に説明するモデルのセットにわたる所与の予測特徴を有する及び有しないモデルスコアの差の加重平均に基づいて計算した。したがって、特徴のシャープレイ値は、モデルスコア全体に対するその特徴の値の平均寄与 (効果) 20
を表し、これは以下のように計算した。

【0155】

【数17】

$$\phi_i(p) = \sum_{S \subseteq N/i} \frac{|S|!(n-|S|-1)!}{n!} (p(S \cup i) - p(S)) \quad [13]$$

式中、

$\phi_i(p)$ は、モデル予測 p についての特徴 i のシャープレイ値であり、

$S \subseteq N/i$ は、特徴 i を除く、特徴グループ化の全ての可能なセット (S) を表し、

$|S|!$ は、セット S 内の特徴の順列の数を表し、

$(n - |S| - 1)!$ は、考慮される n 個の特徴の合計のうちセット S にはない特徴の順列の数を表し、

$p(S \cup i)$ は、特徴 i を有するモデル予測 p であり、

$p(S)$ は、特徴 i のないモデル予測 p である。

【0156】

回帰フォレストの場合、所与の予測特徴のシャープレイ値は、構成要素木のアンサンブルにおけるその特徴に対応するシャープレイ値の加重和であった。ここで、特に回帰木ベースのモデルのために設計された SHAP (SHapley Additive exP
lana tions) 実装が使用され、これは、アンサンブル内の各回帰木に存在する特徴のサブセットに基づいてシャープレイ値をより迅速に計算する (Lundberg & Lee, 2017) 40
。所与のインデルが実際的である可能性が高いか、人為的である可能性が高いかの特定の予測について、全体モデルスコアを、各予測特徴のシャープレイ値 (予測への寄与) と訓練データから計算されたベースライン (平均) モデルスコアとの和として分解した。インデルが実際的である可能性が高いか人為的である可能性が高いかを高度に予測する特徴の存在又は非存在は、モデルスコアの大きな変化をもたらす。そのような影響力のある特徴の高い値は、スコアに正又は負のいずれかで寄与し、典型的には 0 より大きい又は小さいシャープレイ値と関連し、それぞれ真又は偽のインデルの形質に対応する。逆に、全体モデルスコアに対する無視できる影響を有する特徴は、ほぼ 0 のシャープレイ値を有する。

【0157】

構造バリエーションコーリング：

大きなゲノムセグメントに影響を及ぼす欠失、逆位、縦列重複、及び転座を含む構造バリエーション (Structural variant, SV) が、BRASS (breakpoint variant assembly) アルゴリズムを使用してコールされた (Campbell, P. J. et al., Nat. Genet. 40, 722-729 (2008)) (<https://github.com/cancerit/BRASS>)。次に、3段階プロセスを使用して、BRASSと呼ばれる可能性のある人為的なSVをフィルタリングして除去した。第一に、各マイクロダイセクションにおいて検出されたSV事象を、対応する正常バルク対照試料のパネルに存在するものと比較することに基づいて、LCMライブラリー調製プロトコルによって導入された人為的変異体を同定し、除去するカスタムパイプラインを開発した。第二に、全ての残りのSVの詳細な手動レビューを、ゲノムブラウザ及びバリエーションアノテーションを使用して行った。最後に、SNV及びインデルが供されたサニティチェックと同様に、各SVの存在を、可能であれば近接するマイクロダイセクション間でチェックした。ここで、実際のバリエーションは、このような切断クラスターによって共有されることが予想される。

10

【0158】

より複雑なゲノム再配置事象、例えばクロモソリプシス (Stephens, P. J. et al., Cell 144, 27-40 (2011)) (1つ以上の染色体が数千もの断片に粉碎され、続いて無秩序に融合して戻される) を、SVの最終セットの詳細な手動レビューを通して更に同定した。

20

【0159】

クローン面積の計算：

1, 202個のマイクロダイセクションの正確な空間位置を、LMD7レーザーマイクロダイセクション内蔵カメラを用いて撮影した一連の顕微鏡画像において捕捉した。各切断の外縁のデカルト座標を、画像処理MATLABツールボックスからのエッジ関数において実施されるようなCanny法 (<https://uk.mathworks.com/matlabcentral/fileexchange/46859-canny-edge-detection>) を使用して抽出した。これにより、マイクロダイセクションごとにx-y座標のセットが得られ、それぞれのWGSプロファイルに対応するように手動でアノテーションした。次に、単一ヌクレオチドバリエーション (SNV) クラスターを、各クラスターに割り当てられた突然変異を有するそれらのそれぞれのメンバーのマイクロダイセクションの同定から開始して、個々にプロセッシングした。各SNVクラスターについて、突然変異体CF (すなわち、 $2 \times VAF$ 中央値) を使用して、組織切片画像上の各メンバー切断のx-y座標を調整して、遺伝子クローンサイズをより正確に反映させた。続いて、得られた多角形を、対応するクローン領域を表す単一の实体にマージする前に、所与のSNVクラスターの各メンバーマイクロダイセクションの調整された空間座標を包含するように、最小楕円凸包を描いた。画素からミクロンへの変換を適用して単位を平方ミクロンに変換する前に、クローン面積を最初に平方画素に関して計算した。このために、倍率変換係数は、組織切片の高解像度走査組織学画像の上に重ねられたスケールインジケータの画像を最初に生成することによって計算した。これは、Hamamatsu PhotonicsからのNDP.view2 NanoZoomer Digital Pathology スライドスキャナ画像ビューイングソフトウェアを使用して行った。次いで、スケールインジケータ画像を、magic画像処理パッケージ (<https://cran.r-project.org/web/packages/magic/index.html>) を使用してR統計プログラミング環境にロードして、各組織切片画像についてミリメートル当たりの正確なピクセル数を決定した。この研究では、 $VAF > 0.05$ の突然変異に寄与したマイクロダイセクションのみをクローンサイズ計算に含めた。

30

40

【0160】

対立遺伝子及びインデルのカウント

50

各塩基（A、C、G、T）の生対立遺伝子カウント及びフィルタリングされていないリードの総カバレッジを、最小マッピング及び塩基品質をそれぞれ30及び25に設定しながら、SBS型バリエーションのためのalleleCountソフトウェアを使用して列挙して、ベータ二項ベースのフィルタのための入力を生成した（<http://cancerit.github.io/alleleCount/>）。NDPクラスタリングへの入力及びインデルのカウントのために、cgpVAFソフトウェアを使用した（<https://github.com/cancerit/vafCorrect>）。

【0161】

対立遺伝子カウンターソフトウェアを使用して、LCMマイクロダイセクション及びバルク対照試料の両方のbamファイルから直接、各塩基のフィルタリングされていない生のカウント数を決定した。FOXO1 S22Wホットスポットドライバース突然変異を保有することが見出された各患者についてのこれらのカウントデータから積み重ねられたバープロットを作成して、クローン及びバルク対照（利用可能な場合）にわたるその分布を可視化した。

10

【0162】

クローンサイズの比較：

クローン面積（ μm^2 ）を、この研究において見出されたドライバース突然変異を保有する肝細胞と保有しない肝細胞との間で比較した。具体的には、各ドナー由来の変異クローンの数の間のクローン面積（クローンの変異数によって重み付けされた）を、対応する各ドナーから無作為に選択された野生型クローンの同等数と比較することができるように、各ドライバース突然変異について、ドライバース突然変異についてのクローン野生型を、変異を有する各ドナーから均一かつ無作為にサンプリングした。ggstatsplot Rパッケージを使用してクローン面積の比較を行い、p値に対する多重仮説検定補正のBonferroni法を適用し、ノンパラメトリック対比較のためのデフォルトMann-Whitney U検定を使用した（Patil, I., The Journal of Open Source Software 6, 1-5 (2021)）。

20

【0163】

肝臓質量推定：

この研究において同定されたドライバース突然変異の各々を保有する肝細胞のグラム数の計算において、いくつかの仮定を行った：（1）試料を構成する大多数の細胞種は肝細胞であり（Schulze, R. J. et al., J Cell Biol 218, 2096-2112 (2019)）、ここで、ドライバース突然変異は二倍体ゲノム領域において生じ、したがって、突然変異体肝細胞画分 = $2 \times$ ドライバース対立遺伝子頻度である；（2）各LCMマイクロダイセクションは、100～500個の肝細胞を含むと推定され、（3）典型的な1.5kgのヒト肝臓1g当たり 1.16×10^8 個の肝細胞が存在する（Lipscomb, J. C. et al., Toxicol Appl Pharmacol 152, 376-387 (1998)及びSohlenius-Sternbeck, A.-K., Toxicol In Vitro 20, 1582-1586 (2006)を参照されたい）。本発明者らの研究における各ドナーについて、突然変異した肝細胞の肝臓全体の質量（グラム）を、最初に組織学画像を使用して全ての配列決定されたLCM解剖の面積（ピクセル）を計算することによって、各ドライバース突然変異について推測した。各切断が100～500個の肝細胞を含むと推定されたので、RlinMap関数を使用して線形適合を実行して、この範囲内の全てのLCM切断面積をマッピングし、各LCM切断を構成する肝細胞の数を効果的に推定した。次いで、各ドライバース突然変異のバリエーション対立遺伝子頻度を使用して、各LCM切断における変異体保有細胞の割合を推測した。次に、各ドライバースを含有するドナー当たりの配列決定された材料の割合を、全てのドナー特異的配列決定LCM切断からの推定値を合計することによって計算した。次いで、これらのドナーレベル推定値を使用して、典型的なヒト肝臓における肝細胞の推定数に基づいて各ドライバースを保有する肝細胞の割合を概算した。次いで、これらの値を最終的に使用して、典型的なヒト肝臓の重量が1.5kgであると仮定して、各

30

40

50

ドナーについて各ドライバーを含有する肝臓のグラム数を推定した。

【0164】

突然変異シグネチャーの抽出にHDPを使用する：

HDPアルゴリズムは、HDP Rパッケージ (<https://github.com/nicolaroberts/hdp>) を使用して、正常肝臓及び慢性肝疾患試料において同定された1,013個のSNVクラスターのそれぞれにおいてコールされたSBSのセットを構成する突然変異シグネチャーを抽出した。アルゴリズムへの入力は、突然変異カテゴリーの各々についてのSNVクラスター当たりの突然変異カウントのマトリックスからなり、これは、この場合、6つのSBSタイプ(C>A、C>G、C>T、T>A、T>C、T>G)によって定義される192個のトリヌクレオチド突然変異コンテキスト(SigProfilerMatrixGeneratorソフトウェア(Bergstrom, E. N. et al. BMC Genomics 20, 1-12 (2019))を使用して生成された)からなり、各々は、転写及び非転写鎖について、突然変異塩基(3'及び5')に隣接する塩基(A、C、T、G)の全ての可能な組み合わせによって更に定義された。PanCancer Analysis of Whole Genomes (PCAWG) 研究からの65個の以前に同定された192個のコンテキストベースの突然変異シグネチャーの参照カタログを、以前の情報として使用した(Alexandrov, L. B. et al., Nature 578, 94-101 (2020))。既知の肝臓関連シグネチャーの抽出を容易にするために、肝細胞がん(HCC)試料において以前に観察されたシグネチャー(すなわち、SBS1、SBS3、SBS4、SBS5、SBS6、SBS9、SBS12、SBS14、SBS16、SBS17a、SBS17b、SBS18、SBS19、SBS22、SBS23、SBS24、SBS26、SBS28、SBS29、SBS30、SBS31、SBS35、SBS37、SBS40)に、分析中に1000の偽カウントのデフォルト重み付けを割り当てた。残りの先行シグネチャーを、分析においてそれらを完全に除外しないように、100のより低い重み付けを割り当てた。設計により、HDPにより、以前の情報として供給された既知のシグネチャーのセットとは異なる新規の突然変異シグネチャーについてのある程度のデノボ発見が可能になる。肝臓関連突然変異シグネチャーの抽出を更に導くために、314個のHCC WGSプロファイルも分析に含めた。100,000回の反復のバーンインを使用し、続いて、500回の事後ギブスサンプリング反復を500回の反復間隔で実行し、一方で、クラスター融合対分割の程度(それぞれ、より低い値対より高い値)を制御する濃度パラメータを各反復で合計5回調整し、突然変異が最初にランダムに割り当てられる70個のクラスターから開始した。独立した事後サンプリングの仮定に違反する可能性を最小限に抑えるために、事後試料の広い収集間隔と組み合わせた長いバーンインが選択された。更に、70個の初期クラスターを使用して、突然変異の開始分布が65個全ての以前の参照シグネチャーにわたって広がり、いくつかの追加クラスターを加えて、所与の先行物のセットを超えて新規突然変異シグネチャーの抽出を促進することを確実にした。各反復において、各突然変異は、同じ突然変異カテゴリー、試料、又は親ノードにおいて高い割合の突然変異を有するクラスターに割り当てられる。コサイン類似度が0.9超であるクラスターは、デフォルト設定に従ってマージされ、一方、不確実なクラスターメンバーシップに起因して抽出されたシグネチャーのセットに割り当てられていない残りの突然変異は、結果として生じるモデルによって説明されないデータのパーセンテージを表すために一緒にグループ化される。0.8超のコサイン類似度(phiEntropy Rパッケージ(Drost, H. - G., J. Open Source Softw. 3, 765 (2018))を手動検査とともに使用して、抽出されたシグネチャーのうちのいずれかが既知の先行物のいずれかと一致するかどうかを判定し、わずかに低い類似度閾値を使用して、基準シグネチャーの可能な変動を説明した。摂動(Perturbation)モデル(Qiao, W. et al., PLoS Comput Biol 8, e1002838 (2012))として知られる計算デコンボリューション法を使用して、既知のシグネチャーと抽出されたシグネチャーとの間の類似性の二次尺度として、HDP抽

出シグネチャーのそれぞれを構成するPCA WG突然変異シグネチャーの寄与パーセントを推定した。任意の以前のシグネチャーとの密接な一致を妥当な確実性を伴って割り当てることができない程十分に一意であった抽出シグネチャーは、新規であると見なされた。この分析のために、6つの独立した事後サンプリングチェーンを同時に実行して、全ての突然変異について安定なクラスター割り当てへの収束を測定し、ここで、1000000、2000000、3000000、4000000、5000000及び6000000のランダムシードをそれぞれ割り当てた。シグネチャー抽出に使用される濃度パラメータ設定を含む全体的なHDPノード構造を、図17において概説する。

【0165】

図16には、本発明者らによって注目された新しい突然変異シグネチャーの詳細が示されている。図16Bには、同じ肝臓試料内の近くのクローン間の活動の変動性が示されている。

10

【0166】

患者を、数年間にわたって追跡した。特定の患者については、切除試料を1年間にわたって数回採取した。図16Cは、5年の期間にわたる1人の患者からの3つの肝臓切除試料を示す。各生検における突然変異の更なる詳細を図16D、図16E及び図16Fに示す。

【0167】

図16G、図16H及び図16Iには、正常肝細胞、ARLD罹患細胞、NAFLD罹患細胞の試料における、及び8個全ての解剖学的セグメントがサンプリングされた2人のNAFLD患者におけるシグネチャーの分布が示されている。

20

【0168】

本発明者らは、1人のNAFLD患者において、シグネチャーは、最初の肝臓切除標本には存在しなかったが、5年の研究にわたって強度が次第に増加したことに注目した(図16D~図16F)。したがって、このシグネチャーは、肝疾患が進行するにつれて次第により顕著になり、異物代謝障害の悪化に関連している可能性がある。

【0169】

突然変異シグネチャーの抽出にSigProfilerを使用する：

SigProfilerExtractor pythonパッケージ(Alexandrov, L. B. et al., Nature 578, 94-101 (2020)) (<https://github.com/AlexandrovLab/SigProfilerExtractor>)は、非負行列因子分解アルゴリズムに基づき、突然変異シグネチャー同定のための代替手段として役立った。アルゴリズムは、15個の突然変異シグネチャーを同定し、1,000回の反復で実行するように構成された。HDP及びSigProfiler抽出された192個のトリヌクレオチドコンテキストシグネチャーの比較を、コサイン類似性メトリックを評価することによって実行し、ここで、0.8超の値は、シグネチャーの所与の対が互いに同じ又はわずかに異なるバージョンであることを示すと見なした。

30

【0170】

FOXO1-eGFP画像化及びハイコンテンツ解析：

Hoechst 33342及びSPY-555-アクチン(Spirochrome)で対比染色したFOXO1-eGFP発現細胞のハイコンテンツ及び生細胞分析を、20x空気NA=0.4対物レンズを使用してOperetta CLSシステムで行った。固定した細胞の画像を、Harmonyソフトウェア(PerkinElmer)を用いて分析した。完全な分析配列を得た。簡単に説明すると、非細胞性物質(例えば、カバースリップの縁によって生じた明るい領域)を除去した。核をDAPI蛍光からセグメント化した。各核の周囲からの9ピクセル幅の細胞質環を、GFP蛍光からセグメント化し、バックグラウンド領域を、任意の核から120~150ピクセル離れた任意の無細胞領域からサンプリングした。核面積、真円度、及び幅：長さの比に閾値を設定することによって、断片又は他の非細胞性の小さな物体から核をフィルタリングした。平均核、細胞質

40

50

及びバックグラウンド GFP 蛍光強度を測定し、これらから、バックグラウンドを差し引いた値を用いて各セルについて核細胞質比を計算した。これらの値の \log_{10} を取った。

【0171】

生細胞については、Cell Profiler (McQuin, C. et al., PLOS Biology 16, e2005970 (2018)) を用いて同様の分析を行った。完全な分析配列を得た。簡単に述べると、照明補正画像を、多項式適合によって GFP 及び Hoechst チャンネルの両方について計算し、減算した。Hoechst 画像から核をセグメント化した。細胞質を GFP シグナルからセグメント化し、核の周りの 9 ピクセル幅のリングを使用して測定を核周囲領域に限定した。平均核及び細胞質 GFP 強度を測定し、核を時系列を通して追跡した。核：細胞質比を計算し、これらの値の \log_{10} を取った。

10

【0172】

生細胞からの結果は、4つのウェルからプールされたデータの中央値 ± 分散として表示され、それぞれ8つの視野で、ウェル当たり1000~2000個の細胞が分析され、条件当たり合計6000~7500個の細胞が得られる。

【0173】

免疫プロットティングによるタンパク質発現；

SDS-PAGEゲル上での免疫プロットティングを、以下の抗体：抗 - アクチン (Sigma, A5441, 1:5000)；抗 Akt (Cell Signaling, 2938, 1:1000)；抗ホスホ - Akt (T308) (Cell Signaling, 4056, 1:1000)；抗 GFP (Abcam, ab6556, 1:1000)；抗 FOXO1 (Cell Signaling, 2880, 1:1000)；抗ホスホ FOXO1 (T24) (Cell Signaling (9464, 1:1000)) を使用して、Hoare et al. (Hoare et al., Nat. 18, 979-992 (2016)) によって報告されたように実行した。

20

【0174】

メタボロミクス；

野生型 FOXO1-eGFP 又は FOXO1^{S22W} のいずれかを発現する HepG2 細胞を無血清培地中で一晚培養した後、100 nM インスリンの存在下又は非存在下で3時間刺激してから回収した。細胞を PBS で洗浄した後、ドライアイス上で50%メタノール、30%アセトニトリル (両方とも Fisher)、20%超純水及び5 µM バリン d 8 (内部対照、CK 同位体) 中で抽出及び溶解した。次いで、細胞溶解物上清を、LC-MS ベースのメタボロミクスを実行するまで -80 °C で保存した。

30

【0175】

代謝産物の HILIC クロマトグラフィー分離は、二成分溶媒系を有する 2.1 x 20 mm ガードカラム (両方とも粒径 5 mm) を備えた Millipore Sequant ZIC-pHILIC 分析カラム (5 µm, 2.1 x 150 mm) を使用して達成した。溶媒 A は、20 mM 炭酸アンモニウム、0.05% 水酸化アンモニウムであった。溶媒 B はアセトニトリルであった。カラムオープン及びオートサンプレートイをそれぞれ40及び4に保持した。クロマトグラフィー勾配を、以下のように0.200 mL / 分の流速で流した：0~2分：80% B；2~17分：80% B から20% B への直線勾配；17~17.1分：20% B から80% B への直線勾配；17.1~22.5分：80% B で保持。試料を無作為化し、LC-MS を用いて盲検様式で分析し、注入量は5 µl であった。プール試料を、全ての個々の試料の等しい混合物から生成し、品質管理として試料配列内に規則的な間隔で散在させて分析した。

40

【0176】

代謝産物を、Dionex Ultimate 3000 UHPLC に連結した Thermo Scientific Q Exactive Hybrid Quadrupole-Orbitrap 質量分析計 (HRMS) で測定した。質量分析計は、スプレ

50

ー電圧を + 4 . 5 k V / - 3 . 5 k V に設定し、加熱キャピラリーを 3 2 0 に保持し、補助ガスヒーターを 2 8 0 に保持して、フルスキャン、極性切り替えモードで操作した。シースガス流量を 2 5 単位に設定し、補助ガス流量を 1 5 単位に設定し、スイープガス流量を 0 単位に設定した。HRMS データ取得は、 $m/z = 70 \sim 900$ の範囲で行い、分解能は 7 0 , 0 0 0 に設定し、AGC 標的は 1×10^6 に設定し、最大注入時間 (maximum injection time、Max IT) は 1 2 0 m s に設定した。代謝産物の同一性を、2 つのパラメータを使用して確認した：(1) 前駆イオン m/z は、化学式によって予測される理論質量の 5 ppm 以内で一致した；(2) 代謝産物の保持時間は、同じクロマトグラフィー法で行った精製標準物質の保持時間の 5 % 以内であった。Thermo Fisher ソフトウェア Tracefinder 5 . 0 を使用して、クロマトグラムレビュー及びピーク面積積分を実行し、検出された各代謝産物のピーク面積を、その試料の総イオンカウント (total ion count、TIC) に対して正規化して、機器分析を通して試料操作から導入された任意の変動を補正した。正規化された面積を、更なる統計データ分析のための変数として使用した。

10

【 0 1 7 7 】

メタボロミクスデータの統計解析は、予測変数としてインスリン (あり又はなし) 及び FOXO 1 状態 (突然変異体又は野生型)、並びに従属変数として正規化代謝産物レベルを用いた線形モデルを使用して実行した。尤度比検定を使用して p 値を生成し、次いでこれを、Benjamini - Hochberg 法を使用して多重仮説検定について補正した。有意性については $q < 0 . 0 1$ の閾値を用いた。

20

【 0 1 7 8 】

HepG2 細胞株：RNA 配列決定データ前処理：

HepG2 細胞株試料 ($n = 30$) を、Illumina HiSeq 4 0 0 0 プラットフォームを使用して 2 レーンの 1 5 0 塩基対ペアエンド RNA 配列決定に供した。使用したヒト参照ゲノムは、Ensembl リリース 7 5 データに基づく遺伝子アノテーションを有する、1 0 0 0 Genomes Project からの hs37d5 であった。アダプター及び低品質リードを、Trim Galore (<https://github.com/FelixKrueger/TrimGalore>) を、以下のパラメータ：- q 2 0 - fastqc - ペアードストリンジェンシー 1 - 長さ 2 0 - e 0 . 1 で使用した。Spliced Transcripts Alignment to a Reference (STAR) aligner を使用して、生の配列リードを GRCh37 (hg19) ヒト参照ゲノムにマッピングした (Dobin, A. et al., Bioinformatics 2 9 , 1 5 - 2 1 (2 0 1 3))。置換が、Haplotype Caller を使用してコールされた (Poplin, R. et al., bioRxiv 1 - 2 2 (2 0 1 8) . doi : 1 0 . 1 1 0 1 / 2 0 1 1 7 8)。特徴カウントソフトウェア (Liao, Y., et al., Bioinformatics 3 0 , 9 2 3 - 9 3 0 (2 0 1 4)) を使用して、遺伝子発現値を要約し、一方、EdgeR R パッケージからの cpm 関数を使用して、データを対数カウント・パー・ミリオンスケールに正規化した (Robinson, M. D. et al., Bioinformatics 2 6 , 1 3 9 - 1 4 0 (2 0 1 0))。全てのヒートマップは、pheatmap R パッケージ (<https://cran.r-project.org/web/packages/pheatmap/jindex.html>) を使用して生成した。

30

40

【 0 1 7 9 】

遺伝子セット濃縮分析：

遺伝子セット濃縮分析 (GSEA v 3 . 0) (Subramanian, A. et al., PNAS 1 0 2 , 1 5 5 4 5 - 1 5 5 5 0 (2 0 0 5)) を、遺伝子の事前にランク付けされた遺伝子リスト、2 0 0 0 の順列、並びに多くて 5 0 0 の遺伝子を有する全ての遺伝子オンロジー及びリアクトーム関連遺伝子セット (2 0 2 1 年 6 月 1 日バージョン、http://download.baderlab.org/EM_Genesets/ からダウンロード) を使用して実行した。具体的には、各遺伝子について、2 つ

50

の線形モデルを、R統計プログラミング環境においてIm関数を使用して構築し、一方のモデルは、独立変数としてFOXO1ドライバー及びインスリン状態（すなわち、存在又は非存在のいずれか）の両方を有し、他方のモデルはインスリン状態のみを含んだ。両方のモデルにおける従属変数は、モデルにおける遺伝子の発現である。次に、尤度比検定を使用して、遺伝子当たりのネスティングされたモデルの各対間のp値を計算した。続いて、このp値に、各遺伝子のドライバーを有するモデルにおける突然変異状態の回帰係数の符号を乗じた。最後に、遺伝子リストをp値のこのセットに従って次のようにランク付けした： - (0) . . . - 0 . 0 5 . . . - 0 . 9 9 . . . 0 . 9 9 . . . 0 . 0 5 . . . 0 において、リストの最下部の遺伝子は、インスリン状態の影響を説明しながら、FOXO1ドライバーの存在と最も関連すると予想される。

10

【0180】

実施例1：肝臓試料におけるドライバー突然変異の同定

Brunner et al (Nature, 2019, 574 (7779) : 538 - 542) は、健康な肝臓及び罹患した肝臓からの482個の全ゲノムを配列決定したが、選択圧下での遺伝子の明確な同定のための統計的検出力を欠いていた。Brunnerらによって使用された方法よりもドライバー突然変異（すなわち、選択的利点を与える突然変異）を検出するための力を改善するために、Brunnerらからのデータを、20個の肝臓試料からの追加の1108個の全ゲノム配列とともに使用した。これらの試料は、主にNAFLDを有する患者からのものであったが、疾患重症度によって異なっていた。図1に示す階層的な実験計画を使用して試料を分析した。簡単に説明すると、約1cm³の肝臓組織試料ごとに、21～52個の別々のマイクロダイセクションを捕捉し、そのそれぞれが100～500個の肝細胞を含んでいた。NAFLDを有する2人の患者において、それらの外植された肝臓の8つ全てのクイノー解剖学的セグメントから別個の試料を採取し、各セグメントからの22～28個のマイクロダイセクションを配列決定した。この階層的実験設計は、患者当たりの複数のゲノムから生じる非独立性を管理するために特殊化した統計分析を使用して、個々のクローン、解剖学的セグメント及び患者にわたる体細胞突然変異の不均一性の定量を可能にした。

20

【0181】

Brunnerらからのデータと組み合わせた場合、データセットは、事前のネオアジュバント療法を受けていない5人の正常肝臓対照、アルコール関連肝疾患（ARLD）を有する10人及びNAFLDを有する19人を含む、34人の肝臓試料からの1590個のゲノムを含んでいた。ARLD又はNAFLDを有する全ての患者は、HCC、肝不全又は両方を有し、組織は、肝臓切除又は移植に由来した。全体として、9つの試料は、同期性HCC及び根底にある肝硬変を有する患者からのものであった。5年の期間にわたる1人の患者からの3つの肝臓切除試料を含む、更なる8つの試料が、根底にある肝硬変を伴わないHCCを有していた。全ての試料は、専門の肝臓病理学者による中央組織学的レビューを受け、患者の組織学的及び臨床的特徴は、根底にある疾患プロセスについて予想されるものと一致した。マイクロダイセクションを31xの平均深度まで配列決定した。

30

【0182】

正の選択下でタンパクコード突然変異を同定するために、dNdScvアルゴリズム (Martinocorena et al., Cell, 2017, 171, 1029 - 1041) を用いて、遺伝子中の有意に過剰な非同義突然変異についてスクリーニングした。

40

【0183】

非コードドライバー突然変異を検索するために、NBRアルゴリズム (Reinba et al., Nature, 2020, 578, 102 - 111) を使用して、プロモーター、エンハンサー、5' - UTR、3' - UTR長鎖非コードRNA遺伝子及びマイクロRNAをスクリーニングした。

【0184】

結果：

50

全てのタンパク質コード遺伝子にわたって、遺伝子 ACVR2A、CLCN5、FOXO1、CIDEB、ALB、GPAM及びTNRC6B遺伝子は、多重仮説検定の補正後に突然変異の有意な再発を有することが見出された ($q < 0.05$ 、すなわち、FOXO1 ($q < 2 \times 10^{-16}$)、CIDEB ($q < 2 \times 10^{-16}$)、ACVR2A $q = 7 \times 10^{-9}$)、ALB (8×10^{-10})、GPAM $q = 1 \times 10^{-5}$) 及び TNRC6B ($q = 0.04$) (MartincorenaCell et al. 171, 1029 - 1041 (2017))。非コード突然変異がNEAT1において同定された。突然変異遺伝子の詳細を以下に示す。

【0185】

a) コード突然変異：

FOXO1：

FOXO1遺伝子において同定された突然変異は、非常に有意で過剰なミスセンス突然変異であることが見出された ($q < 2 \times 10^{-16}$)。全体として、独立したFOXO1突然変異を獲得した26個の別々のクローンが同定された。これらは、8人の患者からの45個の個々のマイクロダイセクション間に分布していた。これらのうち、24個のクローンは、S22Wアミノ酸置換を生じると予測される同一の塩基変化を含んでいた (図2A)。他の2つの突然変異は、R21L置換及びS22*ナンセンス突然変異を生成するであろう。他の2つの突然変異は、R21L置換及びS22*ナンセンス突然変異を生じると予測された。後者は、正常な肝臓試料からの単一のマイクロダイセクションにおけるものであり、その生物学的重要性は不明である。S22W突然変異は、ARLD又はNAFLDを有する患者においてのみ見られた。FOXO1遺伝子の第1のイントロン内の体細胞構造バリエーションが、NAFLDを有する患者からのマイクロダイセクションにおいて観察された。これは、第13染色体に影響を及ぼす染色体異常事象の状況において生じた (図2D)。

【0186】

FOXO1 S22W突然変異を有する7人の患者のうち、6人は、異なるクローンにおける突然変異の複数の独立した獲得についての明確な証拠 (すなわち、収束進化) を有した (図3C)。

【0187】

肝臓の8つ全ての解剖学的セグメントがサンプリングされた2人の患者において、FOXO1 S22W突然変異が、一方の患者において3つの異なるセグメントで、他方において4つのセグメントで見出された。更に、単一のセグメント内であっても、複数の独立して獲得されたFOXO1突然変異が存在し、その結果、これらの2人の患者のうちの1人は、サンプリングされた領域の中にFOXO1 S22Wを有する9個の独立したクローンを有した。

【0188】

単一セグメントからの試料を分析した患者のうち、1人は、約1 cm³以内のFOXO1ミスセンス突然変異の5回の独立した獲得を有した。更なる患者は3回の別個の発生を有し、2人の患者は2回の別個の獲得を有した。

【0189】

これらの患者の系統樹を対応する肝臓試料の組織学上に階層化すると、平行進化を介して獲得されたFOXO1突然変異が、典型的にはFOXO1の野生型の介在領域によって分離された、地理的に異なる領域に位置することが明らかになった (図3A ~ 図3C)。

【0190】

CIDEB：

肝臓及び脂肪組織内の脂質代謝に関する3つのCIDEファミリー遺伝子のうちの1つであるCIDEB内に有意に過剰な体細胞突然変異が存在することが観察された ($q < 2 \times 10^{-16}$)。CIDEBは、肝細胞において活性な主要なファミリーメンバーであり、ノックアウトマウスモデルは、食事性脂肪性肝炎に対する抵抗性及びインスリン感受性の増加を示す (Liet al. Diabetes, 2007, 56, 2523 - 25

10

20

30

40

50

32。これらのタンパク質は、異なる液滴上のCIDEタンパク質間のホモ二量体の形成によって媒介される、細胞内脂肪液滴の融合を調節する(Schulze, R.J. et al., J Cell Biol 218, 2096-2112 (2019) 及び Lipscomb, J.C. et al., Toxicol Appl Pharmacol 152, 376-387 (1998))。ホモ二量体化は、一方の脂肪液滴からのCIDEタンパク質上の正に荷電した残基と他方の負に荷電した残基との間の静電接触を介して起こる。

【0191】

CIDE Bでは、2つのナンセンス突然変異及び1つの終止欠失突然変異に加えて、18個のミスセンス突然変異が観察された(図4A)。ミスセンス突然変異がCIDEタンパク質のホモ二量体化に關与する2つのドメインに主に位置し、それらの多くが荷電残基を中性残基に切り替える(R45W、R45Q、K140N、R144P)か、又は電荷を逆転させる(D42H、K62E、E78K)かのいずれかであった。以前のインビトロ突然変異誘発研究は、本発明者らの患者において突然変異した残基の一部を含む、これらの重要な保存残基上の電荷を変化させることが、ホモ二量体化を無効にし、細胞内での脂肪液滴の融合及び増殖を妨げることを示した(Schulze, R.J. et al., J Cell Biol 218, 2096-2112 (2019) 及び Lipscomb, J.C. et al., Toxicol Appl Pharmacol, 1998, 152, 376-387)。

【0192】

FOXO1について見られるように、CIDE Bにおける突然変異は、1人の患者の肝臓内の複数の独立したクローンにおいて頻繁に獲得された。例えば、肝臓の8つのクローネーセグメント全てがサンプリングされたNAFLD患者の1人において、CIDE Bにおいて非同義突然変異を有する14個のクローンが見出され、8つのセグメントのうち6つに分布していた(図4B~図4C)。

【0193】

G P A M :

別の有意に変異した遺伝子は、ミトコンドリアグリセロール-3-リン酸アシルトランスフェラーゼをコードするG P A Mであった。この酵素は、トリアシルグリセロール合成における律速ステップ、すなわちグリセロール-3-リン酸による長鎖アシル-CoAのエステル化を触媒する(Bergstrom, E.N. et al., BMC Genomics 20, 1-12 (2019) 及び Alexandrov, L.B. et al., Nature 578, 94-101 (2020) を参照されたい)。他のグリセロール-3-リン酸アシルトランスフェラーゼイソ酵素と比較して、G P A Mは、グルコースから合成された遊離脂肪酸を貯蔵トリグリセリドに組み込む、肝臓におけるデノボ脂質生成に特に重要である(Alexandrov, L.B. et al., Nature 578, 94-101 (2020))。この機能と一致して、G P A M転写は、インスリンシグナル伝達によって急速に上方制御される(Ericsson et al., J. Biol. Chem. 1997 (272): 7298-7305)。

【0194】

12個のミスセンス変異及び3個のタンパク質短縮型変異がG P A Mにおいて観察され($q = 1 \times 10^{-5}$)、7人の患者に影響を及ぼした(図5A)。1つのマイクロダイセクションにおいて遺伝子の20 kb上流から始まる縦列重複(不確定な重要性)(図5B)もまた観察された。ナンセンス及びフレームシフト突然変異と組み合わせた突然変異のホットスポット残基の不在は、これらの突然変異の可能性のある結果がタンパク質機能の障害であることを示唆している。FOXO1及びCIDE B突然変異に関しては、同じ患者の肝臓試料からの異なるクローンにおけるG P A Mにおける体細胞突然変異の収束獲得の証拠があった。例えば、1人の患者は、G P A Mに影響を及ぼす7つの別個の事象を有していたが(図5C)、別の患者は、2つの別個の事象を有していた(図5D及び図5E)。

10

20

30

40

50

【0195】

ACVR2A :

突然変異をコードするタンパク質を含む4つの遺伝子のうちの1つは、ACVR2Aであった。この遺伝子は、既知のHCC遺伝子に限定されることが報告されているTGF-スーパーファミリー中のアクチビン-Aの受容体をコードする(Brunner et al. (2019))。ACVR2A遺伝子は、HCCの5~10%において突然変異しており、タンパク質短縮事象の優勢は、それが腫瘍抑制遺伝子として作用することを示唆している。最近のメタ分析は、ACVR2Aにおける突然変異が、ウイルスHCCよりもNAFLD又はARLD関連HCCの症例においてより頻繁に見られることを示唆している(Chaudhary et al., Clin. Cancer Res. 2019 (25) : 463 - 472)。本研究では、ACVR2Aにおける13個のミスセンス突然変異、2個のナンセンス突然変異及び1個のスプライス部位インデル($q = 7 \times 10^{-9}$)、並びに構造変異による4個の大規模欠失が観察された(図6A及び図6B)。ARLDを有する1人の患者において、2つの独立したミスセンス点突然変異及び独立したクローンにおけるACVR2Aにおける大規模欠失が見られ、これは、この患者内のドライバー突然変異の収束進化と一致した。

10

【0196】

TNRC6B :

TNRC6Bは、マイクロRNAプロセッシングに關与するタンパクをコードする(Meister et al. Curr. 2005 (15) : 2149 - 2155)。TNRC6Bにおいて3つのナンセンス、2つの必須スプライス部位及び1つの大きなインフレーム欠失、並びに3つのミスセンス突然変異が観察された($q = 0.04$; 図9)。タンパク質切断変異体のこの優勢は、遺伝子の不活性化が肝細胞に正の選択的利点を付与することを示唆する。驚くべきことに、NAFLDを有する1人の患者は、TNRC6Bにおいて5つの異なる突然変異を有し、これは、独立した肝実質細胞クローンにおける収束進化と一致した。

20

【0197】

CLCN5 :

CLCN5遺伝子はクロライドチャンネルをコードし、これはX連鎖腎石灰化症を引き起こすが、生殖系列において突然変異した場合に既知の肝臓表現型を引き起こさない。CLCN5において3つのミスセンス及び4つのナンセンス変異($q = 0.04$)が観察され、患者全体に分布していた(図7)。これには、2人の健康な対照が含まれていた。したがって、CLCN5突然変異の慢性肝疾患への関連性は不明である。

30

【0198】

b) 非コード突然変異 :

長鎖非コードRNAであるNEAT1は、多重仮説検定に対する補正後のバックグラウンド期待値と比較して有意に過剰な突然変異を示した($q < 1 \times 10^{-10}$; 図8)。この遺伝子は、HCCを含むヒトがんの範囲において反復的に突然変異するが、これはおそらく、正の選択下にあるのではなく、ゲノムの超変異性領域内に存在するためである。重要なことに、HCC及び異形成結節において一般的であるTERTプロモーターに影響を及ぼす標準的ホットスポット突然変異又は構造バリエーションはいずれもが存在せず、一方、同じプロトコルを使用して配列決定された同期HCCの11/22(50%)がこれらのドライバーを有していた。

40

【0199】

実施例2 : FOXO1突然変異の評価 :

FOXO1は、インスリンシグナル伝達の下流の重要な転写因子である。絶食状態では、インスリンなしで、FOXO1は肝細胞の核において活性であり、糖新生、解糖及び脂肪分解経路における遺伝子の発現を上方制御する。インスリンがその受容体に結合すると、AKTはPI3Kを介して活性化される。続いて、AKTは核内のFOXO1をリン酸化し、T24のスレオニン3つの既知のAKTリン酸化標的の1つである。慢性肝疾患

50

において本発明者らによって観察された突然変異は、進化を通じて高度に保存されている R 2 1 及び S 2 2 アミノ酸に影響を及ぼし（図 2 A）、R S x T x P モチーフにおいて最初の 2 つを形成する。スレオニン T 2 4 がリン酸化されると、モチーフは 1 4 - 3 - 3 タンパク質の標準的な結合部位となり、F O X O 1 を核から細胞質へと輸出し、そこでタンパク質はユビキチン化及びプロテアソーム分解を受ける。S 2 2 残基は、それ自体が A M P K によってリン酸化され、2 つの機構：A K T による T 2 4 リン酸化の低減及び立体障害による 1 4 - 3 - 3 結合の遮断によって F O X O 1 の輸出を阻害する（S a l i n e e t a l . J . B i o l . C h e m . 2 0 1 9 , 2 9 4 , 1 3 1 0 6 - 1 3 1 1 6 ）。本発明者らは、S 2 2 を嵩高いトリプトファンで置換すると、F O X O 1 の核外輸送が同様に阻害されると仮定する。

10

【 0 2 0 0 】

この仮説を評価するために、H e p G 2（図 1 0 A、図 1 0 B、図 1 0 D、図 1 0 E）、H e p 3 B（図 1 0 F）及び P L C / P R F / 5（図 1 0 G）H C C 細胞株を、以下に記載されるように、C 末端緑色蛍光タンパク質（green fluorescent protein、G F P）に融合された野生型、R 2 1 L 又は S 2 2 W 変異体を含む F O X O 1 のレトロウイルス構築物で形質導入した。

【 0 2 0 1 】

H e p G 2、H e p 3 B 及び P L C / P R F / 5 細胞は A T C C から入手し、5 % C O₂ 大気中のダルベッコ改変イーグル培地（Dulbecco's modified Eagle's medium、D M E M）/ 1 0 % ウシ胎児血清（F C S）中で培養した。細胞同一性を、S T R（短い縦列リピート）遺伝子型決定によって確認した。細胞をマイコプラズマ汚染について定期的に試験し、常に陰性であることが見出された。インスリン（S i g m a）刺激は、1 0 0 n M の最終濃度で 1 5 分間インスリンを添加する前に、無血清 D M E M 中で 1 6 時間細胞を培養することによって実行した。

20

【 0 2 0 2 】

細胞を、V e c t o r B u i l d e r I n c . から購入した野生型 F O X O 1（N M _ 0 0 2 0 1 5 . 4）（V B 1 9 0 7 0 9 - 1 0 3 0 p w k）、F O X O 1^{R 2 1 L}（V B 1 9 0 7 0 9 - 1 0 2 8 b j m）又は F O X O 1^{S 2 2 W}（V B 1 9 0 7 0 9 - 1 0 3 2 n w a）を含有するレトロウイルスベクター（p M S C V - h F O X O 1 - e G F P : P 2 A : ピューロマイシン）で形質導入した。

30

【 0 2 0 3 】

結果：

血清飢餓下では、野生型及び変異体 F O X O 1 - G F P の両方が、インスリンシグナル伝達の非存在下で予想されるように、主に核に局在した。インスリン又は血清の添加により、野生型 F O X O 1 - G F P で形質導入された細胞は、インスリンの既知の効果と一致して、タンパク質の迅速な核外輸送を示した。しかしながら、インスリン又は血清の存在下でさえ、変異体 F O X O 1 - G F P で形質導入された細胞は、高い核 - 細胞質比で、実質的なレベルの核タンパク質を維持した。F O X O 1 中のリン酸化 T 2 4 に対する抗体は、突然変異体構築物への結合を示さなかった（図 1 0 H）。

【 0 2 0 4 】

1 0 5 個の代謝産物のレベルを、野生型又は S 2 2 W F O X O 1 - G F P を有する H e p G 2 細胞（インスリンあり及びなし）について、5 つの独立した複製において測定した（図 1 0 C）。全体として、4 3 個の代謝産物が、S 2 2 W 構築物と野生型構築物との間で有意に異なり、解糖 / 糖新生及びペントースリン酸経路における多くの中間体が、ヘキソースリン酸、ペントースリン酸、セドヘプツロース 7 - リン酸、グリセルアルデヒド 3 - リン酸及びジヒドロキシアセトンリン酸を含む、突然変異体 F O X O 1 - G F P を有する細胞内で上昇した。

40

【 0 2 0 5 】

形質導入された H e p G 2 細胞株の R N A 配列決定は、細胞周期（q < 0 . 0 0 0 1）、脂質異化作用（q < 0 . 0 0 0 1）及び F O X O 媒介転写標的（q = 0 . 0 0 8）に関

50

与する遺伝子セットの有意な上方制御；並びにアポトーシス促進プロセス ($q = 0.0004$) 及び標準的な解糖 ($q < 0.0001$) に関連する遺伝子セットの下方制御を明らかにした。(図10H~図10L)。これは、慢性肝疾患において見られる再発性FOXO1突然変異が機能的であり、肝細胞代謝の変化を誘導し、肝細胞生存の変化に寄与することを示唆する。

【0206】

実施例3：ドライバー突然変異を有するクローン及び患者の特性

肝臓の8つのクイノーセグメントの全てがサンプリングされた2人の患者について、特定されたドライバー突然変異及び他のゲノム特徴が複数の領域にわたって複製されたことを見出され、このことは、単一の試料からの知見が肝臓全体を広く代表することを示唆している。したがって、分析した肝臓試料から外挿することにより、有意な遺伝子のそれぞれにドライバー突然変異を有する全肝腫瘍を推定した(図12A)。これは、第一に、ドライバー突然変異を有するクローンが一部の患者において数百グラムの肝臓質量を占め得ること、及び第二に、ドライバーの分布が、どの遺伝子が影響を受け、どのレベルの関与が観察されたかについて、かなりの患者間変動を示すことを明らかにした。興味深いことに、FOXO1 ($p = 0.005$ 、ウィルコクソン検定)、CIDEB ($p = 0.001$) 及びACVR2A ($p = 0.001$) における突然変異を保有するクローンの推定サイズは、野生型クローンよりも平均して大きく(図12B)、このことは、これらの突然変異がクローンに選択的利点を付与し、それらの優先的拡大を可能にすることを示している。

10

20

【0207】

FOXO1におけるS22Wホットスポット突然変異、並びにCIDEB及びGPAMにおける突然変異は、健康な対照肝臓試料では観察されなかったが、ARLD及びNAFLDの両方からの肝臓組織で見られ、これらの選択的利点が疾患の発症又は進行の間に現れることを暗示している。中等度のコホートサイズにもかかわらず、FOXO1、CIDEB及びGPAMにおける突然変異は、広範な患者の特徴：両方の性別；広い年齢範囲；2型糖尿病の有無；及び様々な重症度の組織学的異常、にわたって見られた(図12C~図12E)。これは、コホートからの結果がARLD及びNAFLD患者にわたって一般化されることを示唆する。

【0208】

実施例4：HCCの選択的ランドスケープとの比較

選択的ランドスケープをHCCと比較するために、ICGCデータポータルに記録されたフランス、日本、中国及び米国から1670個のHCCについての突然変異コールにアクセスした。以前に報告されているように、ACVR2Aは、HCCにおける既知の腫瘍抑制遺伝子であり(例えば、The Cancer Genome Atlas Research Network. Comprehensive and Integrative Genomic Characterization of Hepatocellular Carcinoma. Cell 169, 1327-1341 (2017) を参照されたい)、1670個のHCCにおいて報告された77個の非同義突然変異を有する(本発明者らの1590個の慢性肝疾患ゲノムにおける16個の点変異及び4個の構造バリエーションと比較して；補足表S8~S9)。TNRC6Bはまた、HCCにおいて有意に突然変異しており($q = 0.0001$)、タンパク質短縮突然変異が優勢である(24個のバリエーション)。このことは、それがHCC並びにARLD及びNAFLDからの再生肝臓における腫瘍抑制遺伝子であることを示唆している。

30

40

【0209】

コホートにおいて非常に頻繁に変異した3つの代謝遺伝子FOXO1、CIDEB及びGPAMは、HCCにおいて有意に変異しなかった(それぞれ $q = 1.0$ 、 1.0 及び 0.6)。FOXO1 S22W突然変異は、1670個のHCCのうち3個のみで見出され(0.18% 、 $Cl_{95\%} = 0.05 \sim 0.6\%$)、コホートにおいてこの突然変異を保有する24個のクローンよりも有意に低かった($p = 2 \times 10^{-5}$ 、Fisherの正

50

確率検定)。GPAMについて、同義突然変異に対する非同義突然変異の分布は、単一遺伝子として試験した際の予測とは有意に異なっていたが ($p = 0.006$)、多重仮説検定の補正ではもはや有意ではなかった ($q = 0.6$)。

【0210】

実施例5：テロメア長及び遺伝率

Schulze, K. et al. Nat. Genet. 47, 505 - 511 (2015)、Fujimoto, A. et al. Nat. Genet. 44, 760 - 4 (2012)、Fujimoto, A. et al., Nat. Genet. 48, 500 - 509 (2016)、及びLi, Y. et al. Nature 578, 112 - 121 (2020)に報告されているように、大多数のHCCは、テロメラーゼ遺伝子であるTERTを活性化させる突然変異を有する。しかしながら、ここで研究した再生結節では、TERTプロモーター中の標準的ホットスポットに影響を及ぼす1つの突然変異のみが観察された。試料中のテロメア動態を評価するために、各マイクロダイセクションについてのテロメア長 (Farmer, J. H. R. et al. Sci. Rep. 2018, 8, 1 - 17) を推定した。

10

【0211】

テロメアリード (TTAGGG及びCCCTAA六量体を含む) について対応するWGSデータを分析することによって、研究した各マイクロダイセクションのテロメア長 (塩基対の単位で) を推定した。これを達成するために、Telomerecat v3.4.0ソフトウェアを使用し (Farmer et al. (2019))、長さ補正を可能にし、シミュレーションの数を100に設定して長さ推定値の不確実性を抑制した。各SNVクラスターに、最も高いVAF中央値を有するメンバーマイクロダイセクションに対応するテロメア長を割り当てた。

20

【0212】

テロメア長は、ベイズ混合効果モデルを使用してモデル化され、これらは、本発明者らがテロメア長に対する年齢、クローンサイズ及び疾患の効果を評価することを可能にし、同時に、クローン間の系統発生的関係及び患者内の非独立性から生じる相関を制御及び定量した。使用した特定のアルゴリズムは、RパッケージMCMCglmm (Hadfield, J. Stat. Softw. 2010, 33, 1 - 22) であった。

【0213】

これらのモデルを効果的に使用するために、系統樹上のクローン間の関連性マトリックスを構築した。これは、木をウルトラメトリックにすることを必要とした (すなわち、全ての先端が同じ点で終了する) ことを必要とした。木をウルトラメトリックにするための既存の方法があるが、それらは、正常組織における体細胞突然変異についての現在の目的に特に適していなかった。本発明者らは、初期突然変異の真の数を推定するためのより多くの検出力を確立した (初期突然変異は、2回以上のマイクロダイセクションにおいて、平均して、より高いパリアント対立遺伝子画分において見られるため)。本発明者らによって見出された後期のシングルトン枝の枝長の推定値は、それほど確かではない。したがって、本発明者らは、各分岐点 (合体) を分子時間の所与の割合に置くための再帰関数を開発した。木の根から開始し、先端に向かって進行して、各合体の位置は、根とその合体との間で獲得された変異の数を、その数に合体から下降する枝における突然変異の数の平均を加えたもので割ったものとして推定される。

30

40

【0214】

ベイズ線形混合効果モデルを使用して、テロメア長をモデル化した。従属変数は、各クローンの平均テロメア長 (「長さ」) であり、塩基対で測定され、ガウス分布に従うと予想された。固定効果は以下の通りであった：病因 (ARLD及びNAFLDについてのダミー変数)；患者の年齢 (歳)；そのクローンにおける突然変異の数；及びクローン面積。本発明者らは、各患者についてのランダム効果と、ブロック対角行列の形態で符号化された系統関係についてのランダム効果とを適合させた。先行物は、情報価値のない逆ウィッシュアート分布であった。MCMCチェーンは、バーンインとしてこれらのうちの1, 0

50

00, 000個を用いて11, 000, 000回の反復について実行し、1000回の反復ごとに間引きした。

【0215】

MCMCチェーンに関する通常の診断をチェックし、これらは、(1)バーンインが完了した後に変数に系統的な傾向がなかったことを示した(バーンインは十分な長さであった); (2)自己相関は、チェーンの1000番目のインスタンスごとに間引きされた隣接推定値の間で比較的低かった; (3)収束の診断試験は全て合格であったことを示した。

【0216】

テロメア長の「体細胞遺伝率」の推定値は、系統発生関係を記述する変数に起因するランダム効果における分散成分対残留分散及び患者間効果に起因する分散を計算することによって導出された。テロメア長についての「体細胞遺伝率」の推定値は、21%(95%事後間隔 = 0 ~ 44%)であった。すなわち、年齢、疾患及びクローンサイズの影響が説明されると、コホートにわたるテロメア長の残留変動の約21%は、クローン間の系統発生関係によって説明することができる。しかしながら、この推定値にはいくらかの不安定性があり、おそらくMCMC鎖の反復の約26%は体細胞遺伝率を示唆しない。いずれにしても、この試料間相関の交絡効果は、モデリングフレームワーク内で補正されている。

【0217】

結果:

全ゲノム配列決定データからの各マイクロダイセクションについてのテロメア長を推定した。コホート全体にわたってテロメア長にかなりの個体間及び個体内変動が観察され、正常肝臓と比較してNAFLD及びARLDではテロメアが明らかに短かった(図13A、図13B及び図B15)。これらのテロメア長を系統樹上に階層化すると、平均して、より密接に関連するクローンが、関連しないクローンよりも類似したテロメア長を有することが明らかになった(図14A及び図14B)。

【0218】

ベイズ混合効果モデルを使用してテロメア長をモデル化した。これは、テロメア長に対する年齢、クローンサイズ及び疾患の効果の評価を可能にし、同時に、クローン間の系統発生的関係から生じる相関を制御した。正常な肝臓と比較して、テロメア長は、ARLDを有する患者においてかなり短かった(平均posterior = -440 bp; 95%事後間隔(P195%) = -750 : 130 bp; pposterior = 0.009)及びNAFLD(平均posterior = -550 bp; P195% = -820 : 270 bp; pposterior = 0.0006; 図12C)。疾患についてのこれらの効果の大きさは、年齢に伴うテロメア長の比較的わずかな短縮をかなり上回った(平均posterior = -6 bp / 10歳; P195% = -15 : +2 bp / 10歳; pposterior = 0.1)。更に、テロメアは、クローンのサイズが増加するにつれて次第に短くなった(平均posterior = -100 bp / log₁₀(µm²); P195% = -160 : -50 bp / log₁₀(µm²); pposterior < 0.0001)。

【0219】

これらのデータは、ARLD及びNAFLDがテロメアの実質的な減少に関連しており、年齢に伴うテロメア長の比較的小さな短縮を上回っていることを示唆している。更に、テロメアは、クローンのサイズが増加するにつれて次第に短くなるが、これはおそらく、疾患進行中の肝細胞再生に関連する余分な細胞分裂を反映している。

10

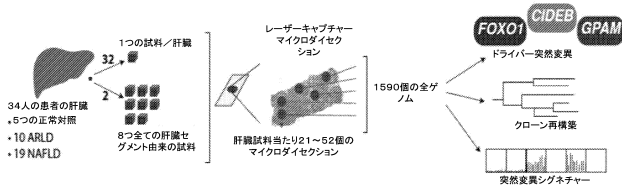
20

30

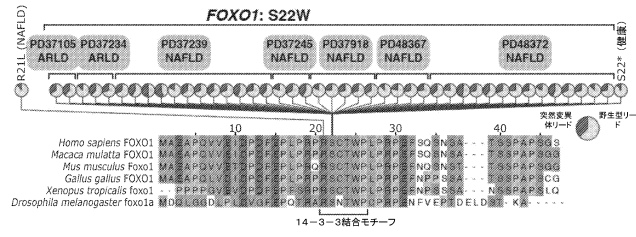
40

【 図 面 】

【 図 1 】

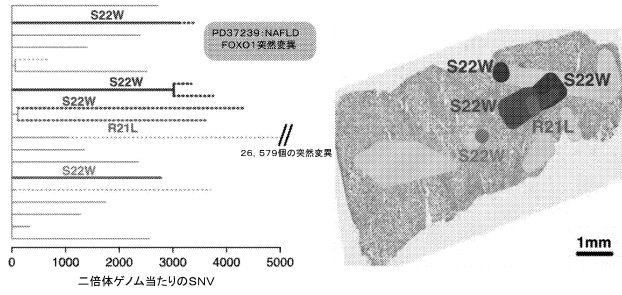


【 図 2 A 】

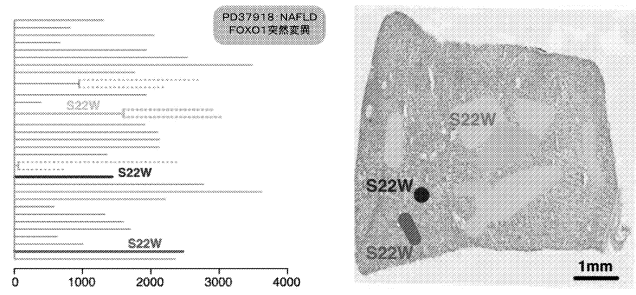


10

【 図 2 B 】



【 図 2 C 】



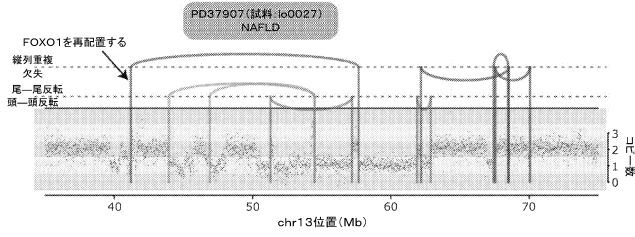
20

30

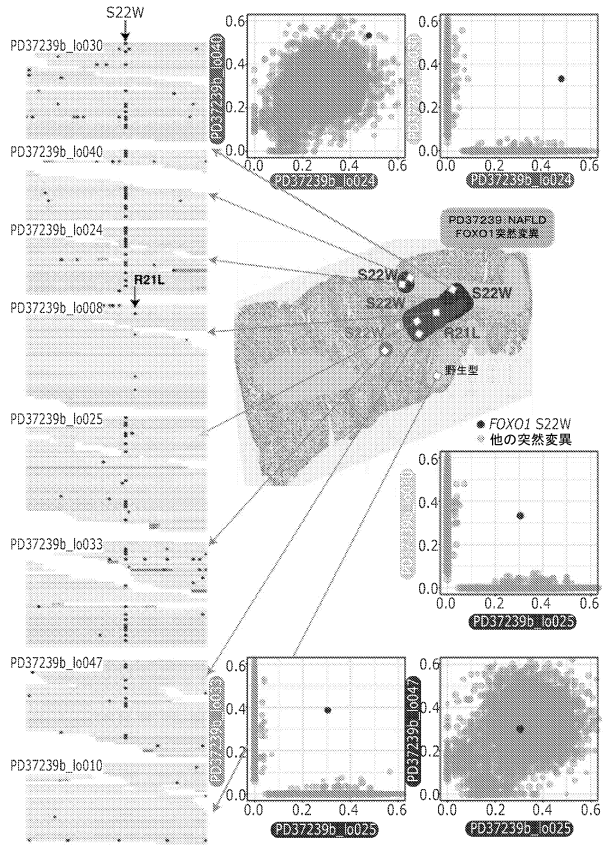
40

50

【 2 D 】



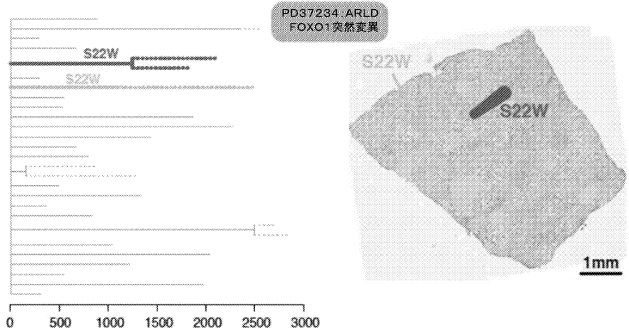
【 2 E 】



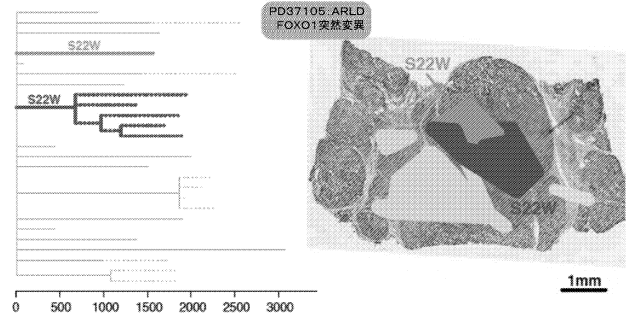
10

20

【 3 A 】



【 3 B 】

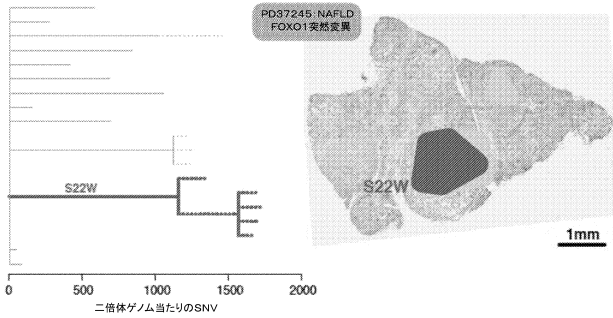


30

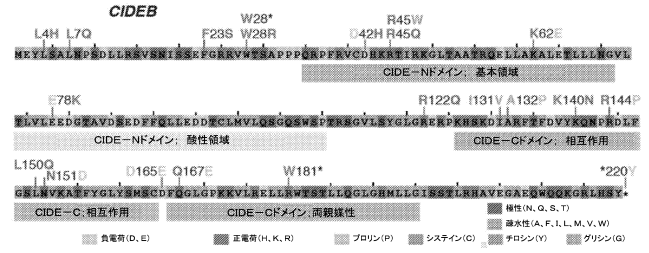
40

50

【図 3 C】

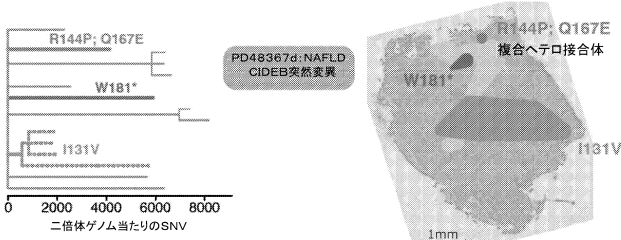


【図 4 A】

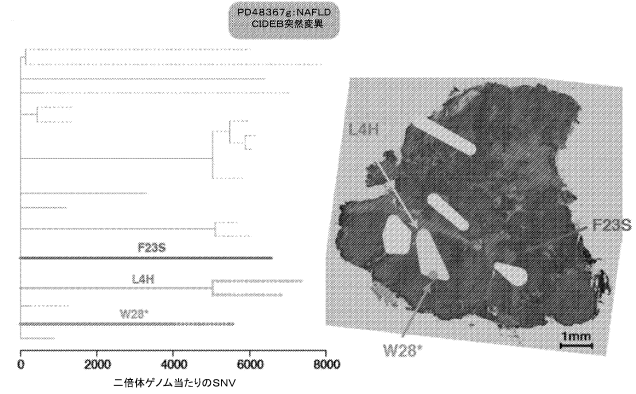


10

【図 4 B】

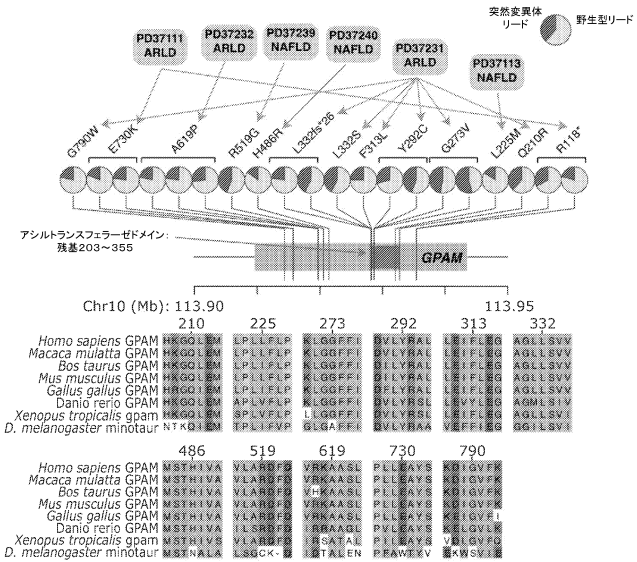


【図 4 C】

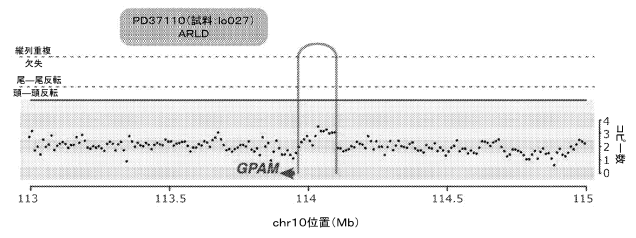


20

【図 5 A】



【図 5 B】

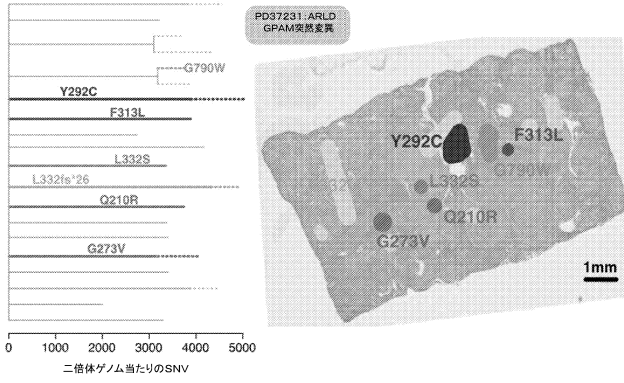


30

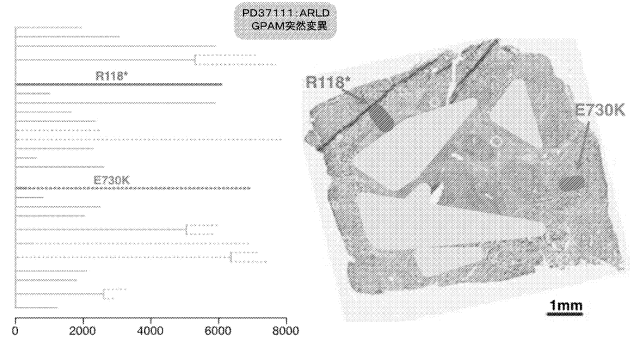
40

50

【 図 5 C 】

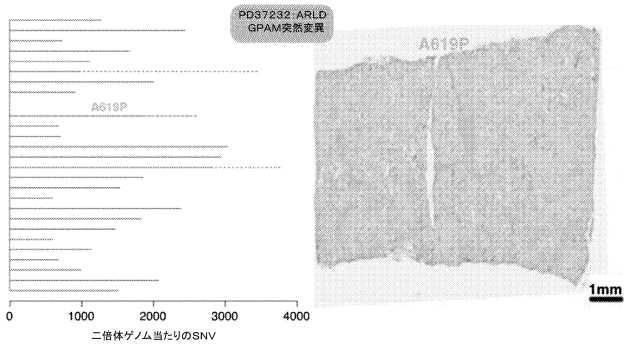


【 図 5 D 】

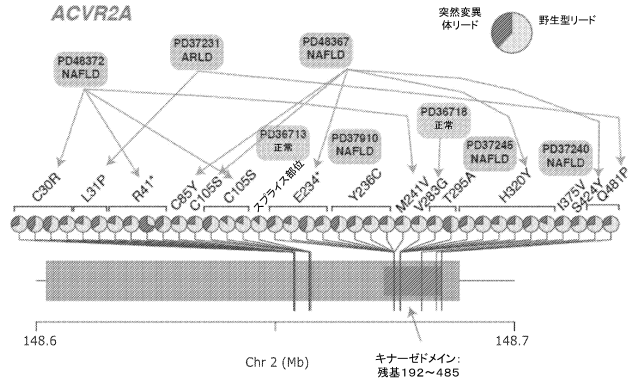


10

【 図 5 E 】

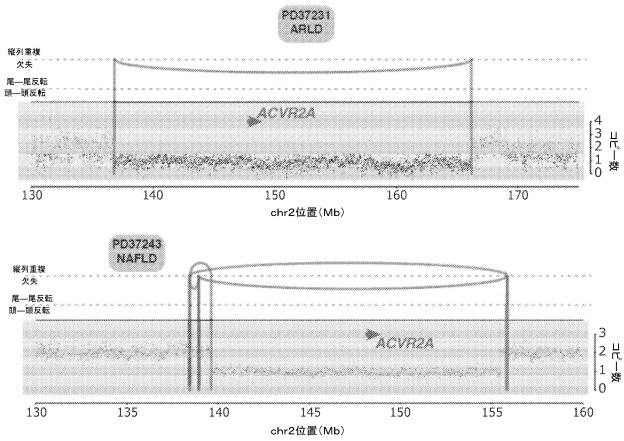


【 図 6 A 】

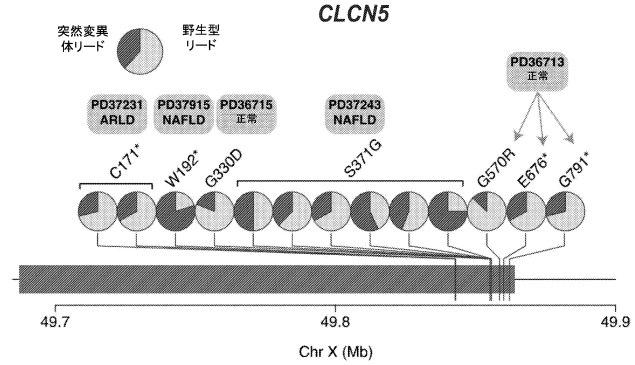


20

【 図 6 B 】



【 図 7 】

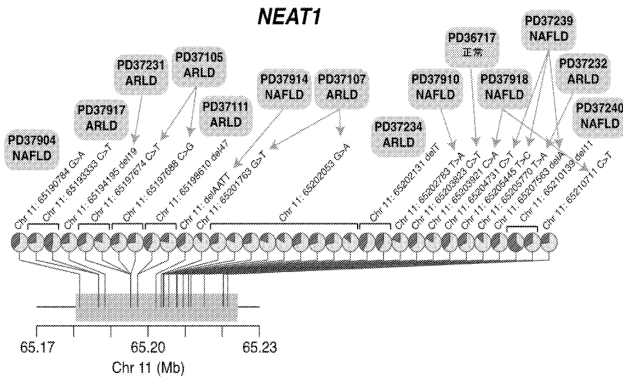


30

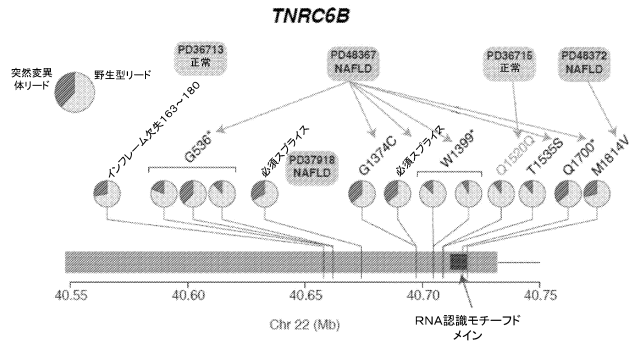
40

50

【 図 8 】

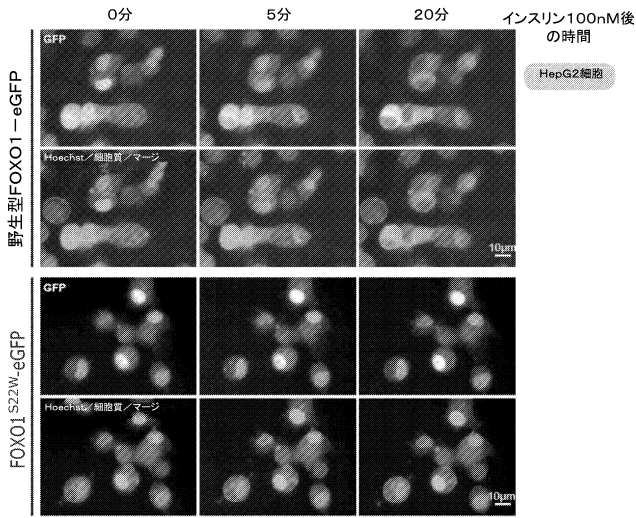


【 図 9 】

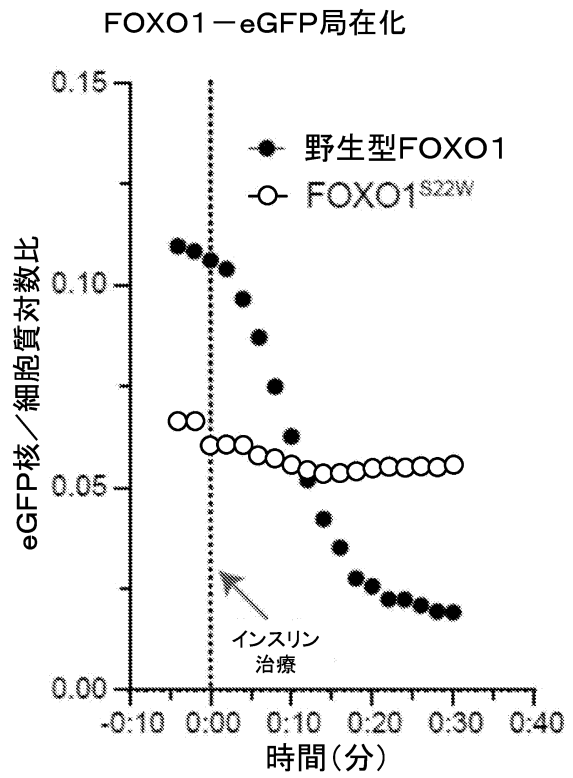


10

【 図 10 A 】



【 図 10 B 】



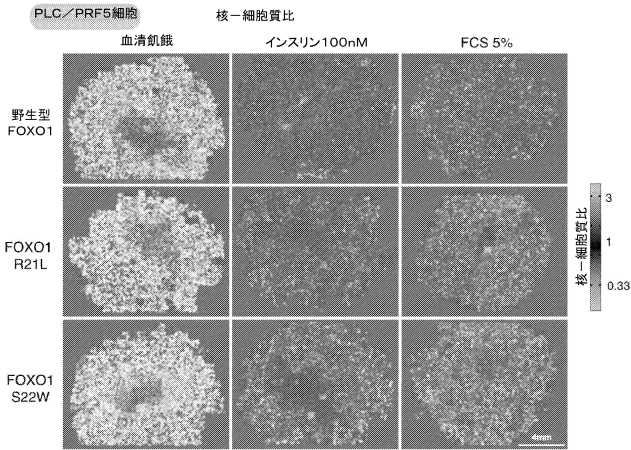
20

30

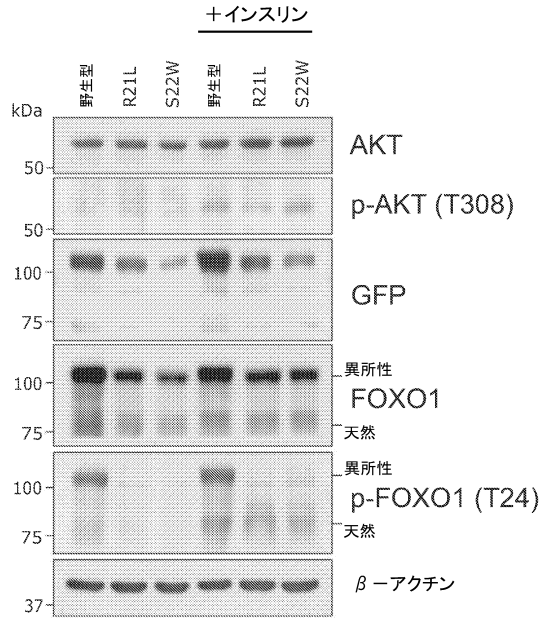
40

50

【図 1 0 G】



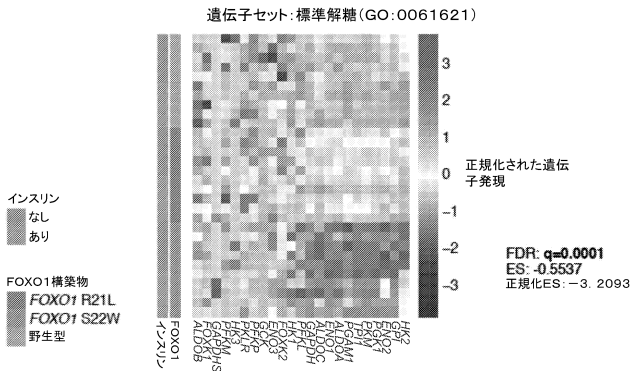
【図 1 0 H】



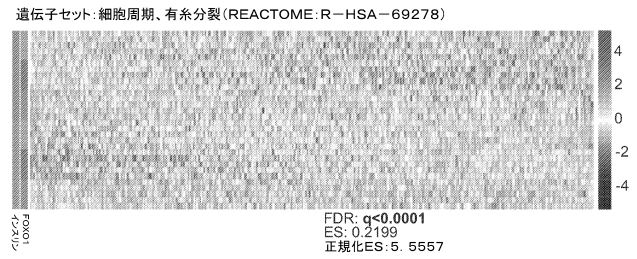
10

20

【図 1 1 A】



【図 1 1 B】

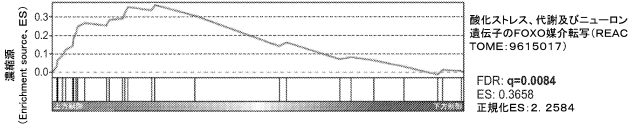


30

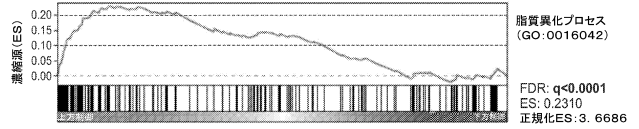
40

50

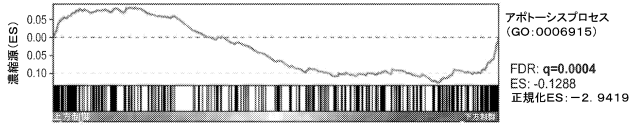
【図 1 1 C】



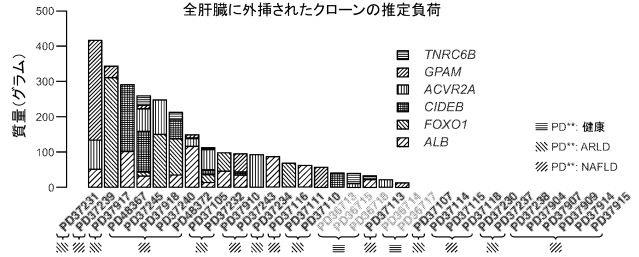
【図 1 1 D】



【図 1 1 E】

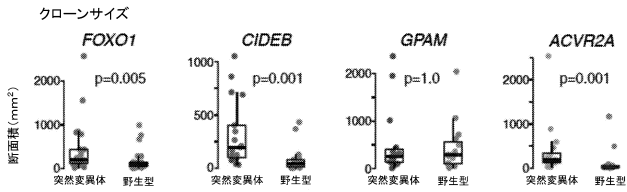


【図 1 2 A】

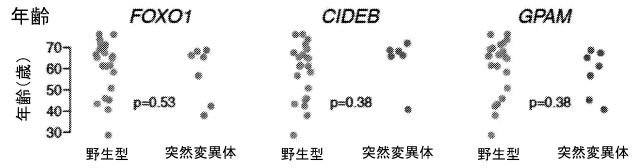


10

【図 1 2 B】

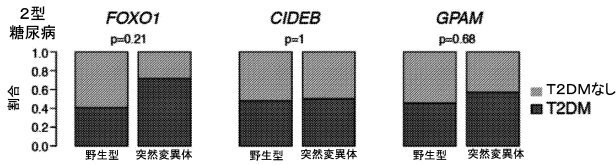


【図 1 2 C】

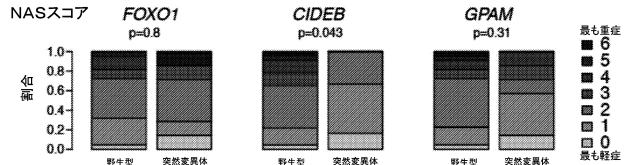


20

【図 1 2 D】



【図 1 2 E】

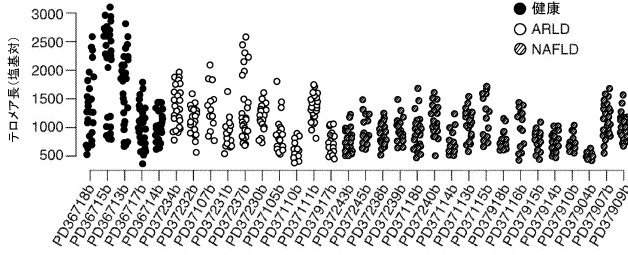


30

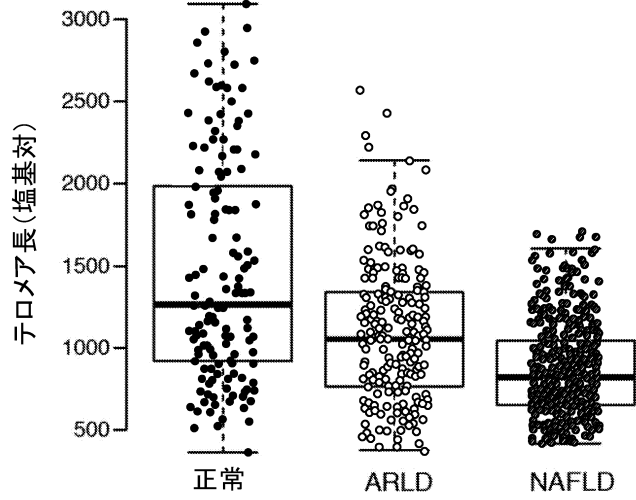
40

50

【図 1 3 A】

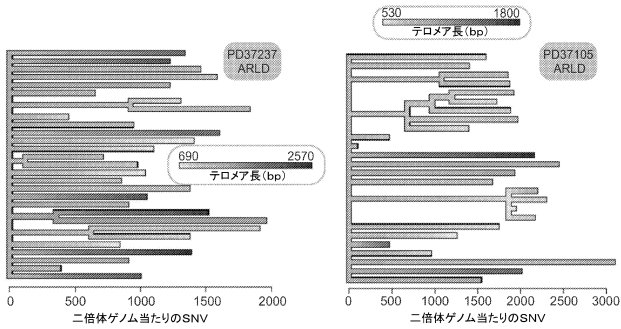


【図 1 3 B】

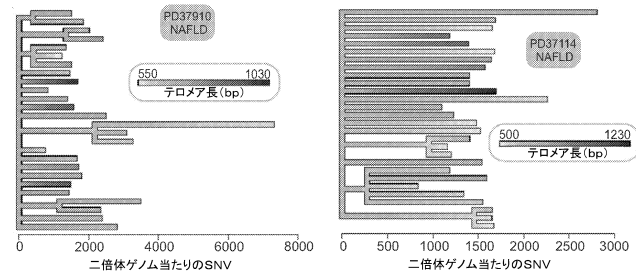


10

【図 1 4 A】



【図 1 4 B】



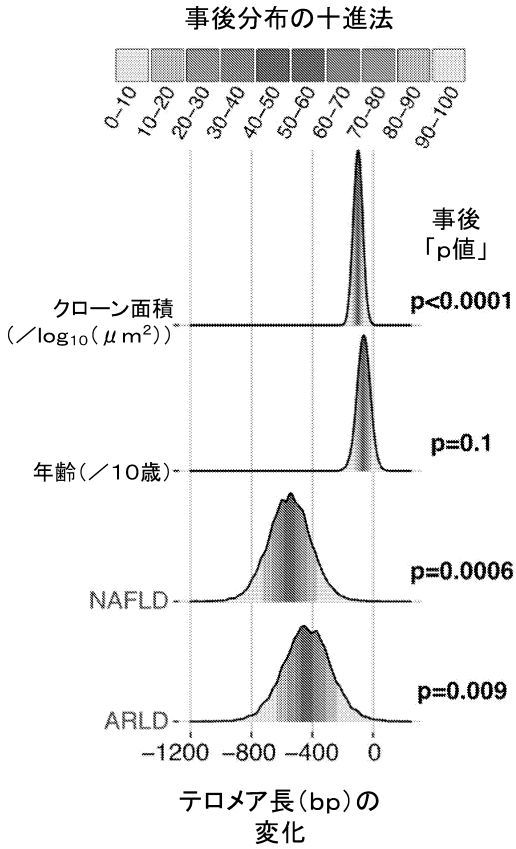
20

30

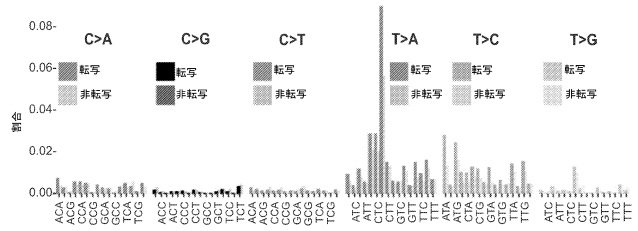
40

50

【 図 1 5 】



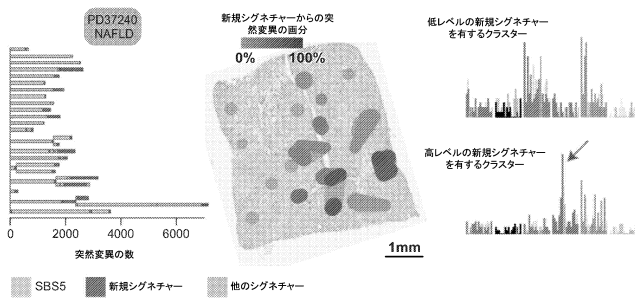
【 図 1 6 A 】



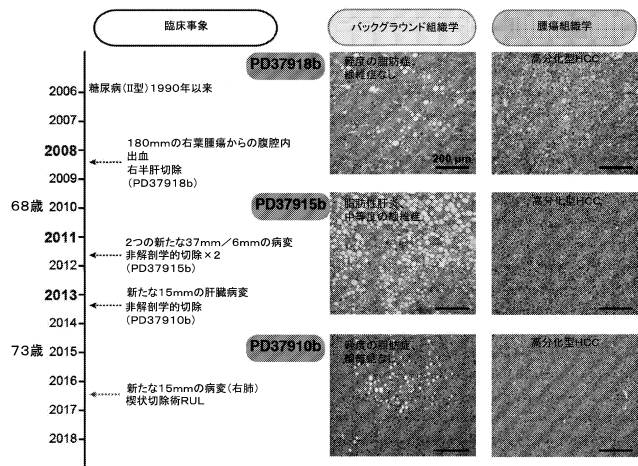
10

20

【 図 1 6 B 】



【 図 1 6 C 】



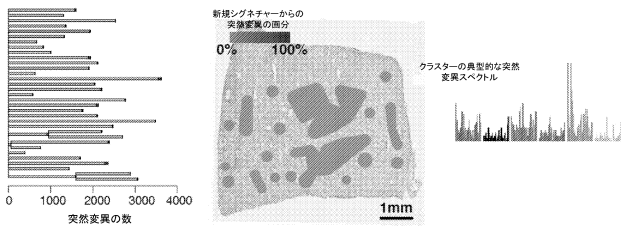
30

40

50

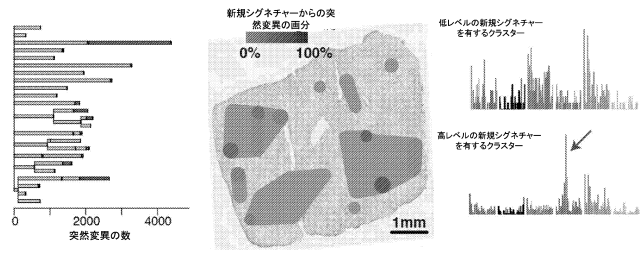
【図 16 D】

PDS37918b: 最初の生検 - 2008年、年齢66歳



【図 16 E】

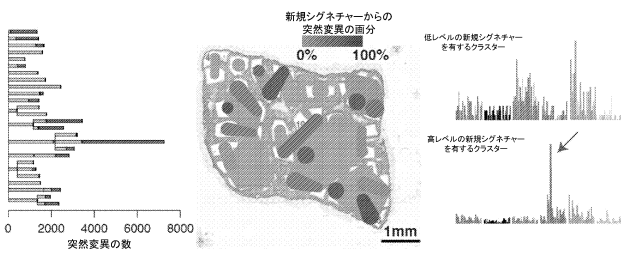
PDS37915b: 2回目の生検 - 2011年、年齢69歳



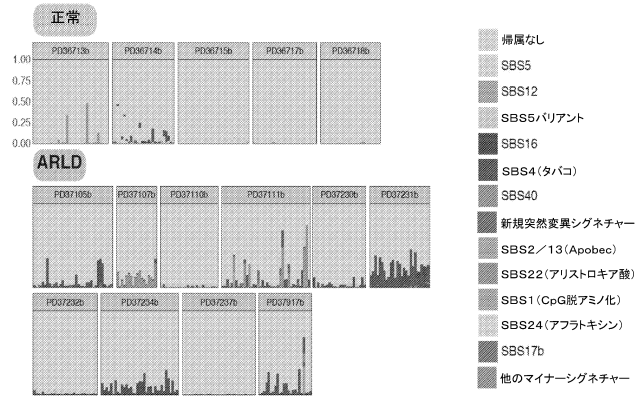
10

【図 16 F】

PD37910b: 3回目の生検 - 2013年、年齢71歳



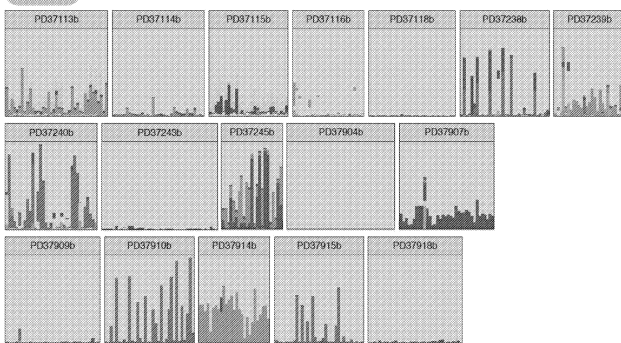
【図 16 G】



20

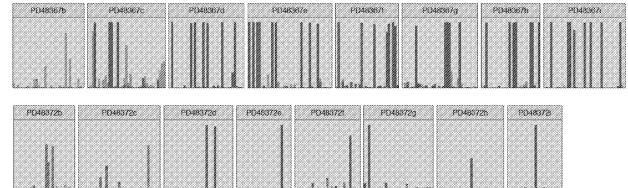
【図 16 H】

NAFLD



【図 16 I】

8つ全ての解剖学的セグメントがサンプリングされた2人のNAFLD患者

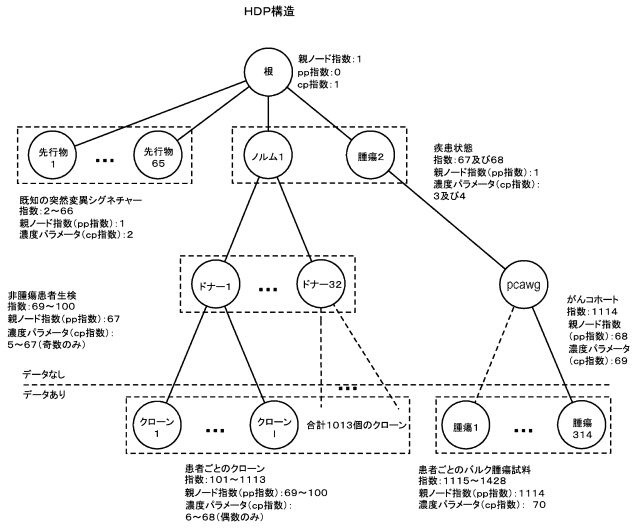


30

40

50

【 図 17 】



10

20

30

40

50

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT

International application No PCT/GB2021/052469
--

A. CLASSIFICATION OF SUBJECT MATTER INV. C12Q1/6883 ADD.		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) C12Q		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) EPO-Internal, BIOSIS, Sequence Search, EMBASE, WPI Data		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	BRUNNER SIMON F ET AL: "Somatic mutations and clonal dynamics in healthy and cirrhotic human liver", NATURE, NATURE PUBLISHING GROUP UK, LONDON, vol. 574, no. 7779, 1 October 2019 (2019-10-01), pages 538-542, XP036913421, ISSN: 0028-0836, DOI: 10.1038/S41586-019-1670-9 [retrieved on 2019-10-23] the whole document	1, 2, 14-16, 19, 21-23, 26
A	US 2017/175195 A1 (LO YUK MING DENNIS [CN] ET AL) 22 June 2017 (2017-06-22) the whole document	1-23, 26
	----- -/--	
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 16 December 2021		Date of mailing of the international search report 04/01/2022
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040. Fax: (+31-70) 340-3016		Authorized officer Betz, Jürgen

2

Form PCT/ISA/210 (second sheet) (April 2005)

10

20

30

40

50

INTERNATIONAL SEARCH REPORT

International application No PCT/GB2021/052469

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>VALENTI L ET AL: "726 GENOTYPE OF THE TRANSCRIPTION FACTOR FOXO1 MEDIATING INSULIN SIGNALING INFLUENCES THE SUSCEPTIBILITY TO AND SEVERITY OF NONALCOHOLIC FATTY LIVER DISEASE BY MODULATING FOXO1 MRNA LEVELS", JOURNAL OF HEPATOLOGY, ELSEVIER, AMSTERDAM, NL, vol. 50, 1 April 2009 (2009-04-01), page S265, XP026496339, ISSN: 0168-8278, DOI: 10.1016/S0168-8278(09)60728-5 [retrieved on 2009-04-01] the whole document</p> <p>-----</p>	1-4, 14-16, 19-23,26
A	<p>REHMAN MUHAMMAD YASIR ABDUR ET AL: "Transcriptome responses in blood reveal distinct biological pathways associated with arsenic exposure through drinking water in rural settings of Punjab, Pakistan", ENVIRONMENT INTERNATIONAL, PERGAMON PRESS, US, vol. 135, 18 December 2019 (2019-12-18), XP085976566, ISSN: 0160-4120, DOI: 10.1016/J.ENVINT.2019.105403 [retrieved on 2019-12-18] the whole document</p> <p>-----</p>	1-23,26
A	<p>PUENTES YINDRA ET AL: "Novel Approach Reveals That Genetic Variation in/Near FXR/RXR Pathway Genes Predisposes to Nonalcoholic Fatty Liver Disease in Humans", GASTROENTEROLOGY, vol. 146, no. 5, Suppl. 1, May 2014 (2014-05), page S923, XP009527781, & 55TH ANNUAL MEETING OF THE SOCIETY-FOR-SURGERY-OF-THE-ALIMENTARY-TRAC T (SSAT) / DIGESTIVE DISEASE WE; CHICAGO, IL, USA; MAY 03 -06, 2014 DOI: 10.1016/S0016-5085(14)63353-9 the whole document</p> <p>-----</p>	1-23,26
A	<p>USTIUGOVA ALINA S. ET AL: "Functional SNPs in the Human Autoimmunity-Associated Locus 17q12-21", GENES, vol. 10, no. 2, 23 January 2019 (2019-01-23), page 77, XP055808934, DOI: 10.3390/genes10020077 the whole document</p> <p>-----</p>	1-23,26

2

-/--

Form PCT15A/210 (continuation of second sheet) (April 2005)

10

20

30

40

50

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2021/052469

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	HITOMI YUKI ET AL: "Identification of the functional variant driving ORMDL3 and GSDMB expression in human chromosome 17q12-21 in primary biliary cholangitis", SCIENTIFIC REPORTS vol. 7, no. 1 1 December 2017 (2017-12-01), XP055808939, DOI: 10.1038/s41598-017-03067-3 Retrieved from the Internet: URL:https://www.nature.com/articles/s41598-017-03067-3.pdf [retrieved on 2021-05-31] the whole document -----	1-23,26
A	CHENG ZHIYONG ET AL: "Foxo1 integrates insulin signaling with mitochondrial function in the liver", NATURE MEDICINE vol. 15, no. 11 1 November 2009 (2009-11-01), pages 1307-1311, XP055808944, New York ISSN: 1078-8956, DOI: 10.1038/nm.2049 Retrieved from the Internet: URL:https://www.nature.com/articles/nm.2049.pdf [retrieved on 2021-05-31] the whole document -----	1-23,26
A	ADACHI ET AL: "The Forkhead Transcription Factor FoxO1 Regulates Proliferation and Transdifferentiation of Hepatic Stellate Cells", GASTROENTEROLOGY, ELSEVIER INC, US, vol. 132, no. 4, 26 April 2007 (2007-04-26), pages 1434-1446, XP022046106, ISSN: 0016-5085, DOI: 10.1053/J.GASTRO.2007.01.033 the whole document -----	1-23,26
A	MICHELINO DI ROSA ET AL: "Genetic variants in candidate genes influencing NAFLD progression", JOURNAL OF MOLECULAR MEDICINE, SPRINGER, BERLIN, DE, vol. 90, no. 2, 6 September 2011 (2011-09-06), pages 105-118, XP035010066, ISSN: 1432-1440, DOI: 10.1007/S00109-011-0803-X the whole document -----	1-23,26

10

20

30

40

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2021/052469

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2019/234779 A1 (OSPEDALE PEDIATRICO BAMBINO GESU [IT]) 12 December 2019 (2019-12-12) the whole document -----	1-23, 26

10

20

30

40

INTERNATIONAL SEARCH REPORT

International application No.
PCT/GB2021/052469

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

- 1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
 - a. forming part of the international application as filed:
 - in the form of an Annex C/ST.25 text file.
 - on paper or in the form of an image file.
 - b. furnished together with the international application under PCT Rule 13ter.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
 - c. furnished subsequent to the international filing date for the purposes of international search only:
 - in the form of an Annex C/ST.25 text file (Rule 13ter.1(a)).
 - on paper or in the form of an image file (Rule 13ter.1(b) and Administrative Instructions, Section 713).
- 2. In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.
- 3. Additional comments:

10

20

30

40

50

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/GB2021/052469

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2017175195 A1	22-06-2017	AU 2010294193 A1	05-04-2012
		CN 102741426 A	17-10-2012
		EP 2475781 A1	18-07-2012
		EP 3318644 A1	09-05-2018
		HK 1254807 A1	26-07-2019
		JP 5917397 B2	11-05-2016
		JP 6203803 B2	27-09-2017
		JP 6523375 B2	29-05-2019
		JP 2013504310 A	07-02-2013
		JP 2016063813 A	28-04-2016
		JP 2017212993 A	07-12-2017
		JP 2019062899 A	25-04-2019
		TW 201122481 A	01-07-2011
		US 2011098192 A1	28-04-2011
		US 2015284801 A1	08-10-2015
		US 2017175195 A1	22-06-2017
		WO 2011029899 A1	17-03-2011

WO 2019234779 A1	12-12-2019	NONE	

10

20

30

40

50

フロントページの続き

(51)国際特許分類	F I	テーマコード (参考)
A 6 1 K 31/7088(2006.01)	A 6 1 K 39/395	N
A 6 1 K 31/713(2006.01)	A 6 1 K 39/395	P
A 6 1 K 31/7105(2006.01)	A 6 1 K 31/7088	
A 6 1 K 48/00 (2006.01)	A 6 1 K 31/713	
G 0 1 N 33/50 (2006.01)	A 6 1 K 31/7105	
G 0 1 N 33/68 (2006.01)	A 6 1 K 48/00	
C 1 2 N 15/12 (2006.01)	G 0 1 N 33/50	P
A 6 1 P 3/06 (2006.01)	G 0 1 N 33/68	
	C 1 2 N 15/12	
	A 6 1 P 3/06	

MK,MT,NL,NO,PL,PT,RO,RS,SE,SI,SK,SM,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,KM,ML,MR,N
E,SN,TD,TG),AE,AG,AL,AM,AO,AT,AU,AZ,BA,BB,BG,BH,BN,BR,BW,BY,BZ,CA,CH,CL,CN,CO,CR,CU,
CZ,DE,DJ,DK,DM,DO,DZ,EC,EE,EG,ES,FI,GB,GD,GE,GH,GM,GT,HN,HR,HU,ID,IL,IN,IR,IS,IT,JO,JP,K
E,KG,KH,KN,KP,KR,KW,KZ,LA,LC,LK,LR,LS,LU,LY,MA,MD,ME,MG,MK,MN,MW,MX,MY,MZ,NA,N
G,NI,NO,NZ,OM,PA,PE,PG,PH,PL,PT,QA,RO,RS,RU,RW,SA,SC,SD,SE,SG,SK,SL,ST,SV,SY,TH,TJ,TM,
TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN,WS,ZA,ZM,ZW

(特許庁注：以下のものは登録商標)

1. M A T L A B

ケンブリッジシャー ケンブリッジ ロビンソン ウェイ シーノオー キャンサー リサーチ ユーケ
ー ケンブリッジ インスティテュート

(72)発明者 キャンベル ピーター
イギリス国 シービー 1 0 1 エスエー ケンブリッジシャー ヒンクストン シーノオー キャンサー
ゲノム プロジェクト ウェルカム トラスト サンガー インスティテュート

(72)発明者 エヌジー スタンレー
イギリス国 シービー 1 0 1 エスエー ケンブリッジシャー ヒンクストン シーノオー キャンサー
ゲノム プロジェクト ウェルカム トラスト サンガー インスティテュート

F ターム (参考) 2G045 AA25 CB01 DA13 DA14 DA36
4B063 QA01 QA13 QA17 QA19 QQ08 QQ42 QQ52 QQ79 QS10 QS14
QS39
4C084 AA02 AA13 AA17 BA03 NA05 ZA751 ZC331 ZC412
4C085 AA13 AA14 BB11 BB22 EE01
4C086 AA01 EA16 MA01 MA04 NA05 ZA75 ZC33