

明 細 書

発明の名称：

ゲノムDNAに大規模な欠失を生じさせる方法およびゲノムDNAの分析方法

技術分野

[0001] 本発明は、ゲノムDNAに大規模な欠失を生じさせる方法およびゲノムDNAの分析方法に関する。

背景技術

[0002] 高等生物においては、ゲノム配列の1－2%しかタンパク質をコードするために使用していないのに対し、哺乳類では近位（プロモーター）または長距離（エンハンサー、サプレッサー、インシュレーターなど）の調節エレメントのために大規模かつ遺伝子間配列が割り当てられている。したがって、機能性人工細胞創出のためのゲノムの設計に際しては、遺伝子の選択に加えて、遺伝子間領域の設計が鍵となる。Encyclopedia of DNA Elements（ENCODE）プロジェクトでは、ヒト培養細胞において、遺伝子間領域の大部分が特定の転写因子結合やヒストン修飾を示すことが示唆されており、合理的な遺伝子間領域設計の指針となる可能性が示唆されている。しかし、これらの転写およびエピジェネティックな活性がすべて正常な細胞機能に必要なものなのかどうかは不明である。哺乳類の細胞ネットワークは、遺伝子ファミリーの拡大と遺伝子制御ネットワークの冗長性の増大により、進化的に強固なものとなった。そのため、非必須遺伝子だけでなく、遺伝子間領域の大部分が細胞の生存や維持のためには必ずしも必要ではない可能性がある（非特許文献1～3）。

先行技術文献

非特許文献

[0003] 非特許文献1: Xavier, J. C., Patil, K. R. & Rocha, I., Microbiol Mol Biol Re

v, 78, 487-509 (2014).

非特許文献2: Posfai, G., Science, 312, 1044-1046 (2006).

非特許文献3: Hutchison, C. A. et al., Science, 351, aad6253-aad6253 (2016).

発明の概要

[0004] 本発明は、ゲノムDNAに欠失を生じさせる方法およびゲノムDNAの分析方法を提供する。

[0005] 本発明者らは、標的配列を配列特異的に切断することができる配列特異的核酸切断分子を用いて、2つの標的配列間においてDNA領域の欠失を誘導する方法を開発した。本発明者らはまた、ゲノムDNA中において欠失させるDNA領域を大きくするときに、細胞の生存および／または増殖に必要な遺伝子を欠失させると、細胞の生存および／または増殖が負の影響を受けることを見出した。これにより、本発明者らは、細胞の生存および／または増殖に必要な遺伝子の場所および当該遺伝子を特定する方法を見出した。本発明者らはまた、細胞の生存および／または増殖に必要な遺伝子を、ゲノムDNAに異所的に発現可能に挿入することにより、DNA領域の欠失を有しながらも生存および／または増殖が可能な細胞を作製する方法を見出した。また、本発明者らは、ネガティブ選択用マーカー遺伝子が挿入されたDNA領域を欠失させる方法を見出した。本発明は、これらの知見に基づくものである。

[0006] 本発明によれば、以下の発明が提供され得る。

(1) インビトロの方法であって、

(a) 単離された細胞を含む細胞集団を用意することと、

(b) 前記細胞集団中の前記細胞のゲノムDNAに、ゲノムDNA上の2箇所の標的配列を配列特異的に切断することができる配列特異的核酸切断分子をそれぞれ作用させて、ゲノムDNA上の2箇所の標的配列においてそれぞれ切断を生じさせ、これによって、細胞集団中の少なくとも一部の細胞にお

いて、前記 2 箇所の切断部位の間の領域において DNA 領域の欠失を生じさせることと、

(c) その後、得られた細胞集団を培養し、細胞の増殖または生存に対する前記 DNA 領域の欠失の影響を決定することと、

を含む、

方法。

(2) 前記 (c) において、前記決定が、DNA 領域の欠失効率またはその推定値を決定することと、その後の培養後の細胞集団における DNA 領域の欠失を有する細胞の割合と、決定された欠失効率またはその推定値とを比較することとにより行われる、上記 (1) に記載の方法。

(3) 前記 (c) において、欠失効率またはその推定値の決定、および、培養後の DNA 領域の欠失を有する細胞の割合が、それぞれ細胞集団を含む懸濁液中に含まれる全細胞に対する欠失を有する細胞の割合として決定される、上記 (2) に記載の方法。

(4) 上記割合が、懸濁液中に含まれる、欠失を有するゲノム DNA および有しないゲノム DNA の計数技術により決定される、上記 (3) に記載の方法。

(5) 上記 (1) ~ (4) のいずれかに記載の方法であって、

(d) 対照となる細胞集団と比較して、前記 (c) における細胞の生存および／または増殖に対する影響の有無または大きさに基づいて、対照のゲノム DNA と比較して欠失した DNA 領域が細胞の生存および／または増殖を制御する遺伝子を含むか否かを決定すること

をさらに含む、方法。

(6) 上記 (5) に記載の方法であって、対照となる細胞集団が、前記 (b) により、より大きな DNA 領域、より小さな DNA 領域、または異なる DNA 領域の欠失を有する細胞を含む細胞集団である、方法。

(7) 上記 (5) または (6) に記載の方法であって、対照のゲノム DNA と比較して欠失した DNA 領域に存在する遺伝子から、細胞の生存および／

または増殖を制御する遺伝子を少なくとも1つ特定することをさらに含む、方法。

(8) 上記 () ~ (7) のいずれかに記載の方法であって、

(f) 欠失したDNA領域に存在する遺伝子から、細胞の生存および／または増殖を制御する遺伝子を少なくとも1つ特定することをさらに含む、方法。

(9) (g) 制御配列に作動可能に連結した細胞の生存および／または増殖を制御する遺伝子の少なくとも1つを、当該DNA領域の欠失を有するゲノムDNAに異所的に導入することをさらに含む、上記(7)または(8)に記載の方法。

(10) 前記(b)において欠失するDNA領域のサイズが、0.5 Mbp以上である、上記(1)~(9)のいずれかに記載の方法。

(11) 前記(b)において欠失するDNA領域のサイズが、1 Mbp以上である、上記(1)~(10)に記載の方法。

(12) インビトロの方法であって、

(α) 単離された細胞を含む細胞集団を用意することと、ここで単離された細胞は、欠失させる領域中にネガティブ選択用マーカー遺伝子を含み、

(β) 前記細胞集団中の前記細胞のゲノムDNAに、ゲノムDNA上の2箇所の標的配列を配列特異的に切断することができる配列特異的核酸切断分子をそれぞれ作用させて、ゲノムDNA上の2箇所の標的配列においてそれぞれ切断を生じさせ、これによって、細胞集団中の少なくとも一部の細胞において、前記2箇所の切断部位の間の領域においてDNA領域の欠失を生じさせることと、ここで、前記2箇所は、前記ネガティブ選択用マーカー遺伝子を挟む位置に設計され、

(γ) ネガティブ選択マーカー遺伝子を有しない細胞を選択することと、を含む、方法。

(13) 前記ネガティブ選択マーカー遺伝子が、欠失させる領域中に外来的に挿入されたものである、上記(12)に記載の方法。

(14) 前記ネガティブ選択マーカー遺伝子が、ゲノムDNA中の内在的な遺伝子である、上記(12)に記載の方法。

(15) 細胞が真核細胞である、上記(1)～(14)のいずれかに記載の方法。

(16) 工程(c)が、工程(b)で得られた細胞集団から一部の細胞を選別する工程を経ずに実施される、上記(1)～(11)のいずれかに記載の方法。

図面の簡単な説明

[0007] [図1] 図1は、H P R T 1 遺伝子を含むDNA領域の欠損を伴うヒト細胞の作製方法を示す。パネルaは、ネガティブ選択用マーカー遺伝子を用いたDNA領域の欠失スキームを示す。ネガティブ選択用マーカー遺伝子を含むDNA領域が欠失すると、細胞は生存し易くなる。したがって、ネガティブ選択用マーカー遺伝子の発現を指標として細胞を選別することによって、DNA領域の欠失を有する細胞を取得することができる。パネルaでは、配列特異的核酸切断分子としてC R I S P R - C a s 9 システムを用いた例を図示している。細胞に対して、ネガティブ選択用マーカー遺伝子を挟むように配置された2箇所の標的配列に対するgRNAとC a s 9 エンドヌクレアーゼを発現させるプラスミドをトランスフェクションする。2日後、一部の細胞を取得してゲノムの欠失領域のジャンクションPCRをデジタルPCR法により行う。コロニーが形成されるまで細胞を選択培地で培養する。これにより、ネガティブ選択用マーカー遺伝子を有する細胞、すなわち、欠失が生じなかった細胞は死滅し、欠失を有する細胞のみが生存し、増殖して、コロニーを形成する。コロニー数を計数し、再びジャンクションPCRをデジタルPCR法によって行い、欠失の存在を確認する。パネルbは、X染色体上の内在性のH P R T 1 遺伝子座周辺を欠失させ、その後、デジタルPCR法によってジャンクション領域をPCRした結果である。デジタルPCRでは、100ゲノムを1PCRの鋳型とし、48PCRを行った。14例で増幅産物が得られたことから、欠失効率は、 $14 / 48 / 100 \times 100 = \text{約} 0.3$

(%)である。パネルcは、6-チオグアニン(6-TG)による選択後に形成されたコロニーおよびその数を示す。パネルdは、L1-R1による欠失領域の前後の配列のシーケンスを示す。概ね、2つの標的配列の間のDNA領域が欠失したことが分かる。

[図2]図2は、HPR T1遺伝子の周辺のゲノム領域のマップと、当該領域のメガスケールの欠失の作製を示す。パネルaは、X染色体のHPR T1遺伝子の周辺のRefSeq遺伝子を示す。OGEEデータベースにより推定される細胞増殖のための必須遺伝子が丸で示される。パネルbとcは、セントロメア方向およびテロメア方向への欠失領域を格段させた実験の結果を示す。配列特異的核酸切断分子の標的配列が記号L1~L5およびR1~R5を伴う垂直の破線で示されている。それぞれのDNA領域の欠失を誘発させた細胞集団の6-TG選択後のコロニー形成数と欠失効率(%)が示されている。

[図3]図3は、L4-R4による欠失を有するクローンのゲノム分析の結果を示す。パネルaは、HAP1細胞(WT)とL4-R4による欠失を有するクローンがそれぞれ有する遺伝子の存在を明らかにするPCRの結果である。パネルbは、L4-R4による欠失を有するクローンの相対的な増殖能を示す。相対的な増殖能は、野生型細胞に対する欠失を有する細胞の割合として求め、欠失後0日目および16日目で測定し、当該割合が変化しなかったことを示す。パネルcは、L4-R4による欠失を有するクローンにおける遺伝子発現のマイクロアレイによる分析結果を示す。

[図4]図4は、必須遺伝子を含むメガスケールの欠失と、欠失のレスキュー実験を示す。パネルaは、必須遺伝子を挟む上流及び下流3kbの配列を含むpHY262ベクターを用いたGAPDH遺伝子座の3'非翻訳領域(UTR)への遺伝子挿入のスキームを示す。パネルbは、レスキューされたRBMX2およびMMGT1遺伝子の発現をリアルタイムPCRにより分析した結果を示す。パネルcおよびdはそれぞれ、内在性のRBMX2およびMMGT1遺伝子を欠損した株と、RBMX2およびMMGT1がレスキューさ

れた株のコロニー形成数および欠失効率を示す。

[図5]図5は、メガスケールの欠失による必須遺伝子を含む領域の欠失の結果を示す。パネルaとcはそれぞれ、RBMX2およびMMGT1遺伝子を含む領域が欠失したが、GAPDH遺伝子座下流へに各遺伝子が再導入され、さらに欠失領域が拡大された実験の結果を示す。パネルbは、パネルa実験の後の相対的コロニーサイズの測定結果を示す。p値は、3回の独立した実験結果から求められた。

[図6]図6は、HCT116細胞におけるネガティブ選択用マーカー遺伝子を用いたメガスケールの欠失の作製を示す。パネルaは、チミジンキナーゼ（TK）発現カセット有するpHYT271ベクターを用いてHCT116細胞のOCLR遺伝子の下流にTKを挿入するスキームを示す。ジャンクションPCRによって増幅産物のサイズが大きくなったことから、挿入が確認されている。パネルbは、HCT116細胞中の欠失を生じさせる遺伝子座の遺伝子マップを示す。パネルcは、ネガティブ選択用マーカー遺伝子であるTKを用いたメガスケールのDNA領域の欠失のようすを示す。欠失はセントロメア方向に拡大された。各欠失に対するセクション後のコロニー形成数および欠失効率が示される。数値は、3回の独立した実験の平均および標準偏差を示す。

[図7]図7は、ネガティブ選択をせずにメガスケールの欠失を生じさせる方法を示す。HAP1細胞において、標的領域A～Cの3種類の標的領域を欠失させた。ゲノムの欠失を誘発させた2日後と17日後に、1ウェル100ゲノムに対するデジタルPCR法によって欠失の有無を判定し、細胞集団中の欠失を有する細胞の割合を求めた。その結果、標的領域BおよびCを欠失させた場合に、細胞集団中の欠失を有する細胞の割合が低下した。この低下は、欠失させた領域に細胞の生存および／または増殖に必要な遺伝子が含まれていることを示すものである。

[図8]図8は、ヒトX染色体上で、細胞の生存に必須ではないとされていた領域について本発明の方法で図に示す領域をそれぞれ欠失させて、細胞の生存

率に対する各欠失の影響を調べた実験の結果を示す。

発明の具体的な説明

- [0008] 本明細書では、「細胞」とは、少なくともゲノムDNAと細胞質と、これらを包む膜構造を有する生命の基本単位をいう。細胞としては、特に限定されないが、原核生物の細胞と真核生物の細胞とが挙げられる。ゲノムDNAは、その細胞の内因性のDNAを含むが、必ずしもその細胞の内因性の因子のみから構成されるものではない。
- [0009] 細胞は、細胞のゲノムDNAを含むが、外来の侵入者（例えば、病原体）のゲノムDNAをさらに含む場合がある。本明細書では、細胞自体のゲノムDNAを「宿主ゲノムDNA」という。侵入者のゲノムDNAは、宿主のゲノムDNAとは別体として細胞内で存在し得るが、宿主のゲノムDNAに組み込まれていることもある。宿主ゲノムDNAは、外来性の因子（例えば、ウイルス等のゲノムDNAの全部または一部の挿入）を含んでいることがある。
- [0010] 本明細書では、「細胞集団」とは、複数の細胞を含む組成物を意味する。
- [0011] 本明細書では、「単離」とは、目的の細胞を少なくとも1つの他の構成要素から分離することを意味する。単離は、例えば、自然な存在状態の細胞を自然な存在状態において一緒に存在する他の構成要素から分離して取り出すことにより実施され得る。単離は、例えば、多細胞生物から一部の細胞を分離して取り出すことにより実施され得る。単離された細胞を扱う技術を本明細書では、インビトロの技術という。
- [0012] 本明細書では、「精製」とは、単離された目的の細胞を、一緒に存在する他の構成要素からさらに分離することを意味する。精製は、例えば、目的の細胞を、形態や表面マーカーに基づいて他の構成要素から分離することにより実施され得る。精製は、細胞の限外希釈および／またはクローニングにより実施され得る。精製は、目的の細胞の株化により実施され得る。精製は、目的の細胞が薬剤耐性遺伝子や蛍光タンパク質をコードする遺伝子などのマーカー遺伝子を有する場合には、当該マーカー遺伝子の発現に基づいて実施

することができる。本明細書では、「富化」(e n r i c h)とは、目的の細胞の存在密度を向上させることを意味する。

[0013] 本明細書では、「ポリヌクレオチド」及び「核酸」という用語は、相互に互換的に使用され、ヌクレオチドがホスホジエステル結合によって結合したヌクレオチドポリマーを指す。「ポリヌクレオチド」及び「核酸」は、DNAであってもよく、RNAであってもよく、DNAとRNAとの組み合わせから構成されてもよい。また、「ポリヌクレオチド」及び「核酸」は、天然ヌクレオチドのポリマーであってもよく、天然ヌクレオチドと非天然ヌクレオチド(天然ヌクレオチドの類似体、塩基部分、糖部分及びリン酸部分のうち少なくとも一つの部分が修飾されているヌクレオチド(例えば、ホスホチオエート骨格)等)とのポリマーであってもよく、非天然ヌクレオチドのポリマーであってもよい。

[0014] 本明細書では、「ポリヌクレオチド」又は「核酸」の塩基配列は、特に明示しない限り、一般的に認められている1文字コードで記載される。特に明示しない限り、塩基配列は、5'側から3'側に向かって記載する。「ポリヌクレオチド」又は「核酸」を構成するヌクレオチド残基は、単に、アデニン、チミン、シトシン、グアニン、又はウラシル等、あるいはそれらの1文字コードで記載される場合がある。

[0015] 本明細書では、「遺伝子」という用語は、特定のタンパク質をコードする少なくとも1つのオープンリーディングフレームを含むポリヌクレオチドを指す。遺伝子は、エクソン及びイントロンの両方を含み得る。

[0016] 本明細書では、「ポリペプチド」、「ペプチド」及び「タンパク質」という用語は、相互に互換的に使用され、アミド結合によって結合したアミノ酸のポリマーを指す。「ポリペプチド」、「ペプチド」又は「タンパク質」は、天然アミノ酸のポリマーであってもよく、天然アミノ酸と非天然アミノ酸(天然アミノ酸の化学的類似体、修飾誘導体等)とのポリマーであってもよく、非天然アミノ酸のポリマーであってもよい。特に明示しない限り、アミノ酸配列は、N末端側からC末端側に向かって記載する。

[0017] 本明細書では、「アレル」という用語は、染色体ゲノム上の同一座位に存在する塩基配列のセットを指す。ある態様では、2倍体の細胞では同一座位に2アレル存在し、3倍体の細胞では同一座位に3アレル存在する。また、ある態様では、染色体の異常なコピーまたは当該座位の異常な追加のコピーによって追加のアレルが形成されている場合がある。

[0018] 本明細書では、「ゲノム改変」又は「ゲノム編集」という用語は、相互に互換的に用いられ、ゲノム上の所望の位置（標的領域）に変異を誘導することを指す。ゲノム改変は、標的領域DNAを切断するように設計された配列特異的核酸切断分子（例えば、配列特異的または配列依存的エンドヌクレアーゼ）の使用を含み得る。好ましい実施形態において、ゲノム改変は、標的領域のDNAを切断するように操作されたヌクレアーゼの使用、を含み得る。好ましい実施形態において、ゲノム改変は、標的領域中の特定の塩基配列を有する標的配列を切断するように操作されたヌクレアーゼ（例えば、TALENやジンクフィンガーヌクレアーゼ（ZFN））の使用を含み得る。特に好ましい実施形態において、ゲノム改変は、標的領域中の特定の塩基配列を有する標的配列を切断するように操作されたヌクレアーゼ（例えば、CRISPR-Cas9システム）の使用を含み得る。好ましい実施形態において、ゲノム改変は、標的領域中の特定の塩基配列を有する標的配列を切断するように、メガヌクレアーゼなどのゲノムに1つしか切断部位を有しない制限酵素（例えば、16ベースの配列特異性を有する制限酵素（理論上は 4^{16} 塩基に1つの割合で存在する）、17ベースの配列特異性を有する制限酵素（理論上は 4^{17} 塩基に1つの割合で存在する）、および18ベースの配列特異性を有する制限酵素（理論上は 4^{18} 塩基に1つの割合で存在する））などの配列特異的エンドヌクレアーゼを用いることができる場合もある。典型的には、部位特異的ヌクレアーゼの使用により、標的領域のDNAに二本鎖切断（DSB）が誘導され、その後、相同組み換え修復（Homologous Directed Repair: HDR）及び非同源末端再結合（Non-Homologous End-Joining Repair: N

HEJ)のような、細胞の内因性プロセスによってゲノムが修復される。NHEJは、ドナーDNAを用いずに二本鎖切断された末端を連結する修復方法であり、修復の際に挿入及び／又は欠失(indel)が高頻度で誘導される。HDRは、ドナーDNAを用いた修復機構であり、標的領域に所望の変異を導入することも可能である。ゲノム改変技術としては、例えば、CRISPR/Casシステムが好ましく例示される。メガヌクレアーゼとしては、例えば、I-SceI、I-SceII、I-SceIII、I-SceIV、I-SceV、I-SceVI、I-SceVII、I-CeuI、I-CeuAIIIP、I-CreI、I-CrepsbIP、I-CrepsbIIIP、I-CrepsbIIIP、I-CrepsbIIP、I-TliI、I-PpoI、PI-PspI、F-SceI、F-SceII、F-SuvI、F-TevI、F-TevII、I-Amal、I-Anil、I-ChuI、I-Cmoel、I-Cpal、I-Cpall、I-Csml、I-CvuI、I-CvuAIP、I-DdiI、I-DdiII、I-DiriI、I-DmOI、I-HmUI、I-HmUII、I-HsNIP、I-LIaI、I-MsOI、I-Naai、I-Nani、I-NcIIP、I-NgrIP、I-NitI、I-NjaI、I-Nsp236IP、I-PakI、I-PboIP、I-PcuIP、I-PcuAI、I-PcuVI、I-PgrIP、I-PobIP、I-Pori、I-PoriIIP、I-PbpIP、I-SpBetaIP、I-ScaI、I-SexIP、I-SneIP、I-SpomI、I-SpomCP、I-SpomIP、I-SpomIIP、I-SquIP、I-Ssp68031、I-SthPhiJP、I-SthPhiST3P、I-SthPhiSTe3bP、I-TdeIP、I-TevI、I-TevII、I-TevIII、I-UarAP、I-UarHGPAIP、I-UarHGPA13P、I-VinIP、I-ZbiIP、PI-MtuI、PI-MtuHIP、PI-MtuHIIP、PI-PfuI、PI-PfuII、PI-PkoI、PI-PkoII、PI-Rma43812IP、PI

—SpBetalP、Pl—SceI、Pl—TfuI、Pl—TfuII、Pl—ThyI、Pl—TliI、およびPl—TliII、並びにこれらの機能的な誘導体制限酵素からなる群から選択されるメガヌクレアーゼおよびその切断部位（または認識部位）、好ましくは、18塩基以上の配列特異性を有する制限酵素であるメガヌクレアーゼおよびその切断部位（または認識部位）、特に細胞のゲノムを1箇所または複数箇所以上切断しないメガヌクレアーゼおよびその切断部位を用いることができる。

[0019] 「標的領域」という用語は、ゲノム改変システムによりターゲティングされる領域をいう。本発明では、2箇所の標的領域（例えば、第1の標的領域および第2の標的領域）の間に位置するゲノム上のDNA領域を欠失させ得る。

[0020] 「配列特異的核酸切断分子」という用語は、特定の核酸配列を認識し、前記特定の核酸配列で核酸を切断することができる分子を指す。配列特異的核酸切断分子は、配列特異的に核酸切断する活性（配列特異的核酸切断活性）を有する分子である。

[0021] 「標的配列」という用語は、配列特異的核酸切断分子による切断の対象となるゲノム中のDNA配列を指す。配列特異的核酸切断分子がCasタンパク質である場合、標的配列は、Casタンパク質による切断の対象となるゲノム中のDNA配列を指す。Casタンパク質としてCas9タンパク質を用いる場合、標的配列は、プロトスペーサー隣接モチーフ（PAM）の5'側に隣接する配列である必要がある。標的配列は、通常、PAMの5'側直前に隣接する17～30塩基（好ましくは18～25塩基、より好ましくは19～22塩基、さらに好ましくは20塩基）の配列が選択される。標的配列の設計には、CRISPR DESIGN (crispr.mit.edu/) 等の公知のデザインツールを利用することができる。

[0022] 「Casタンパク質」という用語は、CRISPR関連（CRISPR-associated）タンパク質を指す。好ましい態様において、Casタンパク質は、ガイドRNAと複合体を形成し、エンドヌクレアーゼ活性又

はニッカーゼ活性を示す。C a s タンパク質としては、特に限定されないが、例えば、C a s 9 タンパク質等が挙げられる。C a s タンパク質は、ガイドRNAと協働してエンドヌクレアーゼ活性又はニッカーゼ活性を示す限り、野生型C a s タンパク質及びそのホモログ（パラログ及びオーソログ）、並びにそれらの変異体を包含する。

好ましい態様において、C a s タンパク質は、クラス2のCRISPR／C a s 系に關与するものであり、より好ましくはII型のCRISPR／C a s 系に關与するものである。C a s タンパク質の好ましい例としては、C a s 9 タンパク質が例示される。

[0023] 「C a s 9 タンパク質」という用語は、II型のCRISPR／C a s 系に關与するC a s タンパク質を指す。C a s 9 タンパク質は、ガイドRNAと複合体を形成し、ガイドRNAと協働して標的領域のDNAを切断する活性を示す。C a s 9 タンパク質は、前記の活性を有する限り、野生型C a s 9 タンパク質及びそのホモログ（パラログ及びオーソログ）、並びにそれらの変異体を包含する。野生型C a s 9 タンパク質は、ヌクレアーゼドメインとしてR u v Cドメイン及びHNHドメインを有するが、本明細書におけるC a s 9 タンパク質は、R u v Cドメイン及びHNHドメインのいずれか一方が不活性化されたものであってもよい。R u v Cドメイン及びHNHドメインのいずれか一方が不活性化されたC a s 9 は、二本鎖DNAに対して一本鎖切断（ニック）を導入する。そのため、R u v Cドメイン及びHNHドメインのいずれか一方が不活性化されたC a s 9 を二本鎖DNAの切断に用いる場合には、センス鎖とアンチセンス鎖それぞれに対してC a s 9 の標的配列を設定し、センス鎖およびアンチセンス鎖のニックが十分に近い位置で生じ、それにより二本鎖切断が誘発されるように、改変システムを構成することができる。

C a s 9 タンパク質が由来する生物種は特に限定されないが、ストレプトコッカス（*Streptococcus*）属、スタフィロコッカス（*Staphylococcus*）属、ナイセリア（*Neisseria*）属、又は

トレポネーマ (*Treponema*) 属に属する細菌等が好ましく例示される。より具体的には、*S. pyogenes*、*S. thermophilus*、*S. aureus*、*N. meningitidis*、又は *T. denticola* 等に由来する Cas9 タンパク質が好ましく例示される。好ましい態様において、Cas9 タンパク質は、*S. pyogenes* 由来の Cas9 タンパク質である。

[0024] 「ガイドRNA」及び「gRNA」という用語は、相互に互換的に使用され、Cas タンパク質と複合体を形成し、Cas タンパク質を標的領域に誘導することができるRNAを指す。好ましい態様において、ガイドRNAは、CRISPR RNA (*crRNA*) 及びトランス活性化型CRISPR RNA (*tracrRNA*) を含む。*crRNA*は、ゲノム上の標的領域への結合に関与し、*tracrRNA*は、Cas タンパク質との結合に関与する。好ましい態様において、*crRNA*は、スペーサー配列とリピート配列とを含み、スペーサー配列が標的領域において標的配列の相補鎖と結合する。好ましい態様において、*tracrRNA*は、アンチリピート配列と3' テイル配列とを含む。アンチリピート配列は*crRNA*のリピート配列と相補的な配列を有し、リピート配列と塩基対を形成し、3' テイル配列は通常3つのステムループを形成する。

ガイドRNAは、*crRNA*の3' 末端に*tracrRNA*の5' 末端を連結した単一ガイドRNA (*sgRNA*) であってもよく、*crRNA*及び*tracrRNA*を別々のRNA分子とし、リピート配列及びアンチリピート配列で塩基対を形成させたものであってもよい。好ましい態様において、ガイドRNAは*sgRNA*である。

[0025] *crRNA*のリピート配列及び*tracrRNA*の配列は、Cas タンパク質の種類に応じて適宜選択することができ、Cas タンパク質と同じ細菌種に由来するものを用いることができる。CRISPR-Cas9 システムとしては、*S. pyogenes* 由来の Cas9 タンパク質、*crRNA* および *tracrRNA* (または *sgRNA*) を用いることができる。*sgR*

NA設計のためのc rRNAリピート配列及びt r a c rRNAの配列は、種々提案されており、当業者は、公知技術に基づいてs gRNAを設計することができる（例えば、Jinek et al. (2012) Science, 337, 816-21; Mali et al. (2013) Science, 339: 6121, 823-6; Cong et al. (2013) Science, 339: 6121, 819-23; Hwang et al. (2013) Nat. Biotechnol. 31: 3, 227-9; Jinek et al. (2013) eLife, 2, e00471）。

[0026] ポリヌクレオチドに関して用いる「作動可能に連結」という用語は、第一の塩基配列が第二の塩基配列に十分に近くに配置され、第一の塩基配列が第二の塩基配列又は第二の塩基配列の制御下の領域に影響を及ぼしうることを意味する。例えば、ポリヌクレオチドがプロモーターに機能的に連結するとは、当該ポリヌクレオチドが、当該プロモーターの制御下で発現するように連結されていることを意味する。

[0027] 「発現可能な状態」という用語は、ポリヌクレオチドが導入された細胞内で、該ポリヌクレオチドが転写され得る状態にあることを指す。

「発現ベクター」という用語は、対象ポリヌクレオチドを含むベクターであって、該ベクターを導入した細胞内で、対象ポリヌクレオチドを発現可能な状態にするシステムを備えたベクターを指す。例えば、「Casタンパク質の発現ベクター」とは、該ベクターを導入した細胞内で、Casタンパク質を発現可能なベクターを意味する。また、例えば、「ガイドRNAの発現ベクター」とは、該ベクターを導入した細胞内で、ガイドRNAを発現可能なベクターを意味する。

[0028] 本明細明細書において、塩基配列どうし又はアミノ酸配列どうしの配列同一性（又は相同性）は、2つの塩基配列又はアミノ酸配列を、対応する塩基又はアミノ酸が最も多く一致するように、挿入及び欠失に当たる部分にギャップを入れながら並置し、得られたアラインメント中のギャップを除く、塩

基配列全体又はアミノ酸配列全体に対する一致した塩基又はアミノ酸の割合として求められる。塩基配列又はアミノ酸配列どうしの配列同一性は、当該技術分野で公知の各種相同性検索ソフトウェアを用いて求めることができる。例えば、塩基配列の配列同一性の値は、公知の相同性検索ソフトウェアBLASTNにより得られたアライメントを元にした計算によって得ることができ、アミノ酸配列の配列同一性の値は、公知の相同性検索ソフトウェアBLASTPにより得られたアライメントを元にした計算によって得ることができる。

[0029] <本発明の方法>

本発明によれば、

(a) 細胞を含む細胞集団を用意することと、

(b) 前記細胞集団中の前記細胞のゲノムDNAに、ゲノムDNA上の2箇所の標的配列を配列特異的に切断することができる配列特異的核酸切断分子をそれぞれ作用させて、ゲノムDNA上の2箇所の標的配列においてそれぞれ切断を生じさせ、これによって、細胞集団中の少なくとも一部の細胞において、前記2箇所の切断部位の間の領域においてDNA領域の欠失を生じさせることと、

を含む方法

が提供される。本発明の方法は、インビトロの方法であり得る。本発明の方法はまた、単離された細胞を好ましく用い得る。

[0030] したがって、本発明によれば、インビトロの方法であって、

(a) 単離された細胞を含む細胞集団を用意することと、

(b) 前記細胞集団中の前記細胞のゲノムDNAに、ゲノムDNA上の2箇所の標的配列を配列特異的に切断することができる配列特異的核酸切断分子をそれぞれ作用させて、ゲノムDNA上の2箇所の標的配列においてそれぞれ切断を生じさせ、これによって、細胞集団中の少なくとも一部の細胞において、前記2箇所の切断部位の間の領域においてDNA領域の欠失を生じさせることと、

を含む方法が提供される。

[0031] 本発明の一態様によれば、本発明の方法は、

(c) 前記DNA領域の欠失を生じた細胞を含む細胞集団において、全細胞に対する前記DNA領域の欠失を生じた細胞の割合を決定すること
をさらに含む。

[0032] 上記(a)において、細胞集団に含まれる細胞は、単一種類であってもよいし、複数種類であってもよい。好ましくは、細胞集団に含まれる細胞は、単一種類（例えば、細胞株、クローン化された細胞）であり得る。細胞集団は、浮遊細胞の集団であってもよいし、接着細胞の集団であってもよい。細胞集団は、単一細胞化された細胞の集団であってもよいし、細胞塊を含む細胞集団であってもよい。細胞集団は、細胞と生理学的に許容可能な賦形剤とを含み得る。生理学的に許容可能な賦形剤としては、細胞の維持に適切な条件を有する賦形剤であり、例えば、水、塩、pH緩衝剤、等張剤などが挙げられる。細胞集団は、例えば、 10^2 以上、 10^3 以上、 10^4 以上、 10^5 以上、または 10^6 以上の細胞を含み得る。

[0033] 本発明のある態様では、細胞は、単離された細胞であり得る。細胞は、精製された細胞であってもよい。細胞は、単細胞生物の細胞であり得る。細胞は、細胞株であり得る。細胞は、クローン化された細胞（細胞クローン）であり得る。細胞は、単一細胞化された細胞であり得る。細胞は、浮遊細胞であり得る。細胞は、接着細胞であり得る。細胞は、細胞塊を形成していてもよい。細胞は、コロニーを形成していてもよい。細胞は、その維持または増殖に適した条件下におかれる。

[0034] ある態様では、細胞は、多能性細胞、および多能性幹細胞（胚性幹細胞および誘導多能性幹細胞など）からなる群から選択される細胞であり得る。ある態様では、細胞は、組織幹細胞であり得る。ある態様では、細胞は、体細胞であり得る。ある態様では、細胞は、組織前駆細胞であり得る。ある態様では、細胞は、生殖系列細胞（例えば、生殖細胞）であり得る。ある態様では、細胞は、細胞株であり得る。ある態様では、細胞は不死化細胞であり得る。

る。ある態様では、細胞はがん細胞であり得る。ある態様では、細胞は、非がん細胞であり得る。ある態様では、細胞は、疾患患者の細胞であり得る。ある態様では、細胞は健常者の細胞であり得る。ある態様では、細胞は、外来の病原体に感染した細胞であり得る。ある態様では、細胞は、非感染細胞であり得る。ある態様では、細胞は、動物細胞（例えば、鳥類細胞、例えば、魚類細胞、例えば、両生類細胞、例えば、は虫類細胞、例えば、哺乳動物細胞、例えば、齧歯類細胞、例えば、霊長類細胞、例えば、ヒト細胞）であり得る。ある態様では、細胞は、例えば、昆虫細胞（例えば、カイコ細胞）、HEK293細胞、HEK293T細胞、Exp293F（商標）細胞、FreeStyle（商標）293F細胞、チャイニーズハムスター卵巣細胞（CHO細胞）、CHO-S細胞、CHO-K1細胞、およびExp1CHO細胞、ならびにこれらの細胞からの派生細胞からなる群から選択される細胞（特に、組換えタンパク質産生細胞）であり得る。ある態様では、細胞は、植物の細胞であり得る。ある態様では、細胞としてはまた、微生物の細胞、例えば、糸状菌、放線菌などのグラム陽性菌、大腸菌等のグラム陰性菌、および酵母などの真菌などの微生物の細胞が挙げられる。

[0035] 本発明で用いられる細胞は、好ましくは、1倍体の細胞であっても、2倍体の細胞であってもよい。細胞はまた、それ以外の倍数体の細胞であっても、特に限定無く用いることができる。本発明のすべての実施態様において1倍体の細胞は好ましく用いられ得る。

[0036] 上記（b）において、細胞のゲノムDNA（特に宿主ゲノムDNA）に対して、ゲノムDNA上の2箇所の標的配列を配列特異的に切断することができる配列特異的核酸切断分子をそれぞれ作用させることができる。標的配列は、用いる配列特異的核酸切断分子の切断特性に鑑みて、当業者であれば適宜設定することができる。標的配列は、少なくとも1つのアレルの塩基配列に存在し、好ましくは、2つ以上またはすべてのアレルに共通して存在する配列であり得る。ゲノム上の複数の標的配列を切断するには、例えば、CRISPR-Cas9システム、TALEN、およびジンクフィンガーヌクレ

アーゼはいずれも好ましく用いられ得る。具体的には、用いるガイドRNA (crRNAおよびtracrRNAの組合せ、またはsgRNA) を標的配列の増加に応じて増やすだけで、複数箇所の切断ができるためである (例えば、WO2014/093661参照)。切断にCRISPR-Cas9システム以外のゲノム改変システムを用いる場合は、標的配列毎に切断分子を用意する。

[0037] ゲノムDNA (特に宿主ゲノムDNA) に切断が生じると、細胞のゲノム修復機構により、切断部位が再び連結される。ゲノムDNAに2箇所の切断を生じさせると、ゲノム修復機構により一定の確率でゲノムDNAのテロメア側の切断末端とセントロメア側の切断末端とが直接連結され、2箇所の切断部位の間のDNA領域は、ゲノムDNAから欠失し得る。

[0038] したがって、上記(b)において、前記2箇所の切断部位の間の領域において生じたDNA領域の欠失を有する少なくとも1つの細胞 (好ましくは複数の細胞) を含む細胞集団が取得される。

[0039] 上記(b)において、2箇所の切断部位は、1つまたは複数の遺伝子または遺伝子候補を含む領域を挟むように設計され得る。2箇所の切断部位に含まれる1つまたは複数の遺伝子または遺伝子候補は、1つまたは複数の細胞の生存および／または増殖を正に制御する遺伝子 (例えば、細胞の生存および／または増殖に必須の遺伝子) を含んでいてもよい。2箇所の切断部位に含まれる1つまたは複数の遺伝子または遺伝子候補は、1つまたは複数の細胞の増殖を抑制する遺伝子などの、欠失すると細胞増殖が促進される1つまたは複数の遺伝子を含んでいてもよい。細胞の生存および／または増殖を正に制御する遺伝子 (例えば、細胞の生存および／または増殖に必須の遺伝子) は、当業者は予め把握しておく必要は必ずしもない。しかし、細胞の生存および／または増殖を正に制御する遺伝子 (例えば、細胞の生存および／または増殖に必須の遺伝子) は、Online Gene Essentiality (OGEE) データベース (<http://ogee.medicgenius.info/browse/>) などの公開データベースを用いて予

め推定しておくことができる。

- [0040] ある態様では、1つまたは複数の遺伝子または遺伝子候補を含む領域は、細胞の生存および／または増殖を正に制御する遺伝子（例えば、細胞の生存および／または増殖に必須の遺伝子）と推定される遺伝子を含んでいてもよい。ある態様では、1つまたは複数の遺伝子または遺伝子候補を含む領域は、細胞の生存および／増殖を負に制御する遺伝子またはそのように推定される遺伝子を含んでいてもよい。ある態様では、1つまたは複数の遺伝子または遺伝子候補を含む領域は、細胞の生存および／または増殖を正に制御する遺伝子（例えば、細胞の生存および／または増殖に必須の遺伝子）またはそのように推定される遺伝子と、細胞の生存および／増殖を負に制御する遺伝子またはそのように推定される遺伝子とを含んでいてもよい。
- [0041] 上記（b）において、2箇所の切断部位は、特に限定されないが、0.1 Mb以上、0.2 Mb以上、0.3 Mb以上、0.4 Mb以上、0.5 Mb以上、0.6 Mb以上、0.7 Mb以上、0.8 Mb以上、0.9 Mb以上、1 Mb以上、2 Mb以上、3 Mb以上、4 Mb以上、または5 Mb以上の領域を挟むように設計されていてもよい。すなわち、ゲノムDNAから欠失するDNA領域は、0.1 Mb以上、0.2 Mb以上、0.3 Mb以上、0.4 Mb以上、0.5 Mb以上、0.6 Mb以上、0.7 Mb以上、0.8 Mb以上、0.9 Mb以上、1 Mb以上、2 Mb以上、3 Mb以上、4 Mb以上または5 Mb以上の長さを有していてもよい。
- [0042] これにより、上記（b）では、前記2箇所の切断部位の間の領域において生じたDNA領域の欠失を有する細胞を含む細胞集団が得られ得る。2箇所の切断を有するゲノムDNAなどのDNAは、細胞内で修復され得るが、この修復により、一定確率で、前記2箇所の切断部位の間の領域において生じたDNA領域の欠失を有するゲノムDNAなどのDNAを有する細胞が生じ得る。したがって、上記により、前記2箇所の切断部位の間の領域において生じたDNA領域の欠失を有する細胞を含む細胞集団が得られるのである。
- [0043] 細胞集団から、前記2箇所の切断部位の間の領域において生じたDNA領

域の欠失を有する細胞を取得することができる。取得は、例えば、細胞集団を限外希釈してクローニングすることによって行うことができる。DNA領域を欠失したゲノムDNAにおいては、欠失した部分の前後が直接連結されていると考えられる。したがって、その連結部（ジャンクション）を挟むように増幅用プライマーを設計して、DNA領域を欠失したゲノムDNAを鋳型としたPCR法（ジャンクションPCR）によって、DNA領域の欠失の有無を決定することができる。また、連結部をシーケンシングすることによって、DNA領域の欠失の有無を決定することができる。

このとき、欠失したDNA領域が、細胞の生存および／または増殖を正に制御する遺伝子（例えば、細胞の生存および／または増殖に必須の遺伝子）を含まない場合には、当該DNA領域の欠失を有する細胞を取得することができる。また、欠失したDNA領域が、細胞の生存および／または増殖を正に制御する遺伝子（例えば、細胞の生存および／または増殖に必須の遺伝子）を含む場合には、当該欠失を有する細胞は、時間と共に増殖を停止するか、死滅し、細胞集団に占める細胞の割合を減少させる。あるいは、欠失したDNA領域が、細胞の生存および／増殖を負に制御する遺伝子を含む場合には、当該欠失を有する細胞は、時間と共に増殖し、細胞集団に占める割合を増加させる。

[0044] 上記（c）では、DNA領域の欠失を生じた細胞の割合に基づいて細胞の増殖または生存に対する前記DNA領域の欠失の影響を決定することができる。

細胞の増殖または生存に対する前記DNA領域の欠失の影響は、種々の方法により決定することができる。

[0045] 前記（c）は、工程（b）に引き続き行うことができる。ここで、工程（b）と（c）の間では、細胞の選別（特定の細胞を他の細胞から単離すること、例えば、細胞集団から一部の細胞を選別すること）を実施しないことができる。すなわち、工程（b）において得られた細胞集団をそのまま培養して、工程（c）を実施することができる。培養は、欠失前の細胞の培養（ま

たは増殖)に適した条件下で行うことができる。あるいは、工程(b)と(c)の間では、細胞の選別工程を実施してもよい。

[0046] 前記(c)において、前記決定は、DNA領域の欠失効率またはその推定値を決定することと、その後の培養後の細胞集団におけるDNA領域の欠失を有する細胞の割合と、決定された欠失効率またはその推定値とを比較することにより行われ得る。

[0047] 前記(c)において、欠失効率またはその推定値の決定、および、培養後のDNA領域の欠失を有する細胞の割合は、それぞれ細胞集団を含む懸濁液中に含まれる全細胞に対する欠失を有する細胞の割合として決定され得る。上記割合は、懸濁液中に含まれる、欠失を有するゲノムDNAおよび有しないゲノムDNAの計数技術により決定され得る。

[0048] 細胞の増殖または生存に対するDNA領域の欠失の影響は、欠失の導入効率(またはその推定値)と、続く培養後のDNA領域の欠失を有する細胞の割合とから決定することができる。例えば、上記(c){以下、DNA領域の欠失の影響を決定する工程でも同様である}は、欠失の導入効率(またはその推定値)を算出することを含み得る。DNA領域の欠失の導入効率は、上記(b)の処理後、DNA領域の欠失の効果が細胞に生じる前に行うことができる。DNA領域の欠失をした直後には、細胞内に当該欠失したDNA領域に由来する転写産物や翻訳産物が残存しているため、DNA領域の欠失の影響は少ない。したがって、上記(b)の処理後、例えば、2時間後～3日後、4時間～3日後、6時間～3日後、8時間～3日後、12時間～3日後、18時間～3日後、1日後～2日後、1日後～3日後、1日～60時間後、4時間～60時間後、4時間～48時間後、4時間～36時間後、4時間～30時間後、または4時間～2日後にDNA領域の欠失の導入効率を確認することができる。上記(c){以下、DNA領域の欠失の影響を決定する工程でも同様である}はまた、得られた細胞を培養することを含み得る。培養は、DNA領域の欠失処置前の細胞の培養に適した条件下で行われ得る。例えば、上記(b)の後で得られる、前記DNA領域の欠失を有する細胞

を含む細胞集団を培養し、一定細胞数あたりの増殖した細胞の数、または増殖した細胞が形成するコロニーの数を計数することによって評価することができる。具体的には、一定細胞数あたりの増殖した細胞の数、または増殖した細胞が形成するコロニーの数が増加したこと（または、欠失を有するコロニーの割合または細胞の割合が欠失の導入効率より低下したこと）は、前記欠失したDNA領域に、細胞の生存および／または増殖を制御する遺伝子が含まれることを意味する。より具体的には、一定細胞数あたりの増殖した細胞の数、または増殖した細胞が形成するコロニーの数が増加したことは、前記DNA領域の欠失は、細胞の増殖または生存に正の影響を与えたと評価すること、または前記DNA領域が、細胞の生存および／または増殖を負に制御する遺伝子を含んでいたと評価することができる。また、一定細胞数あたりの増殖した細胞の数、または増殖した細胞が形成するコロニーの数が小さいことまたは減少したことは、前記DNA領域の欠失は、細胞の増殖または生存に負の影響を与えたと評価すること、または前記DNA領域が、細胞の生存および／または増殖を正に制御する遺伝子（特に、細胞の生存および／または増殖に必須の遺伝子）を含んでいたと評価することができる。前記培養は、改変前の細胞の培養に適した条件下においてなされ得る。形成されたコロニー数に対するDNA領域の欠失を有するコロニーの割合を調べてもよい。

[0049] また、例えば、欠失の導入効率（またはその推定値）および細胞集団中の前記DNA領域の欠失を有する細胞数の割合をデジタルPCRや分子バーコードによる核酸のデジタル計数技術などの計数手法を用いて決定することもできる。前記計測手法は、ある細胞を他の細胞から分離すること、および／または、細胞にコロニーを形成させることを含まないでよい。すなわち、欠失の導入効率（またはその推定値）および細胞集団中の前記DNA領域の欠失を有する細胞数の割合を細胞懸濁液を用いて（または細胞懸濁液から抽出したゲノムDNAまたはその増幅産物を用いて）算出することができる。

[0050] デジタルPCRは、核酸（ここではゲノムDNA）を1ウェルあたりのゲ

ノム数が一定となる濃度および量で複数の微細な仕切りに核酸を分注し、PCR反応に供して、PCR反応を生じたウェル数を計数することにより、鋳型核酸の絶対定量を行う手法である。マイクロ流体デバイス、油中液滴を用いたドロップレット法を用いて核酸を1分子を含む多数の画分に分画することができ、それぞれの1分子を含む画分で並行して核酸増幅反応を引き起こす。増幅する鋳型が存在しない場合には、増幅産物が生じず、増幅する鋳型が存在する場合には、増幅産物が生じることを利用して、画分毎に増幅の有無を計数することにより、デジタル的にサンプル中に増幅する鋳型が存在していたか、またはその存在量を決定することができる。PCRプライマーは、DNA領域の欠失を生じないときに増え、欠失を有するときに増えないように設計することができ、または、DNA領域の欠失を生じないときに増えず、欠失を有するときに増えるように設計することができる。

分子バーコードによる核酸のデジタル計数技術は、核酸の断片に対して、1分子毎に固有の分子バーコード（具体的には、固有の塩基配列）を付加し、その後、その配列をシーケンシングすることによって、核酸分子の数を絶対定量する技術である。分子バーコードによる核酸のデジタル計数技術においては、分子バーコードの種類の数、核酸分子数に対応する。

[0051] 前記DNA領域が、細胞の生存および／または増殖を正に制御する遺伝子（特に、細胞の生存および／または増殖に必須の遺伝子）を含んでいる場合には、培養と共に細胞は死滅する。したがって、例えば、その死滅の前後で、細胞集団中の前記DNA領域の欠失を有する細胞数の割合を決定すること（または、欠失を有するコロニーの割合または細胞の割合が欠失の導入効率より低下したこと）によって、細胞の増殖または生存に対する前記DNA領域の欠失の影響を評価することができる。

[0052] 細胞の増殖または生存に対する前記DNA領域の欠失の影響は、対照と比較して決定することができる。例えば、細胞の増殖または生存に対する前記DNA領域の欠失の影響は、対照の細胞の増殖または生存と比較して決定することができる。対照の細胞としては、例えば、前記DNA領域の一部また

は全部を有する細胞とすることができる。

[0053] 本発明のある態様では、上記（c）{以下、DNA領域の欠失の影響を決定する工程でも同様である}は、上記（b）で得られた細胞を培養することを含む。当該培養は、例えば、所定の条件下で行うことができる。所定の条件下としては、通常の細胞の培養条件下に加えて、例えば、ストレス存在下、増殖刺激存在下、分化誘導刺激存在下、低酸素条件下、増殖因子存在下、増殖阻害剤存在下、分化誘導因子存在下、分化阻害因子存在下、および薬剤存在下が挙げられる。このようにして、上記（b）で得られた細胞を所定の条件下で培養することによって、当該細胞の当該所定の条件下での振る舞いに対する、欠失したDNA領域の影響（役割）を決定することができる。また、特定のDNA領域の欠失を有する細胞の振る舞いに対する、培養条件（所定の条件）の影響を調べることもできる。例えば、特定のDNA領域の欠失を有する細胞の振る舞いに対する、薬剤添加の影響を調べることによって、薬剤の特性分析を行うことや薬剤のスクリーニングを行うことができる。

[0054] 本発明の方法は、上記（a）～（c）と、

（d）対照となる細胞集団と比較して、前記（c）における生存および／または増殖に対する影響の有無または大きさに基づいて、対照のゲノムDNAと比較して欠失したDNA領域が細胞の生存および／または増殖を制御する遺伝子を含むか否かを決定すること
とを含んでいてもよい。

[0055] 上記（d）において、対照または陰性対照は、未改変または改変前の細胞または細胞集団であり得る。上記（d）において、対照（または陰性対照）は、より小さなDNA領域の欠失を有する細胞または当該細胞を含む細胞集団であり得る。上記（d）において、対照（または陽性対照）は、より大きなDNA領域の欠失を有する細胞または当該細胞を含む細胞集団であり得る。上記（d）において、対照は、異なるDNA領域の欠失を有する細胞または当該細胞を含む細胞集団であり得る。上記細胞集団は、上記（b）により得られ得る。

[0056] 上記（d）において、未改変または改変前の細胞または細胞集団を対照とすることによって、欠失したDNA領域に含まれる遺伝子の、細胞の生存および／または増殖に対する影響を評価することができる。

上記（d）において、より小さなDNA領域の欠失を有する細胞または当該細胞を含む細胞集団を対照とすることによって、当該対照が含む遺伝子であって、評価対象となるDNAの欠失により欠失した遺伝子の、細胞の生存および／または増殖に対する影響を評価することができる。

上記（d）において、より大きなDNA領域の欠失を有する細胞または当該細胞を含む細胞集団を対照とすることによって、当該対照が含まない遺伝子であって、評価対象となるDNAの欠失を有する細胞が含む遺伝子の、細胞の生存および／または増殖に対する影響を評価することができる。

上記（d）において、異なるDNA領域の欠失を有する細胞または当該細胞を含む細胞集団を対照とすることによって、（d-1）当該対照が含む遺伝子であって、評価対象となるDNAの欠失により欠失した遺伝子および／または（d-2）当該対照が含まない遺伝子であって、評価対象となるDNAの欠失を有する細胞が含む遺伝子の、細胞の生存および／または増殖に対する影響を評価することができる。

[0057] 本発明はある態様では、上記（a）～（c）と、

（e）前記細胞の割合を少なくとも異なる2時点で決定することをさらに含み、時間経過に伴う前記（c）における細胞の生存および／または増殖に対する影響の有無または大きさに基づいて、欠失したDNA領域が少なくとも1つの細胞の生存および／または増殖を制御する遺伝子を含むか否かを決定すること
とを含み得る。

[0058] 上記（b）により上記DNA領域の欠失を含む細胞を含む細胞集団が得られる。細胞集団に含まれる当該欠失を含む細胞の割合は、ゲノムの改変効率に依存する。細胞集団に含まれる当該欠失を含む細胞は、培養することによってその数を増加、減少、または維持し得る。細胞は、ゲノムに改変が生じ

てからも、しばらくは細胞質内のゲノムDNAからの転写物（例えば、mRNA）の残存や、転写物からの翻訳産物（例えば、タンパク質）の残存により、改変前と同様の振る舞いを示し得る。しかしながら、しばらく培養すると、上記転写物の残存量やタンパク質の残存量が減少し、DNAの欠失による遺伝子型が表現型として発現するようになる。そして、その後、細胞の生存および／または増殖を制御する遺伝子による表現型が明らかになる。

[0059] 上記（e）において、2時点（第1の時点および第2の時点）に関して、第1の時点は、DNAの欠失による遺伝子型が表現型として発現する前であり得、第2の時点は、DNAの欠失による遺伝子型が表現型として発現した後であり得る。例えば、第1の時点は、ゲノム改変から3日以内であり得る。また例えば、第2の時点は、ゲノム改変から3日目以降であり得る。第1の時点と第2の時点との間隔は、例えば、1日以上、2日以上、3日以上、4日以上、5日以上、6日以上、または1週間以上の期間であり得る。当業者であれば、細胞の状態等に鑑みて適宜第1の時点と第2の時点を決定し、本発明を実施することができる。

[0060] 上記（e）において、前記細胞の割合を少なくとも異なる2時点で決定することをさらに含み、時間経過に伴う前記（c）における細胞の生存および／または増殖に対する影響の有無または大きさを評価する。

上記（e）において、時間経過に伴う前記（c）における細胞の生存および／または増殖に対する影響が認められた場合には、その原因となるDNA領域の欠失が、細胞の生存および／または増殖を制御する遺伝子を含んでいる可能性が示唆される。また、時間経過に伴う前記（c）における細胞の生存および／または増殖に対する影響が認められた場合には、その影響の大きさを評価することによって、その原因となるDNA領域の欠失が含む細胞の生存および／または増殖を制御する遺伝子の細胞の生存および／または増殖への影響力の大きさを評価することができる。影響の大きさは、対照と比較して評価することができる。対照としては、細胞の増殖因子をコードする遺伝子などの陽性対照および／または細胞の増殖抑制因子をコードする遺伝子

などの陽性対照が挙げられる。対照としては、細胞の生存および／または増殖に影響を与えないことが知られている因子をコードする遺伝子などの陰性対照を用いてもよい。

[0061] 欠失したDNA領域が複数の推定遺伝子（または遺伝子）を含む場合には、（f）そこからいずれの遺伝子が、細胞の生存および／または増殖を制御する遺伝子であるかを決定することを実施してもよい。いずれの遺伝子が、細胞の生存および／または増殖を制御する遺伝子であるかを決定することは、様々な方法を用いて実施することができる。例えば、特異的な遺伝子破壊、本発明による遺伝子の欠失による欠損領域の細分化、遺伝子クローニングとその遺伝子の機能解析などによって、いずれの遺伝子が、細胞の生存および／または増殖を制御する遺伝子であるかを決定することができる。

[0062] 欠失したDNA領域に、細胞の生存および／または増殖を制御する遺伝子が含まれた場合には、当該遺伝子をゲノムDNAの他の場所に異所的に導入することができる。したがって、本発明は、

（g）制御配列に作動可能に連結した細胞の生存および／または増殖を制御する遺伝子（特に細胞の生存および／または増殖を正に制御する遺伝子）の少なくとも1つを、当該DNA領域の欠失を有するゲノムDNAに異所的に導入すること

をさらに含んでいてもよい。

ここで異所的に導入するとは、内因性に位置する場所とは異なる場所に対して導入することを意味する。

[0063] 上記（b）と上記（g）によって、ゲノムDNAにDNA領域の欠失を生じさせながらも、細胞の生存および／または増殖を制御する遺伝子（特に細胞の生存および／または増殖を正に制御する遺伝子、特に細胞の生存に必須の遺伝子）を異所的に前記ゲノムDNAに導入することによって、細胞の生存および／または増殖に対する影響を低減させながら、ゲノムDNAにより大きなDNA領域の欠失を生じさせることができる。このようにして、細胞のゲノムDNAを生存に必須な遺伝子セットを有する状態となるまで最小化

することができる。

[0064] 本発明によればまた、

(α) 細胞を含む細胞集団を用意することと、ここで細胞は、欠失させる領域中にネガティブ選択マーカー遺伝子を含み、

(β) 前記細胞集団中の前記細胞のゲノムDNAに、ゲノムDNA上の2箇所の標的配列を配列特異的に切断することができる配列特異的核酸切断分子をそれぞれ作用させて、ゲノムDNA上の2箇所の標的配列においてそれぞれ切断を生じさせ、これによって、細胞集団中の少なくとも一部の細胞において、前記2箇所の切断部位の間の領域においてDNA領域の欠失を生じさせることと、ここで、前記2箇所は、前記ネガティブ選択マーカー遺伝子を挟む位置に設計され、

(γ) ネガティブ選択マーカー遺伝子を有しない細胞を選択することと、を含む、方法

が提供される。本発明の方法はインビトロの方法であり得、細胞は単離された細胞であり得る。ある態様では、本発明の方法はインビトロの方法であり、細胞は単離された細胞である。

[0065] 上記(α)は、細胞は、欠失させる領域中にネガティブ選択マーカー遺伝子を含むこと以外は、上記(a)と同一であるので、同一部分についての説明を省略する。ここで、選択マーカー発現細胞及び非発現細胞が混在する細胞集団において、選択マーカー非発現細胞を選択する場合、当該選択マーカーを「ネガティブ選択マーカー」または「ネガティブ選択用の選択マーカー」という。

[0066] ネガティブ選択マーカーは、それを発現しない細胞を選択可能なものであれば、特に限定されない。ネガティブ選択マーカー遺伝子としては、例えば、自殺遺伝子(チミジンカイネース(TK))等)、蛍光タンパク質遺伝子、発光酵素遺伝子、発色酵素遺伝子等が挙げられる。ネガティブ選択マーカー遺伝子が、細胞の生存に負の影響を与える遺伝子(例えば、自殺遺伝子)である場合、当該ネガティブ選択マーカー遺伝子は、誘導性プロモーターに

機能的に連結され得る。誘導性プロモーターに機能的に連結することで、ネガティブ選択マーカー遺伝子を有する細胞を除去したいときにのみ、ネガティブ選択マーカー遺伝子を発現させることができる。ネガティブ選択マーカー遺伝子が、蛍光、発光、および発色等の光学的に検出可能なマーカー遺伝子（可視化マーカー遺伝子；*visible marker gene*）である場合など、細胞の生存に負の影響が少ない場合には、恒常的に発現させてもよい。

[0067] 上記（α）では、細胞は、欠失させるDNA領域中にネガティブ選択マーカー遺伝子（ネガティブ選択に用いることができるマーカー遺伝子である）を含む。欠失させるDNA領域は、ネガティブ選択マーカー遺伝子を内在的に有する部分とするか、または欠失させるDNA領域に外来的にネガティブ選択マーカー遺伝子を導入することによって作製され得る。ネガティブ選択マーカー遺伝子を内在的に有する部分は、例えば、X染色体のHPR T1遺伝子（特に、ヒトX染色体のヒトHPR T1遺伝子）の存在する領域（特に、ヒトX染色体のq25-26のDNA領域）であり得る。欠失させる領域への外来的にネガティブ選択マーカー遺伝子の導入は、例えば、遺伝子改変技術（例えば、HDR、ゲノム編集システム）を用いて行うことができる。例えば、ネガティブ選択マーカー遺伝子が可視化マーカー遺伝子である場合には、例えば、可視化マーカー遺伝子は、ゲノムDNAの特定DNA領域に挿入された後に発現させることで、その発光強度から1つのアレルに可視化マーカー遺伝子が導入されたか、複数のアレルに可視化マーカー遺伝子が導入されたかを判断することができる。特定のDNA領域に可視化マーカー遺伝子が挿入された細胞をクローニングすることによって、欠失させる領域中にネガティブ選択マーカー遺伝子を含む細胞を得ることができる。あるいは、複数のアレルに対して、それぞれ相互に区別可能な固有の選択マーカー遺伝子とネガティブ選択マーカー遺伝子とを挿入し、当該それぞれのアレルに組込まれた固有の選択マーカーの発現を指標として細胞を選択することで、複数のアレルにネガティブ選択マーカー遺伝子が組込まれた細胞を取得する

ことができる。特定のDNA領域への遺伝子の挿入は、ゲノム改変システムを用いて当該DNA領域中の標的配列を切断することと、切断部位の上流と相同組換え可能な上流ホモロジーアームと切断部位の下流と相同組換え可能な下流ホモロジーアームとを備え、上流ホモロジーアームと下流ホモロジーアームとの間に、挿入したい遺伝子を含むドナーDNAによるHDRを誘発させることとにより達成できる。ネガティブ選択マーカー遺伝子が自殺遺伝子などの細胞の生存に負の影響を与える遺伝子である場合には、クローニングした後に、それぞれのクローンに自殺遺伝子を発現させて、自殺する細胞は、少なくとも1つのアレルにネガティブ選択マーカー遺伝子が導入された細胞である。他のアレルにもネガティブ選択マーカー遺伝子を導入する場合には、別の誘導性プロモーターに機能的に（作動可能に）連結し、他のアレルにネガティブ選択マーカー遺伝子を導入し、当該プロモーターを駆動した場合にも細胞が死滅しないことを確認すればよい。細胞は、2倍体細胞、好ましくは1倍体細胞を用いることができる。アッセイ系の簡便性の観点から、1倍体の細胞としては、例えば、HAP1細胞を用いることができる。

[0068] 上記（ β ）では、上記（ α ）においてネガティブ選択用マーカー遺伝子が挿入されたDNA領域を挟むように、2箇所の標的配列を決定し、これを配列特異的に切断することができる配列特異的核酸切断分子を設計することができる。ゲノムDNAに配列特異的核酸切断分子を作用させると、2箇所の標的配列の間のDNA領域がゲノムDNAから欠失する。2箇所の標的配列の間のDNA領域は、ネガティブ選択用マーカー遺伝子を含むため、上記欠失によって、ネガティブ選択用マーカー遺伝子も一緒に欠失する。

[0069] 上記（ γ ）では、上記（ β ）によって得られた、2箇所の標的配列の間のDNA領域の欠失を有する細胞を選択する。2箇所の標的配列の間のDNA領域の欠失を有する細胞は、上述のようにネガティブ選択用マーカー遺伝子を有しない。したがって、ネガティブ選択用マーカー遺伝子が誘導される条件下で細胞を維持しても、ネガティブ選択用マーカー遺伝子は発現することがない。したがって、ネガティブ選択用マーカー遺伝子が誘導される条件下

で、細胞を維持し、ネガティブ選択用マーカー遺伝子が発現しないことを指標として、2箇所の標的配列の間のDNA領域の欠失を有する細胞を選択することができる。ネガティブ選択用マーカー遺伝子が、細胞を死滅させる遺伝子（自殺遺伝子）である場合には、これらの遺伝子は、誘導性プロモーターに作動可能に連結されていてよく、当該誘導性プロモーターに対して誘導因子を作用させてネガティブ選択用マーカー遺伝子の発現を誘導して、DNA領域の欠失を有しない細胞を除去することができる。ネガティブ選択用マーカー遺伝子が、可視化マーカー遺伝子である場合には、可視化マーカー遺伝子が発現していることを指標として、DNA領域の欠失を有しない細胞を除去することができる。ネガティブ選択用マーカー遺伝子が、可視化マーカー遺伝子である場合にはまた、可視化マーカー遺伝子が発現していないことを指標として、DNA領域の欠失を有する細胞を選択することもできる。

[0070] <DNA領域の欠損を有する細胞>

本発明によれば、上述した本発明の方法により、標的DNA領域の欠失を有する細胞を得ることができる。したがって、本発明によれば、標的DNA領域の欠失を有する細胞が提供される。本発明によれば、標的DNA領域が細胞の生存に必須の遺伝子を含んでいることができる。したがって、本発明によれば、標的DNA領域の欠失を有する細胞であって、標的DNA領域が細胞の生存に必須の遺伝子を含む細胞が提供される。本発明によれば、標的DNA領域の欠失を有する細胞であって、標的DNA領域が細胞の生存に必須の遺伝子を含む細胞は、生存に必須の遺伝子が異所的に挿入されたゲノムDNAを有し得る。したがって、本発明によれば、標的DNA領域の欠失を有する細胞であって、標的DNA領域が細胞の生存に必須の遺伝子を含み、当該細胞の生存に必須の遺伝子が異所的に挿入されたゲノムDNAを有する細胞が提供される。異所的な挿入箇所は、特に限定されないが、他の遺伝子が存在しない領域であり得る。上記（a）および（α）において、標的DNA領域の欠失を有する細胞であって、標的DNA領域が細胞の生存に必須の遺伝子を含み、当該細胞の生存に必須の遺伝子が異所的に挿入されたゲノム

DNAを有する細胞を細胞として用いてもよい。これにより、上記(b)または(β)により、さらに欠失するDNA領域を拡大することができ、次の細胞の生存および／または増殖を制御する遺伝子(例えば、細胞の生存および／または増殖を正に制御する遺伝子、例えば、細胞の生存および／または増殖に必須の遺伝子)まで欠失を広げることができる。上述した本発明の方法より得られた標的DNA領域の欠失を有する細胞は、欠失前の細胞と比較して50%以上、60%以上、70%以上、80%以上、90%以上、または95%以上の増殖率または生存率を有し得る。ある態様では、上記標的DNA領域の欠失を有する細胞は、元々のゲノムの20%以上、30%以上、40%以上、50%以上、60%以上、70%以上、80%以上または90%以上を欠失している。

[0071] <DNA領域の欠損を有する細胞の製造方法>

上記標的DNA領域の欠失を有する細胞は、上記の本発明の方法により作製することができる。したがって、本発明によれば、標的DNA領域の欠失を有する細胞を製造する方法であって、本発明の方法を実施することを含む、方法が提供される。DNA領域の欠失を有する細胞は、ネガティブ選択用マーカー遺伝子を有しないことで選択されたものであり得る。

実施例

[0072] 材料と方法

細胞培養

HAP1細胞は、5%CO₂を含む雰囲気下、37℃で、10%(v/v)ウシ胎児血

清と100 U/mL ペニシリン／ストレプトマイシンを補充したイスコフ改変ダルベッコ培地(IMDM)で培養した。HCT116細胞はマッコイ5A培地を使用したことを除き、同じ条件で培養した。

[0073] 過去の大規模実験に基づく必須遺伝子の推定

HAP1細胞の必須遺伝子は、以下の2つの基準で定義された：(1) 3つの報告(OGEデータベース¹、CRISPRスクリーニング²、Gen

e - t r a p スクリーニング³) のうち少なくとも1つで必須であるとアノテーションされていること、(2) H A P 1 細胞において転写されていること [Transcript per million > 0; データは Human Protein Atlas (www.proteinatlas.org/humancell) から入手できる]。

[0074] プラスミド構築

24個のgRNA/Cas9発現プラスミド(gpHY001-pJS067)の構築については、CRISPRdirectソフトウェア⁴を用いてgRNA配列を設計した。gRNAオリゴヌクレオチド(表1)をアニーリングし、標準プロトコルを用いて直線化したpX330(Addgene)にライゲーションした⁵。

[0075] [表1]

表1: gRNA/Cas9発現プラスミドの構築に用いたオリゴヌクレオチド

gRNA	Forward primer	Reverse primer
L1	CACCGAATAATTGTAGTTAAGTC	AAACGACTTAACTACACAATTATTC
L2	CACCGTTAATAGCTTGATATAGCTA	AAACTAGCTATATCAAGCTATTAAAC
L3	CACCGGGGAGAGATTATGCTTTACG	AAACCGTAAAGCATAATCTCTCCCC
L4	CACCGTTCTCAGCAATTGTGCGGTC	AAACGACCGCACAAATTGCTGAGAAC
L5	CACCGCATCTAGCACTCCCGTAAC	AAACGTTACCGGGAGTGTAGATGC
L6	CACCGTTCTCAGCAATTGTGCGGTC	AAACGACCGCACAAATTGCTGAGAAC
L7	CACCGTGGGAAGAGCTCACCGCTC	AAACGAGCGGTGAGCTCTTCCCCAC
L8	CACCGATCCCATTAAGTCTGCTGAA	AAACTTCGAGCTAGTAATGGGGATC
L9	CACCGATTCTATGAACACACGTAG	AAACCTACGTGTGTTTCATAGAAATC
L10	CACCGGCATGGAGTGAATACGACTT	AAACAAGTCGTATTCCTCCATGCC
L11	CACCGCAGAGAAATGACGAACGATT	AAACAATCGTTCGTCAATTTCTGTC
L12	CACCGAGCACATTTGCAACGTCACG	AAACCGTGACGTGCAAAATGTGCTC
R1	CACCGGACAGCAGTGAGCGTAGAGT	AAACACTCTACGCTCACTGCTGTCC
R2	CACCGCACTTGATGCATTAATGCCG	AAACCGGCATTAATGCATCAAGTGC
R3	CACCGGAGGGAGAAACAGGTCCGTT	AAACAACGGACCTGTTCTCCCTCC
R4	CACCGAACAGAAAAGTTTACTGCGA	AAACTCGCAGTAAACTTTTCTGTTC
R5	CACCGTTACTAGAAAATGTGAAACG	AAACCGTTCACATTTTCTAGTAAAC
R6	CACCGAGAACAGTGTCCAGATCGA	AAACTCGATCTGGCACACTGTCTTC
R7	CACCGAAGCCAACTACAACACGGA	AAACTCCGTAGTTGTAGTTGGCTTC
R8	CACCGGCTGTGCTATTAAGATGCGC	AAACGCGCATCTTAATAGCACAGCC
R9	CACCGCCATGGACTGGAATAGCCGG	AAACCCGGCTATTCAGTCCATGGC
gRNA#1	CACCGAGCCCCAGCAAGAGCACAAG	AAACCTTGTGCTCTTGCTGGGGCTC
gRNA#2	CACCGGAGATCGAGTGCCGCATCAC	AAACGTGATGCGGCACTCGATCTCC
gRNA#3	CACCGGGGACTGCAGTCAAGTCGAA	AAACTTCGACTTGAAGTGCAGTCCCC

[0076] ライゲーションされた生成物は、DH5 alpha E. coli 細胞に形質転換され、プラスミドは、製造者のプロトコルを使用してEndoF

ree Plasmid mini Kit (Qiagen) を使用して精製された。

pHY262は、pWZ267ベクター⁶に、IRES-puroカセットと、pJS050で標的とするsgRNA (sg-A) サイトと呼ばれる配列をIRES-puroカセットの5' UTRに挿入することにより構築された。pHY262は、TransforMaxTM EPI300TM大腸菌 (Episentre) に形質転換された。

pHY263は、pGEM (商標) -T Easy Vector (Promega) にIRES-GFP-2A-Puroカセットを挿入して構築した。

pHY271 (チミジンキナーゼネガティブセレクションマーカー) は、表2-1および2-2に示す増幅産物をギブソンアセンブリーを用いて結合して構築した。ギブソンアセンブリー産物をDH5 alpha E. coli細胞に形質転換した。プラスミドは、EndoFree Plasmid Midi Kit (Qiagen) を用いて、製造者のプロトコルを用いて精製した。

[0077]

[表2-1]

表2-1: プラスミド構築に用いたプライマー

Plasmid name	Fragment #	Forward primers (name & sequence)	Reverse primers (name & sequence)	Template
pHY262	1	HY1674 ATATGGCCACACCGCCACCATGGGTACCGTAGCAAGGG	HY1675 GACGACTGAGCGCGGAGTCTCTTGACAGCTGCTCCATGCC	pLenti3-MCS-IRES2-EGFP(Add gene)
	2	JS214 GATGGAGTGGCGGATGACCGGCAATGAGCGGTGCTCTCTC	HY1673 GGTGGCGGTGTGGCCATAT	pCEN3-T Easy Vector (Promega)
	3	JS210 AAAGTCATCATGTATCTTAATGGGTGGCGCTACATGCC	JS211 GTATGGCGGCACTGATCTCCGTATGGCGGTATTTCTGCC	HR110PA-1 (Frankoski)
	4	HY2609 CGCGGCTCAGTGTGTCGAATTCGCGGTGGACGGGACCGCGGACCGCGGCTCCACCGGATCTCGGAGGGAGAGGAAGTCTCTCT	JS218 TAGATACATGTGATGAGTTTGGACAAC	HR110PA-1 (Frankoski)
pHY263	1	HY1598 GAATACATCTTTAGCTTCCTTAGCTCTGAAATCTCG	HY1499 GCATCGCAACTGCGACCGGTGGCGGCGACGACATCCA	pHY262
	2	HY1496 GCCCGGGCGATGGCGCGGATGATGACTATCAGCGTGAG	HY1497 GGAAGCTAAATGAGTATTCAACATTTCCGTGT	pLIM375
	3	HY1569 GTGGGATCCAGACTTCGAATTCATGTCGAAGCTACATATAAGGA	HY1570 TCGCGCGGATGGCGCGGCGCGCGCTGATGGGTATTTCTCCTTA	pN308
	4	HY1567 AAGTCATCAATGTATCTTATGCCCTACACAACTAGAAATGG	HY1568 GACATGAATTCGAAGTCTGGATCCCACTATTTATACCATGGGAGG	pombe genome
	5	HY2620 ATATGGCCACACCGCCACCATGACCGAGTACAGC	HY1419 GCTAAGATACATGATGAGTTTGGACA	HR110PA-1 (Frankoski)
	6	JS214 GATCGAGTGGCGCATCCCGGCAATTGACGGGTGCTCCTC	JS213 GTGGCGGTGTGGCGCATATATCATGCTGTTTTTCAAAGG	pLenti3-MCS-IRES2-EGFP(Add gene)
pHY269	1	HY1571 ATCGCTTTCCCTCCCTATAC	HY1572 CCACATTTATACCATGGGAGCGG	pHY262
		HY1587 TCCCATGTATATAATAGTGGGGCTTATAAATCCAGATCTTAC	HY1578 CTCCCAGAAAGGAGAACAGAG	HAP1_gDNA
	2	HY1579 AATACCTCTGCTGAACCTACCA	HY1580 GATGCTTGTGCTCACTAACCT	HAP1_gDNA
		HY1581 GGTGCTAGTTTGGCTTTCTTGT	HY1582 CCTTTGGATTTCCTCAGTTTCT	HAP1_gDNA
	3	HY1583 TGGCTTGAATATGCTTTGCT	HY1427 ACAGCGCATCTCAATATCACAGCAATTTAAACATATATTT	HAP1_gDNA
		HY1428 GTCAATTTGAGATGCTGCTGTAATCTTTTGTATCTTAAACTTC	HY1429 GTCTAGAGCGAGAAAGTGTGTCACAAACATATCCAGGATAC	HAP1_gDNA
pHY270		HY1430 GAACCTTCTGCTGCTAGCACAGATGATCTCTGAACTTTGGTT	HY1584 AGCCCTAAGTGTGTAACTGGAA	HAP1_gDNA
	4	HY1585 CAACACCGAGTACGTTTGT	HY1589 CTATATGTAGCTTCACATGCGCGGTGAGCAGCTCTAAT	HAP1_gDNA
		HY1573 CATGCGAAAGCTACATATAAGGA	HY1574 GCTTCAACCCCTCAACATATAC	pHY262
	1	HY1571 ATCGCTTTCCCTCCCTATAC	HY1294 TAAGTACTTGGGTTCCTCACTATTTATACCATGGGAGGCTCA	pHY262
		HY1295 ATGGTATAAATAGTGGGAAACCCAAAGTACTTATTTGAATGATACATC	HY1296 GCTGGCGACCGGAAAAAGGCGCGTG	HAP1_gDNA
	2	HY1297 CTTCGCTCTCTGTTTATGACT	HY1298 CCAGTCTCTCTGTTTGTGTTGT	HAP1_gDNA
pHY276	3	HY1299 ATGTAAAGGAAATAAAGCCACA	HY1300 AGAGGGATGGAAAGGAGTGAAG	HAP1_gDNA
	4	HY1301 GAGGCAATGAAGAGGTAGAGAGA	HY1302 TCAGAGAAATGATGAACACAGACAG	HAP1_gDNA
		HY1303 ATCCTGCTGCTTAAAGTTTGT	HY1669 GGACTATCTTAAGCAAGCATCTTTTAGGGAGTTACAGATAATACCC	HAP1_gDNA
	5	HY1669 AGAATGCTTCTTACGATATGTC	HY1670 AGATATTTTCGAGAGCCCTAGC	HAP1_gDNA
		HY1663 GCTAGCGCTCTCGAAAATATCTATGATGTGTGTCACCCCTTTTGTGT	HY1304 CTTCGCCCGCACTCCCTTTCTTAAGAA	HAP1_gDNA
	6	HY1663 GCTAGCGCTCTCGAAAATATCTATGATGTGTGTCACCCCTTTTGTGT	HY1305 GTAGCTTTCGACATTTATGATGACAGACATTTTAAAGCAATTGGGAG	HAP1_gDNA
		HY1306 TTTTAAAAATGTGTGATTAATAATGTGCAAGCTACATATAAGGAA	HY1574 GCTTCAACCCCTCAACATATAC	pHY262

[表2-2]

表2-2 (補足表6): プラスミド構築に用いたプライマー

Plasmid name	Fragment #	Forward primers (name & sequence)		Reverse primers (name & sequence)		Template
pHT271	1	HY2639	TCACCTCCGGTTTCCAGTCC	HY2640	TGCGGGTGATCGCGCCACTCTGATCTCCCTGTGATCGGTATTTCCTCC	pGEM-T Easy Vector (Promega)
	2	HY2641	CGAGTGGCCGCAATCCAGGGCAATTGACGGCTGCTCAATTGCAATATCATATGTGGACGTACGTTAAGAGATA	HY2642	CCCTCTGCAATATACAGAGCCGATAAACAACATCGGTCAATTTTACAGATGATTACTTTACGTTACGTTCCACA	No
	3	HY2643	TCGGGTGTGTATATCGAGGGGGCCAGAGCGGACATCCGGCCACAGTCCC	HY2644	GCTTGGCGTCTTGTTGGGAGGGGTGACGTGGGCCATGGTGGCGATTA	HE110P-A-1 (F-malshish)
		HY2645	GCCAGCTACCCCTTCCACCAAGCACAGCCAGCGCTTCGACCAAGGCG	HY2646	TGCTCTCTGTCTCTCTGGGTCATCAGGGCGGGTCTTCTTGTTGCTGTGTGGCGTCTGCTTCT	No
		HY2647	AGACCCAGGACAGACAGAGAGGCCACCGACAGGTGAGACTGGAGAGAGATCCACGACCTGTGATGATGTATAGATCGAGG	HY2648	TCTCTCTGGCCACAGGCGCACACAGAGCTGGGTGGTGGTGTCTTGATCCATGCGGTGGGGGCCGTGATGATGATCACTTCAG	No
		HY2649	GTGGCGCTGGGGCAGAGAGGACGATCTGCTACGTTGCCCGAGCGCCATGAGACTTCTGGAGTGTCTGGGCGTCCAGCGGAC	HY2650	TCCGCGCGCGCTGATCTTCCGCTGGGTTCAGTCTGTGGTGGGTGGTGTGTGTGGCGATGTGGCGATGTGCTCTGCTGGCGCGCCAGCAC	No
	4	HY2651	CCATGAGATATACCGCGCGGACGCGCGCTGTGTGTATGACCAGCGCCCGACATCA	HY2652	CGATGCTCTGCTGGCGCGCCAGCAC	No
		HY2653	CCATCTGGCGGATACAGATACAGATGTGTACGCTGTGCTGGACAGACACCCCATGGCTCTTCAATGCTGTGCTTACCGCCGCGCAG	HY2654	GGGGGGGATCAAGGGCCACAGAGGCCAGCACGGCTTGGGGGGTCAATGCTTGGCCCATCAGATATCTTGGCGGGGGTGTGATGAG	No
		HY2655	TTGTGGGCGCTGATCTGCTCCCGCCACCTCTCCCGGACCAACATATGTCTGGCGGCGGACAGACAGATGCTGTGATGTGTTGTGTCCTC	HY2656	CTGATGGCGGCGACATGGGCGAGTCTCATGTCCTCTCTCTGCGGGTCTCTGTCTCTTGG	No
		HY2657	GCCATCTGGCGCGGATACAGATAGATGTGTACGGCTGTGCTGGCCATCACCTGTGAGATTAAGCTTGGAGGGGGGCGGACGCTGGTG	HY2658	CTCTGGGGGTTCGCGCGCCCTGGGGGGGCGCACGGCGGTGCTCCCTCAGCTTGGCCCACTGTCCTCCACCAAGCTTGGCGCGCCCTG	No
		HY2659	CAGCGCGCGGACGCGCCAGAGCGCCGGGCGCCAGACCCCGGACCATCTGGCGACATCCCTGTTCATACCTGTGTAGAGGCCCGGAG	HY2660	GCCACATGCTCCAGGGGCCAGCGGAGCAATGTTGATCAAGTGTCCGCTGTGGGGGCGCACAGAGCTGGGGGGGCTGTGACAGAGGT	No
		HY2661	TGGGGCGCTGGACGTGGCGAAGAGACTGAGACCTATGGACCTCATGGTGTTCATCTGCACTACAGACAGAGCCCGCGCGGCGCTG	HY2662	CCCGGGGGTGTGTCACGTGGGGTCTTCCACCATCCGCGCTGGTACGCTAGCAGCGGGCGTCTGTCCAGCGCGCGGGGGTCTTGATCAGAGGT	No
	5	HY2663	GCCAGCTTACCTCTGCCACACAGACAGCCGACCGCCATCGACCAAGGC	HY2664	GTTTGGCGTCCGCCATCTCTCTTGGCGAGAGTCTTGGCCAGGTGTCCAGATGTGGGGGATATGCTGGCGGGGGTGGTCACTGGG	No
		HY2665	GAGATGATGGCGGACGACCAACGAGGGCAAGGAAAGTCTTCT	HY2666	TGATGTCAAATGACGCACTGAGACCTCTGTTCTGTGAGGAC	HE110P-A-1 (F-malshish)
	6	HY2667	CATGCGTCACTTTTGTACTCACGCGGTCTGTTATAGTTTCAAAATCATGTCGACTTACCGCAT	HY2668	GGAAAGCGGGCAGTGAAGCTCCCGCTGTGAGGCGCTGCTTGTCAATGCGGTAAAGTGTCACTGAT	No

[0078] pHY269 (RBMX2 遺伝子座クローニング) および pHY270 (MMGT1 遺伝子座クローニング) を構築した。pHY262 (~100 ng) を BamHI および EcoRI (New England Biolabs) により直鎖化し、転写開始部位の上流の 3 kb 領域およびポリアデニル化部位の下流の 3 kb 領域を含む RBMX2 または MMGT1 遺伝子座をカバーする ~100 ng のゲノムの増幅産物 (補足表 6) とともに酵母に共形質転換した。各フラグメントは、~300 bp のオーバーラップを有する。酵母の形質転換は、先に記載されたプロトコルに従って行った⁷。プラスミドを酵母コロニーから抽出し、前記のプロトコルに従って、Transformax™ EPI300™ 大腸菌 (Epicentre) に回収した⁸。大腸菌を CopyControl™ Induction Solution (Epicentre) を用いて LB で培養し、製造者のプロトコルを用いて EndoFree Plasmid Midi Kit (Qiagen) を用いてプラスミドを精製した。構築したプラスミドを確認するために、精製したプラスミドを適切な制限酵素で消化し、アガロースゲル電気泳動で分離した。

[0079] トランスフェクション

HAP1 細胞を 12 ウェルプレートに 6×10^5 細胞/ウェルの密度で播種し、一晩インキュベートした。1. 5 μ g の DNA を含む 2. 5 M CaCl₂ を 50 μ L 調製し、2×BBS バッファー (400 mM ホウ酸, 300 mM NaCl, 5 mM EDTA) 50 μ L と混合した。この溶液を 25℃ で 5 分間インキュベートした後、IMDM 1 mL と混合し、細胞培養物に添加して 4~8 時間インキュベートした。その後、培地を吸引し、細胞を D-PBS (ナカライテスク、#14249-24) で 2 回洗浄した。その後、新鮮な IMDM 培地を加え、37℃、5% CO₂ でインキュベートした。HCT116 細胞を 12 ウェルプレートに 3×10^5 細胞/ウェルの密度で播種し、一晩インキュベートした。FuGENE (商標) HD (Promega) トランスフェクションは、製造者の指示書に従い、3:1 の割

合で行った（1ウェルあたり、FuGENE（商標）HD 3 μ L Transfection Reagent：900 ng DNA）。

[0080] デジタルジャンクションPCR

各標的領域について、48回の反応をネストPCR形式で行った。1回目のPCRでは1反応あたり100ゲノム（半数体ヒト細胞の場合は3.3 pgのゲノムDNA）を使用し、2回目のPCRでは2×Quick Taq HS Dye Mix（東洋紡）と表3-1および3-2に記載のプライマーを用いて、1回目のPCR混合物の1/100希釈液を鋳型として使用した。欠失効率（ λ ）は、ポアソン統計学を用いて計算した： $\lambda = -\ln(1-p)$ {ここで、 λ は100ゲノム中の欠失接合部を持つゲノムの平均数であり、 p は48回のPCRにおける陽性反応の割合である} ⁹。

[0081] [表3-1]

表3-1（補足表5）：ジャンクションデジタルPCRに用いたプライマー

Fig. #	Junctions	1st STEP forward	1st STEP reverse
図2	L1-R1	TGGTAAGGATGAGGGAAGATG	GCCAAGTAGCAAAGAGCGTGT
	L2-R1	TGGCATAATTGTAAATGTCTGCT	GCCAAGTAGCAAAGAGCGTGT
	L3-R1	AAAGAGATCATGGGGTAGCAAA	GCCAAGTAGCAAAGAGCGTGT
	L4-R1	CAGACTTGGATGTGGACCTTG	GCCAAGTAGCAAAGAGCGTGT
	L5-R1	CCCATCCTGTGCTTCTCTCT	GCCAAGTAGCAAAGAGCGTGT
	L4-R2	CAGACTTGGATGTGGACCTTG	TCCTGAACCACACCTTGGAC
	L4-R3	CAGACTTGGATGTGGACCTTG	TTGTAACCTGTGTTTTCATGTTTT
	L4-R4	CAGACTTGGATGTGGACCTTG	GCTTGCTAGGATTTTGGTTTTT
図4	L4-R5	CAGACTTGGATGTGGACCTTG	CACTTCTGGGTCACCTTAGGAA
	L5-R1	CCCATCCTGTGCTTCTCTCT	GCCAAGTAGCAAAGAGCGTGT
図5	L4-R6	CAGACTTGGATGTGGACCTTG	CCAAGGTACCACTTAATGCTTCA
	L5-R1	CCCATCCTGTGCTTCTCTCT	GCCAAGTAGCAAAGAGCGTGT
	L6-R1	CTCCCTGGCTTTTGTGTGG	GCCAAGTAGCAAAGAGCGTGT
	L7-R1	CACCGACCAGACCCCTGGGATCGTCC	GCCAAGTAGCAAAGAGCGTGT
	L4-R6	CAGACTTGGATGTGGACCTTG	CCAAGGTACCACTTAATGCTTCA
	L4-R7	CAGACTTGGATGTGGACCTTG	CTGCTGGACGTTTGGTAATGC
図6	L4-R8	CAGACTTGGATGTGGACCTTG	GGAAGAAGGCAAAATCAAAAGGTC
	L8-R9	TTTGTGGTCTGGCATATTGGG	ATGCCTCCTTGACTTCCATCAG
	L9-R9	TAATTGAAGAAGCCCTCGACCC	ATGCCTCCTTGACTTCCATCAG
	L10-R9	ATGATTGTCCAGCTTTTCCCC	ATGCCTCCTTGACTTCCATCAG
	L11-R9	AGGCAGTTAAGGAAAAGTACCCA	ATGCCTCCTTGACTTCCATCAG
	L12-R9	AAATAAGGCTGGGCATGGTTG	ATGCCTCCTTGACTTCCATCAG

[表3-2]

表3-2 (補足表5): ジャンクションデジタルPCRに用いたプライマー

Fig. #	Junctions	2nd STEP forward	2nd STEP reverse
図2	L1-R1	ATATAAGAAGGACGAGGAGAAATG	GGCCAATATGTGCTTTAGGAA
	L2-R1	ATAAGACCAACTTCAGACACCAG	GGCCAATATGTGCTTTAGGAA
	L3-R1	GTAGCCAAGAGGGCTGAGAGT	GGCCAATATGTGCTTTAGGAA
	L4-R1	GATCATAGCAAAAGCTCAGTTGTC	GGCCAATATGTGCTTTAGGAA
	L5-R1	GCACTCATGTGCTCTGGTCTT	GGCCAATATGTGCTTTAGGAA
	L4-R2	GATCATAGCAAAAGCTCAGTTGTC	ACGAGAGGCAGCTCTATTTCATC
	L4-R3	GATCATAGCAAAAGCTCAGTTGTC	TGGCACAAAGGAGAGAATAACA
	L4-R4	GATCATAGCAAAAGCTCAGTTGTC	CCACAGATTCAATAAAAGGAGGAG
	L4-R5	GATCATAGCAAAAGCTCAGTTGTC	CTAGCTGTACCTTGCACCTTTG
図4	L5-R1	GCACTCATGTGCTCTGGTCTT	GGCCAATATGTGCTTTAGGAA
	L4-R6	GATCATAGCAAAAGCTCAGTTGTC	TACTCGGCAGGTAGTAACTTCT
図5	L5-R1	GCACTCATGTGCTCTGGTCTT	GGCCAATATGTGCTTTAGGAA
	L6-R1	CTCCCTGGCTTTTGTGTTGG	GGCCAATATGTGCTTTAGGAA
	L7-R1	CTGTGGAGATTAGTCAGCCCAA	GGCCAATATGTGCTTTAGGAA
	L4-R6	CTCAGCCATATCTATTCTCCAC	TACTCGGCAGGTAGTAACTTCT
	L4-R7	GATCATAGCAAAAGCTCAGTTGTC	CCAGCTATCCACGCCAATTAG
	L4-R8	GATCATAGCAAAAGCTCAGTTGTC	GCTTGACTCTTCTAATTCTCACTGC
図6	L8-R9	AAAAGCAAAGCAGAAATCCGGA	TAAAGTGACGAGCAAGCAAGGA
	L9-R9	ACCCCTTTTACTATCCACAGGC	TAAAGTGACGAGCAAGCAAGGA
	L10-R9	ACAGGGAAGTTGCAAAGACAAA	TAAAGTGACGAGCAAGCAAGGA
	L11-R9	TCTGATCTATTCTGGGCTGGT	TAAAGTGACGAGCAAGCAAGGA
	L12-R9	GCAAACCAAAACCACGAGATAC	TAAAGTGACGAGCAAGCAAGGA

[0082] 6-TG選択後の生存コロニーの単離

gRNA/Cas9発現ベクターをHAP1細胞にトランスフェクションした後、～2,000個の細胞をトリプシン処理して10cmディッシュに再播種し、5 μ M 6-チオグアニン (6-TG) (Sigma、A4882-100MG) で12日間培養した。その後、10cmディッシュ上の独立したコロニーをピックアップし、24ウェルプレートの各ウェルに別々に展開した。次いで、細胞の一部をゲノムDNAの単離のために12ウェルプレート中で展開し、残りを個々のクローンのさらなる分析のために凍結ストックとした。

[0083] 6-TG選択後のコロニー形成アッセイ

gRNA/Cas9発現ベクターをHAP1またはHCT116にトランスフェクションした後、～ 5×10^4 個の細胞をトリプシン処理して6cmディッシュに再播種し、5 μ M 6-TG (Sigma、A4882-100

MG) で9日間培養した。その後、6 cmディッシュから培地を除去し、E t B r 溶液 (10 mg/mL エチジウムブロマイド (ナカライテスク) 50 %エタノール中0. 5 %) を加え、室温で30秒間インキュベートした。E t B r 溶液を除去し、コロニーをUVイルミネーターで可視化した¹⁰。染色されたコロニーの数と大きさは、それぞれImage Jソフトウェアの「find maxima」と「analysis particle」関数を使用して定量した。

[0084] ゲノムDNA抽出

12ウェルプレートで培養した細胞を1. 5 mLのチューブに採取し、118 μ Lのブレーキングバッファー (10 mM Tris-HCl (pH 8. 0)、100 mM NaCl、0. 04 % (w/v) SDS、0. 2 mg/mL プロテイナーゼK (和光)、2. 5 mg/mL RNase A (ニッポンジーン)) に再懸濁し、37℃で~18時間インキュベートした。フェノール/クロロホルム/イソアミルアルコール (25 : 24 : 1) の1体積を添加し、60 rpmで30分間回転させた。12, 000 rpmで10分間遠心分離した後、上層水層を新鮮なチューブに移し、絶対イソプロパノール1体積を添加した。13, 000 rpmで5分間遠心分離した後、ペレットを300 μ Lの70 %エタノールで洗浄し、風乾した。ゲノムDNAを100 μ LのTE緩衝液に溶解し、-20℃で保存した。

[0085] 欠失ジャンクション部の検出と配列決定

2X Quick Taq HS DyeMix (東洋紡) と補足表5に記載のPCRプライマーを用いて、抽出したゲノムDNAから、ネストPCR形式で欠失を有するジャンクション部 (欠失ジャンクション部; deletion junction) を増幅した。ABI PRISM BigDye Terminator Cycle Sequencing kit (Applied Biosystems社) を用いて、ABI 3100 DNAシーケンサーを用いて全PCR産物の塩基配列を決定した (図1d)。

[0086] 欠失クローン中の遺伝子欠失の検出

遺伝子欠失の検出は、抽出したゲノムDNAを2x Quick Taq HS Dye Mix（東洋紡）と表4-1および4-2に記載のPCRプライマーを用いて増幅し、欠失すべき遺伝子がゲノム上から欠失したことを確認するために、ネストPCR形式で行った。得られたPCR産物を1%アガロースゲル電気泳動により解析した（図3a）。

[0087]

[表4-1]

表4-1 (補足表4) : 遺伝子欠失の検出に用いたプライマー

Gene regions	1st STEP forward	1st STEP reverse
<i>GPR119</i>	GCCCTTACCGTCTTAGCCATCA	CGACATGCTCAAGATTGCCCTCC
<i>RBMX2</i>	CACGTGAAGCATCGCCTGTCTTG	ATGTCGAGTGGCTGTACAAGGT
<i>DENND10P1</i>	GAAAGGTCCCAAGTCAAAGGAA	GGGTCAGAAACACACCAAG
<i>ENOX2, LOC105373338</i>	AAGCCAGTTTTCAGCACGAATG	AACAGAGAGGATGGAGTTTGG
<i>LINC01201, ARHGAP36</i>	TCCTCTAGAACGGACCAAGTAT	CGAAGGGAGGCAGAGACACAC
<i>IGSF1</i>	ACATTACACCCAGGAAGAGCGG	AGGTGCCCTTACTGAGTCCAAT
<i>OR13H1</i>	TCAAACGAGTGTCTCCTTGGCT	TCAGGGTAGGACTTGGACTGAG
<i>FIRRE</i>	CACGTGTGGCACCGTTATAGTTT	TGACCACGCACAAACAGATGAG
<i>MST4</i>	GGGGTGTGGTATGGGACTTCAA	CCCCTTTCCTCCATCCTAGCAC
<i>FRMD7</i>	CGCTATTCTCCCTCCTGTTACACA	TGTTTACGTGTGGCAGTGTGT
<i>RAP2C, RAP2C-AS1, MBNL3</i>	TTACCCCTGACTGACTTTTGCT	TGGGACTGACATTTAAACGCCCT
<i>HS6ST2, HS6ST2-AS1</i>	GTGAGTCTGGTTTGGCTTTCGG	TCAGACCCATTTCAGAGCCAG
<i>USP26</i>	AACGCCATCTCCTCCTGCATCTG	ACTCAGCAGTGTGACGGTATGA
<i>TFDP3</i>	CCCAAAAGTCATGCCCATCCAC	CCTGTGCCGTCTTTCATGAAG
<i>GPC4</i>	TTGCCCGAGTGTGACAGCTAT	TCTGGAGAGACGTGGAGGTGAT
<i>GPC3, GPC3-AS1</i>	CAGGGAGGTGGGTTTCAAGGAA	GTGTGCTTCTTCTCTCTGGTGC
<i>microRNAs</i>	ATCCTCCTTTCCTCCACAGGCC	ACTGCCCTAAATGCCCTTCTTG
<i>CCDC160</i>	CAGATAGCAGCAAGGGAATGGA	TGACTTCTTTGGCTTGTCTGGAG
<i>PHF6</i>	GGGTTACCGCTTGTCTAAGGACT	TGGTACTGAGGTGCTATGGTAGT
<i>HPRT1</i>	CCTTTGGGCGGATTGTGTTTAA	TTTGTAGCTCCGCCAACCCATT
<i>microRNAs</i>	CCTTATGTCAGGGGTTTCATGCT	TGGGGAAAAGAGGCTATCAGGA
<i>MIR503HG, microRNAs</i>	TGGAGGAAAATCTAGGCACACTG	GAAATCCCCATTCTGCTCCCGC
<i>LINC00629</i>	TTTCATGTTCCAGGTGGCAAGG	CATCCAAGGACAAGAGGCAGGG
<i>PLAC1</i>	AGGGACCTGGGTATGCTCTTCT	GATCCTCCTACCTCTGCGTTT
<i>FAM122B, FAM122C</i>	GGCAACCACAAACCCCAATCT	GGTGGCTGGGAGTAGAAGTAGC
<i>MOSPD1</i>	GCAGGGAGGGGCTTGAAGA	CACITTTGATGCTAGGGGTCATGT
<i>LINC02243</i>	CCACTGGCATTTCTACCAACCCA	CAGCACAGCAGTCCATGTCAC
<i>SMIM10</i>	TGATTCTCAACGTCCACCTGCA	GGGAGATTGCGCTGCTGTCTGA
<i>RLT8B</i>	CCTGCTGGCTGACTTCTCACAT	CTGTTACCTTCAGCGAGACCCA
<i>RLT8C</i>	CTTCTGGCCGAGATGAAGCGA	TCTGTCCGTGGGATGCGATG
<i>RLT8A</i>	AGTCTGGAGGAGTGGAGGTGG	TCTCCAACGACGCCCTGAAG
<i>SMIM0L2B</i>	AAATGGGGATAGCAGGTGAGGG	GGGAATCTGGGGATCTCGCTTT
<i>ETDB, SMIM10L2B-AS1</i>	CCCAAGTGACAGAGAGGTGAGA	GCCAGGGTAGTTGAGGGTTCTG
<i>CT55</i>	AGGGGTCTCTGCTTCTCTTTG	TTCTCTGTGTGTCCCTTGTCT
<i>ZNF75D, ETDA, ETDC, ZNF449</i>	CAGCTCCTGTCCCGTTTCG	TGACGCATTGAATCTCACAGGA
<i>LOC100506790</i>	TGTTTGGTAATCAGGGGAGGGT	AGCTGCCTTCTCTCACCTCTG
<i>SMIM10L2A</i>	ATCTGCTCCTGTCCCTGCTTTT	GTCCAGGTCAAGTCGGTCTCA
<i>INTS6L, ISTN6L-AS1</i>	TCCCTCCCCTAGTCTGTCTATC	GTTCATAGAGGGCGACGTGTCT
<i>SAGEP2</i>	GCGAGACTAATTGAGCCCTCCT	CCACTCCATTCCCTACCTTTGG
<i>CT45A1, 2, 3, 5, 6, 7, 8, 9, 10</i>	CCTGAGTCCCACCATAGGCTTA	CGCCACAACITCACTGCCATTT
<i>SAGE1</i>	GTCTTTGGTGGAAACTGTGCCA	CACCTACCTTGTGCTTCTCTG
<i>MMGT1</i>	CCCACCTCCCCTTCTAAGAACACT	TGACCTTGCCCTGTCCAATAAAT
<i>SLC9A6</i>	TTTGTGTGCTAGGGGAGGGAG	AACTTCTGACGACAAAGCCTGC

[表4-2]

表4-2 (補足表4): 遺伝子欠失の検出に用いたプライマー

Gene regions	2nd STEP forward	2nd STEP reverse
<i>GPR119</i>	CTGCTGATCTAGTTGGGGCTCC	GCTGTATTTCACCCCTCACTTCGT
<i>RBMX2</i>	TCACTGAAGCATCGCCTGTCTT	GGAGAACAACCTTGGCAGCAG
<i>DENND10P1</i>	CCAGCCCTACATGTGTTCAGGA	CCGCCACAACCCCATATGTAGA
<i>ENOX2, LOC105373338</i>	CCAAGAGCCCGAGACTGATGAA	TTCCTCTTGGCTTCTTACAG
<i>LINC01201, ARHGAP36</i>	TCTTGCTCACATTCTTGGTCCT	ACCCAGCTAAGAGTGTTCGGT
<i>IGSF1</i>	TGGATATGAGTGCAGCAGATGT	GGATCTGGGTCCGGCTAATTT
<i>OR13H1</i>	AATTIACCTCAGCACATTGTGGG	CAACATAAATTTCCTTATTGCCCGT
<i>FIRRE</i>	TGCACAAATCTTCAACACCACC	GTATTCGTCTCTCTCCCGCC
<i>MST4</i>	TCTAGGTGAAGCATACTCCAGTGT	AGGAGGCAATTCTATAGGCAAGCT
<i>FRMD7</i>	AAGCCCTCATGTGTCACGATGT	ACCTCGTACCAGTGTCTCATCC
<i>RAP2C, RAP2C-AS1, MBNL3</i>	ACGAAACTCAGTTGAACCGAAGC	AGTGGGGAACAGAAGTATATGTGCT
<i>HS6ST2, HS6ST2-AS1</i>	TTCTACAGCTCGGTAAGATTCACC	ACAACCTCTCTGTTCATGCCTGA
<i>USP26</i>	AGCATATGGGATTGAGGTGTCTGT	TCAAACAGTGCCCGAAATCCA
<i>TFDP3</i>	TGCAATGTGTGTCAGTCACT	GCCTCTCAGAACCAGCATTCCT
<i>GPC4</i>	TCTGGACACCTAGTCTCCCT	AAGCAAAACAGGGTCCCTCTT
<i>GPC3, GPC3-AS1</i>	AGTTGTAAAGAGAGACCAGACCAGG	ACTGGCCTATGATCTGGATGTGG
<i>microRNAs</i>	GCTGATCAGGAATGCGCCAAC	AGGCCTTGGCCATGTAAAGTG
<i>CCDC160</i>	CAGATATTGGAAGAGGTGCCGGA	GAGCACATTTCCAACTCCGGT
<i>PHF6</i>	CTTGACAGCACACCTCTCTCT	ACTTGACAAACGTGATCTTAAGGGA
<i>HPRT1</i>	ACTGGAAAAGCAAAATACAAAGCCT	GGAAACAGTGTAAAGCTAGATGGC
<i>microRNAs</i>	CAATGCGTCTTGAGGCCCTG	TCTGAGTTGTGGTATAAAGGTGACC
<i>MIR503HG, microRNAs</i>	AGTTTAAAGGCCACGCTGTCT	GGCCCTAGAAAGTTTCCAGTT
<i>LINC00629</i>	ACTCGATTCTCTCCCTACACTGG	TTTCAAAGTGGGTGAGGAGGT
<i>PLAC1</i>	GGACCAATCATATCATCTGTGTGA	GTCTGTGTTTGGGTTCATCTTGAGG
<i>FAM122B, FAM122C</i>	AGTTGCCATGCTGTCTCTTC	CCCAAAATTAAGTACTACAAGTGGCG
<i>MOSPD1</i>	TGACTTAGGCTCCTCACATGCT	AGGATGCTTCTGAGTGACACT
<i>LINC02243</i>	CCCCAGTCCATTATCTGCCCT	TGCATCGTCACAGGCAGTAAGA
<i>SMIM10</i>	CATCCCTCTGTCCCAAGGTC	CAACTGTTCAAAACCATGCCACA
<i>RLT8B</i>	CAGCTGTTACCAATAATCTGTGTG	AGACGACTACATCATGCTGCC
<i>RLT8C</i>	TGCAGCTGATAAAGGCCCTCC	CCAGCCAGATACCTCCGAG
<i>RLT8A</i>	CTCTGGGGATGGGAGGAAATG	TTCCTTTCCCGAGACGTTTGA
<i>SMIM0L2B</i>	CAAGGTACAGGGTCTCAGGAG	TGACAGAGGAAACCGAGGCTTC
<i>ETDB, SMIM10L2B-AS1</i>	TGGGAACGTAATCACAGTGCT	CATGGCGGACAAAGTTGCTGAT
<i>CT55</i>	TCACCAACATGATGAGTAGCACCT	GGTCCATGCCACAGTGTCT
<i>ZNF75D, ETDA, ETDC, ZNF449</i>	GTTTCAAGTTCCAGCCGTGAG	TGAGCCTTGGAGTAACTGGAGC
<i>LOC100506790</i>	ACCCTATTGAGTCTGAACCCG	TCACACATGGGCTGTATAGGA
<i>SMIM10L2A</i>	CAGATGTGGCCCTTCAAACCC	AGAACAGCAGTACCTCTCGG
<i>INTS6L, INTS6L-AS1</i>	CATTCGAACTGCTTCCAAGTCCT	GGGGTCGCGTAGGAAAGGAG
<i>SAGEP2</i>	GTATGTTTGTGCCATCTCCCT	TAACGCGGGTCCCATCCATTTC
<i>CT45A1, 2, 3, 5, 6, 7, 8, 9, 10</i>	GCACATTCCCACTCCCAAG	AGACAGAGTTCCCAAGGTCACG
<i>SAGE1</i>	ACAGGACATTGTATGCCACCCA	GGCCAGGATTAGGAGGAACAA
<i>MMGT1</i>	TGCATACATCATCTTCTTGCCCC	AGCATCTTCTGGTGTATGTGG
<i>SLC9A6</i>	TGGTTCAGCTCTGTAGACTGGA	CTGACTTTGGTGAGTGTGGAA

[0088] 成長速度の解析

各HAP1欠失クローンまたはWTクローンをhHY224 (IRES-GFP-2A-PuroカセットをHAP1細胞のGAPDH遺伝子座の3' UTRに組み込んだ)と混合し、IMDM培地で培養した。2日間培養した後、細胞を「day_0」サンプルとして回収し、残りの細胞をさらに1

6日間培養して「day__16」サンプルとして回収した。GFP陽性細胞と陰性細胞をフローサイトメトリーで測定した。欠失クローン中のGFP陰性細胞の比率をWTクローン中のGFP陰性細胞の比率で割って相対集団を算出した。フローサイトメトリー解析は、EC800フローサイトメトリーアナライザー（ソニーバイオテクノロジー社製）を用いて行った。

[0089] マイクロアレイ解析

RNeasy Mini RNA isolation kit (Qiagen) を用いてHAP1細胞からTotal RNAを抽出した。マイクロアレイ解析は、SurePrint G3 Human GE 8x60 K Microarray (アジレント) を用いて行った。データ解析は、Agilent Feature Extraction Software (Agilent) を用いて行った。データのプロットには、すべての野生型クローンまたは欠失クローンにおいてシグナル評価スコア=2を有する遺伝子のみを選択した。P値<0.05、2倍変化>1を有意とみなした。

[0090] 必須ゲノム遺伝子座またはIRES-GFP-2A-PuroカセットのGAPDH遺伝子座への移植

pHY269 (RBMX2遺伝子座クローニング)、pHY270 (MMGT1遺伝子座クローニング)、またはpHY263 (hHY224細胞を作るためのIRES-GFP-2A-Puroカセット付きプラスミド)を、pJS039 (GAPDH領域の3' UTRを標的とするgRNA発現ベクター) およびpJS050 (pHY262のsgRNA (sg-A) 部位を標的とするgRNA発現ベクター) を用いて、HAP1細胞に共トランスフェクションした。トランスフェクション後2日目に、細胞を1 ng/mlのピューロマイシンを添加したIMDM培地で~10日間培養した。HAP1細胞の一部の集団は二倍体となるため、SH800Z (ソニーバイオ) を用いてハプロイド集団のみを採取した。ハプロイド細胞は、長期保存のために細胞を凍結保存する前に、拡大培養して7日間培養した。細胞株を解凍し、通常の培地で5日間培養した後、MEGESを適用した。

[0091] TKマーカーのHCT116ゲノムへの統合

pHY271をpJS067（OCRL領域の下流を標的としたgRNA発現ベクター）およびpJS050（pHY271のsg-A部位を標的としたgRNA発現ベクター）を用いてHCT116に共トランスフェクションした。2日間のトランスフェクション後、細胞をマッコイ5A培地（10% FBS、1% ペニシリン／ストレプトマイシン、1 ng/μL ピューロマイシン）中で～10日間培養した。単一コロニーを拡大し、細胞を長期保存のために凍結保存する前に、別の7日間培養した。マーカーの統合を、HY2223およびHY2224プライマーを用いたPCRによって試験した（補足表1）。細胞ストックを解凍し、MEGESを適用する前に、通常の培地で7日間培養した。

[0092] DNA含有量の分析

2×10^6 細胞を4% PFA/PBSの1 mLに懸濁し、4℃で10分間インキュベートした。その後、細胞を500gで3分間スピンドウンさせた。細胞ペレットを70% EtOHに再懸濁し、-20℃で保存した。PI染色は、細胞を500gで3分間スピンドウンし、500 μLのPI溶液（D-PBS（ナカライテスク）、50 μg/mL ヨウ化プロピジウム（和光）、0.25 mg/mL RNase A（ニッポンジーン））に再懸濁した後、37℃で10分間インキュベートした。フローサイトメトリー解析はEC800フローサイトメトリーアナライザー（ソニーバイオ）で行った。

[0093] 定量的RT-qPCR

RNeasy Mini RNA isolation kit（Qiagen）を用いてHAP1細胞からTotal RNAを抽出した。400 ngのRNAをPrimeScript RT Reagent Kit with gDNA eraser（タカラ）を用いて逆転写した。定量PCRはTB Green（商標） Premix Ex Taq™（タカラ）を用いて行った。各データポイントは、重複して行った3つの独立した

実験の平均値である。遺伝子発現レベルのばらつきを正規化するために、ACTB mRNAを内部コントロールとして使用した。プライマー配列を表5に示す。

[0094] [表5]

表5：定量的RT-PCRおよび遺伝子型決定のために用いたプライマー

Primer Name	Sequence (5' to 3')	Description
HK1338	GCGAACCCGAGGAGATGA	Forward primer for qPCR amplifying RBMX2 in Fig. 4b
HK1339	AGATCCAGGCGCTGTCCTT	Reverse primer for qPCR amplifying RBMX2 in Fig. 4b
HK1375	CCTTCGGATACAGCAAATTCCTT	Forward primer for qPCR amplifying MGMT1 in Fig. 4b
HK1376	GCAGGGAGGAGTGTATTATACC	Reverse primer for qPCR amplifying MGMT1 in Fig. 4b
HY2671	CACTGTGTTGGCGTACAGGCTCTT	Forward primer for qPCR amplifying ACTB in Fig. 4b
HY2672	CTGAGGCACCTCTCCAGCCCTTC	Reverse primer for qPCR amplifying ACTB in Fig. 4b
HY2223	ACATTGCATTAGTCTCCCTTCAC	Forward primer for genotyping in Fig. 6a
HY2224	AATTGAGTTGTTTGTGCCTGAA	Reverse primer for genotyping in Fig. 6a

[0095] [結果]

ゲノムの大規模欠失のための方法論の開発

この実験プラットフォームの概念を実証するために、ヒト細胞株HAP1のヒポキサンチンホスホリボシルトランスフェラーゼ1 (HPRT1) 遺伝子を含む染色体X上のゲノム領域を欠失させることを試みた (図1b)。HPRT1は、Online Gene Essentiality (OGEE) のデータベース [REF] に基づいて特定のヒト細胞において必須遺伝子として同定されたが、6-TG (Nature Reviews Genetics, 6, 507-512 (2005)) とのネガティブな選択マーカーとして確立されている。このような場合、HAP1細胞も同様であるとすれば、HPRT1をネガティブセレクションマーカーとして使用することはできない。Blomen's gene trapping³やMair's CRISPR screening²によるHAP1細胞の機能低下スクリーニングでは、HPRT1はHAP1細胞の増殖には不要であることが示されていたが、我々が開発したペアのgRNA (例えば、図1aに示されるgRNA LおよびgRNA R) による欠失は、遺伝子ノックアウトとは異なる方法であり、異なる結果をもたらす可能性がある。そこで、HPRT1を含む0.51Mbの領域が必須であるかどうかを調べるために

、ペアの gRNA を用いたアプローチを行った。HAP1 細胞にペア型 gRNA (R1 と L1) を導入した 2 日後に、細胞の一部を採取し、ジャンクション PCR を行った。テンプレートとして ~100 のゲノムを含む 48 の PCR 反応のうち、14 の反応は増幅を生じさせ、トランスフェクトされた HAP1 細胞の ~0.3% が、任意の増殖効果が生じた後に R1 と L1 切断部位の間の標的欠失を有していたことを示した。gRNA R1 のみによるネガティブコントロールでは、48 の反応のいずれにおいても増幅は起こらなかったが、これは 0.51 Mb の領域が無傷であったため、ジャンクション PCR で増幅するには長すぎたためと説明される。このようにして、後続の 6-TG 選択は、生存細胞コロニーの数に有意な差を与えることが予想された (図 1 c)。上記工程によって HPR T1 を含む 11 の遺伝子とその中の 0.51 Mb 内の非遺伝的領域を除去することができた。また、ペアの gRNA 切断および 6-TG 選択後に生存した 3 クローンから抽出したゲノム DNA の塩基配列を解析したところ、標的領域が明らかに遺伝子座から欠失していることが確認された (図 1 d)。HPR T1-6-TG ネガティブ選択を用いた我々のプラットフォームが、我々が標的とする非必須領域を欠失した HAP1 細胞クローンを得るのに有用であることが検証された。

[0096] ゲノム必須領域検出のための欠失遺伝子座の拡大

次に、HPR T1 から欠失された領域をセントロメア方向とテロメア方向の両方に拡大し、我々のプラットフォームが本質的な遺伝子領域および／または遺伝子間領域の同定に使用できるかどうかを検討した (図 2)。OGEE データベースに基づいて、セントロメア方向の HPR T1 からの最初の必須遺伝子候補は ARHGAP36 であり、これらの間には 22 の潜在的に非必須遺伝子領域とその間に介在するノンコーディング領域 (図 2 b の gRNA L1 サイトと L2 サイトの間の 3.12 Mb の領域に相当) が存在する。次の必須遺伝子候補は RBMX2 である。ARHGAP36 と RBMX2 の間 (図 2 b の gRNA L3 サイトと L4 サイトの間の 0.62 Mb 領域に対応) には、3 つの非必須遺伝子候補が存在し、大部分は遺伝子間領域で

ある。ここでは、gRNA L1とL5の間の約4 Mbの領域を4つのセクション（2つの必須遺伝子候補領域と2つの非必須遺伝子候補クラスター）に分離し、我々のプラットフォームを適用して、必須遺伝子候補および／または広大な遺伝子間領域のいずれかがHAP1の成長に必須であるかどうかを検討した（図2 b）。

[0097] 2日目のデジタルジャンクションPCRでは、L1、L2、L3、L4、L5のそれぞれとR1の5つの異なるgRNAペアによって、標的領域の欠失が同等の効率（0.3-1.1%）で達成されることが確認された。注目すべきは、6-TG選択後、5つのgRNA対のうち4つのgRNA対で数百個のコロニーが形成されたが、L4/L5の間の領域を欠失させると、生存細胞が極端に失われたことである。このデータは、L4とL5のgRNA標的部位の間に存在するRBMX2遺伝子がHAP1の正常な増殖に重要であり、その間にある領域は欠失していることを明らかにする。この結果はまた、Blomenの遺伝子トラップ³やMaierのCRISPRスクリーニング²による機能低下スクリーニングと一致しており、RBMX2遺伝子座には必須要素が含まれているが、ARHGAP36はHAP1細胞には存在しないことが示唆されている。しかし、これらの従来のLOFスクリーニングでは、1つの遺伝子の本質性のみを一度に評価できるのに対し、我々のプラットフォームは、RBMX2を含むMbスケールのゲノム領域全体、他の38の遺伝子、そして大規模な遺伝子間領域の分散性に関する新しい情報を提供している。このようなスケーラビリティの高さから、我々のプラットフォームはMEGESと命名された。

[0098] テロメア方向への欠失の拡大を継続して行った（図2 c）。最初のOGEEにより支持される必須遺伝子候補はSMIM10であり、これまでHAP1細胞では必須性が未検証であった。その次のOGEE候補であるMMGT1はMaierのCRISPRスクリーニングでHAP1細胞では必須であることが示唆されたが、Blomenの遺伝子トラップでは非必須であることが示唆された。HPRT1とMMGT1の間の1.1 Mbの領域（gRNA

R 1 と R 4 サイトの間) にも 3 0 の非必須遺伝子候補とその遺伝子間配列があることに注意すべきである。MEGES は、R 2、R 3、R 4 および R 5 のそれぞれと L 4 の 5 つの異なる gRNA 対を利用し、欠失を誘導した。その 2 日目にデジタルジャンクション PCR によって検証されたように、すべての標的領域の効率的な欠失を誘導した。MMGT 1 は必須であるが、SMIM 1 0 を含む gRNA の L 4 と R 4 部位の間の 5. 4 8 Mb 領域全体が、HAP 1 細胞の成長への大きな影響なしに欠失可能であることが示された。

[0099] L 4 – R 4 欠失クローンの特性評価

R 4 と L 4 のペアリング gRNA を用いて MEGES 後にシングルクローン展開して得られた 4 種類の HAP 1 細胞クローン (hHY 1 3 1, 1 4 5, 1 4 8, 1 4 9 ; L 4 – R 4 欠失クローン) を分析した。得られたクローンがプロピジウムヨード染色を用いたフローサイトメトリーを用いてハプロイドであることを確認した¹⁶。まず、欠失領域内に位置する個々の遺伝子座に対して PCR ジェノタイピングを行った。ペアの gRNA によって切断されたゲノム断片は、異なる染色体部位にランダムに統合されているのではないかと推測され、この領域内の必須遺伝子が非必須遺伝子であると判断された可能性もあるだろう。gRNA の R 4 部位と L 4 部位の間には 6 9 個のアノテーション遺伝子が存在するが、多くの遺伝子が重複しているか、あるいは非常に近い位置に存在する。そこで、L 4 – R 4 欠失クローンから抽出したゲノム DNA を用いて、6 9 個の遺伝子座を全てカバーできる 3 9 個の遺伝子座に対してジェノタイピング PCR を行った。図 3 a に示すように、試験した遺伝子座を全て保持していたオリジナル HAP 1 の 2 つのクローン (hHY 1 5 3、1 5 4) と比較して、L 4 – R 4 欠失クローンは、そのゲノムから試験した遺伝子座を全て欠失していることが分かった。

[0100]

[表6]

表6：拡張された欠失を有する細胞取得効率

欠失領域	R1	L1-R1	L2-R1	L3-R1	L4-R1	L5-R1
欠失の長さ [kbp]	0	508	3,628	3,725	4,347	4,379
分析したクローン数	6	12	12	12	24	6
欠失を有するクローン数	-	11	12	12	24	0
効率 [%]	-	92*	100	100	100	0
欠失領域	L4	L4-R2	L4-R3	L4-R4	L4-R5	
欠失の長さ [kbp]	0	4,542	4,577	5,476	5,498	
分析したクローン数	6	12	12	20	6	
欠失を有するクローン数	-	12	12	20	0	
効率 [%]	-	100	100	100	0	

[0101] 表6に示されるデータは、図2に示されるクローンの遺伝子型を分析することによって得られた。所望の遺伝子欠損を有する細胞クローンを欠失を有する細胞数として計数した。

[0102] また、69遺伝子欠失がトランスクリプトームに及ぼす影響を評価する（図3b）。このトランスクリプトーム解析では、RBMX2とMMGT1は、隣接する5.5Mb領域の欠失による転写レベルの変化がほとんどなく〔図3c〕、この欠失領域にこれらの遺伝子の遠位のシス調節要素が存在しないことが示された。

[0103] 必須要素の染色体間移植

RBMX2遺伝子とMMGT1遺伝子は、遠位のシス制御要素がなければ、細菌や酵母の遺伝子の多くがゲノムの設計および合成においてそうであったように、コンパクトで独立した遺伝子ユニットとして扱うことができるであろう。これを検証するために、両遺伝子を異なる染色体位置に移植することを試みた。BAC-YACベクターを用いて、転写開始部位の上流の～3kbp領域とポリアデニル化（pA）部位の下流の～3kbp領域を含む遺伝子領域全体（「レスキューカセット」）をクローニングした〔図4a〕。閉環型pHY262ベクター上では、IRES配列とpAシグナルとの間に

ピューロマイシンN-アセチルトランスフェラーゼ (PNAT) のコーディング配列が存在する。結果として得られたレスキューベクターをHAP1にトランスフェクトすると、gRNA/Cas9の2つの異なるベクターが共トランスフェクトされる；1つは、12番染色体上のGAPDH遺伝子の3' UTRに標的化されたgRNAであり、もう1つは、レスキュー遺伝子座ベクター上のIRES配列のすぐ外側に標的化されたものである¹⁵。ピューロマイシン選択により、IRES-Puro-pA+レスキューカセットがGAPDH遺伝子座に統合された細胞を得ることができる。リアルタイムRT-PCRにより、得られた細胞クローンは、RBMX2またはMMGT1のmRNAレベルが1.5~2.0倍に増加していることが実証された〔図4b〕。続いて、X染色体の内因性遺伝子座からRBMX2またはMMGT1遺伝子を欠失させるMEGESを行ったところ、同等のgRNAを介在させた欠失効率で、移植細胞はより多くの生存細胞コロニーを示した〔図4c〕。これらのデータから、両遺伝子座は内因性でない染色体部位への移植が可能であり、必須遺伝子としての役割を果たしていることが明らかになった。また、上流と下流の3kbの領域は、正常な細胞増殖に十分なレベルで両遺伝子を発現させるのに十分であることが明らかになった。特に、染色体間で必須遺伝子を移植することで、MEGESはHPRT1遺伝子座のゲノムをさらに欠失させ、次の必須遺伝子候補に到達することができた〔図5〕。

[0104] 非HPRT1遺伝子座でのMEGES

MEGESの範囲を広げるために、別のネガティブセレクションマーカーを用いて、異なるヒト細胞株HCT116において、HPRT1遺伝子座以外のゲノム領域を欠失させる試験を行った。大腸がんの男性患者から樹立されたHCT116細胞は、X染色体が1本であるため、HPRT1遺伝子座は1つである。単純ヘルペスウイルスチミジンキナーゼ (TK) 発現カセットは、TKをコードする配列に続いて、2Aペプチドを挟んだPNATを含み、これらはCMVプロモーターとSV40 pAシグナルによって発現さ

れる [図 6 a]。TK 発現カセットを含むプラスミドを gRNA/Cas9 の 2 つの異なるベクターと共トランスフェクトしたとき、1 つの gRNA は染色体 X 上の HCT 116 の非必須遺伝子 OCRL (Nucleic Acid Research, 2020, doi: 10.1093/nar/gkaa884) の 3' UTR に標的化され、もう 1 つの gRNA は TK 発現カセットベクター上の CMV プロモーターのすぐ上流に標的化された¹⁵。ピューロマイシン耐性のために選択された細胞は、ジャンクション PCR により、標的とされた OCRL 遺伝子座に TK 発現カセットを有することが確認された [図 6 a および 6 b]。次に、これらの細胞を MEGES に使用した。5 つの gRNA 対 (R9 との対では L8、L9、L10、L11 または L12) はすべて、2 日目にジャンクションデジタル PCR によって測定された 0.3 ~ 1.2 の欠失効率によって示されるように、OCRL 遺伝子座を含むゲノム領域を効率的に欠失させた [図 6 c]。TK がゲノム中に存在し、細胞内で発現している場合に細胞増殖障害を誘発するガンシクロビル選択は、L12 と R9 の対を除くすべての gRNA 対で有意に多くのコロニーを生じさせ、L11 と L12 の gRNA 標的部位の間の領域が細胞増殖に重要な要素を含んでいることを示唆した。このように、MEGES の基本戦略は、異なる選択マーカーと異なるヒト細胞を用いることで、どのような染色体部位にも適用可能であることが証明された。

[0105] 負の選択をしない MEGES

これまでに、細胞増殖に重要なゲノム領域のスクリーニングや同定、さらにはメガベーススケールでの標的ゲノム領域の欠失を有する細胞クローンの単離を行い、標的領域の機能的意義をさらに明らかにするために、MEGES の有用性と利点を実証してきた。しかし、必須領域のみをスクリーニングする場合には、事前にネガティブセレクションマーカーを統合する必要があり、また、ネガティブセレクションを用いたコロニー形成アッセイはスループットが低いという問題があった。そこで我々は、ネガティブ選択の代わりに、複数の時点でのデジタルジャンクション PCR を含む MEGES の

改良版を考案した。欠失した領域が細胞増殖に不可欠であるか、または重要である場合、欠失した領域を持つ細胞の割合は、細胞培養が進むにつれて減少する。この割合の変化は、gRNA/Cas9トランスフェクション後、複数のタイムポイントでデジタルジャンクションPCRを用いて決定した欠失効率を比較することで評価することができる。この戦略をデジタルジャンクションPCRベースのMEGES (dMEGES) と呼ぶ。その概念を実証するために、我々はdMEGESをHPRT1遺伝子座に適用した(図7)。我々は、2つのgRNAペア(R1-L4とR1-L5)による切断の後、2日目と17日目にデジタルジャンクションPCRを実行し、期待される結果を実証した。R1-L4が切断された細胞の割合は2日目から17日目にかけてほとんど変化しなかった(1.8%から2.0%)が、R1-L5が切断された細胞の割合は1.0%から0.2%と大幅に減少していた。ここで、RBMX2遺伝子座の内側を切断するR1とL13の別のgRNA対開裂を試験したところ、R1-L5の開裂は1.0%から0.2%まで有意に減少した。これもまた、欠失を有する細胞集団の有意な減少をもたらした(2.1%から0.6%)。これらのデータは、図2bのデータと一致しており、R1とL4の間の4.35Mbの領域は細胞増殖に影響を与えずに欠失させることができたが、R1-L5の領域はRBMX2の本質的な領域のために欠失させることができなかったことが示されている。

[0106] さらにX染色体に対して22領域それぞれの欠失を有する細胞集団を作成して同様に分析した(図8)。その結果、図8に示されるように、領域7および領域16の欠失は細胞の増殖または生存に必須である可能性が明らかになった。また、領域1、9~15、および17~19は、細胞の増殖または生存に正に影響する遺伝子を含む可能性が明らかとなった。これまでのCRISPRスクリーニングでは、生存に必須ではないとされてきた22領域の欠失であったが、本発明によると感度高く必須の染色体領域を決定することができた。

[0107] このようにして本願発明では、ネガティブ選択マーカーを事前に導入せず

にMEGESを実施することができ、この場合には、MEGESによって大規模なゲノム欠失を生じさせることができ、その上、細胞の生存に必須な遺伝子が存在する領域を決定することができる。生存に必須な遺伝子が決定された場合には、これを他の染色体上に統合し、さらにゲノムの欠失を拡張することもできるであろう。dMEGESは、細胞の生存に必須な遺伝子を持つこと、および／または、ゲノムの大規模欠失に有用な技術であり得る。

[0108]

材料と方法に関する参考資料

1. Chen, W.-H., Lu, G., Chen, X., Zhao, X.-M. & Bork, P. OGEE v2: an update of the online gene essentiality database with special focus on differentially essential genes in human cancer cell lines. *Nucleic Acids Res* 45, D940–D944 (2017).
2. Mair, B. et al. Essential Gene Profiles for Human Pluripotent Stem Cells Identify Uncharacterized Genes and Substrate Dependencies. *Cell Reports* 27, 599–615.e12 (2019).
3. Blomen, V. A. et al. Gene essentiality and synthetic lethality in haploid human cells. *Science* 350, 1092–1096 (2015).
4. Naito, Y., Hino, K., Bono, H. & Ui-Tei, K. CRISPRdirect: software for designing CRISPR/Cas guide RNA with reduced off-target sites. *Bioinformatics* 31, 1120–1123 (2015).
5. Ran, F. A. et al. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* 8, 2281–2308 (2013).
6. He, X. et al. Knock-in of large reporter genes in human cells via CRISPR/Cas9-induced homology-dependent and independent DNA repair. *Nucleic Acids Res* 44, e85 (2016).
7. Zhang, W. et al. Engineering the ribosomal DNA in a megabase synthetic chromosome. *Science* 355, (2017).
8. Muller, H. et al. Assembling Large DNA Segments in Yeast. in *Gene Synthesis* (ed. Peccoud, J.) vol. 852 133–150 (Humana Press, 2012).
9. Wu, Z. et al. Absolute quantification of DNA methylation using microfluidic chip-based digital PCR. *Biosensors and Bioelectronics* 96, 339–344 (2017).
10. Guda, K., Natale, L. & Markowitz, S. D. An improved method for staining cell colonies in clonogenic assays. *Cytotechnology* 54, 85–88 (2007).

[0109]

材料と方法以外に関する参考文献

1. Xavier, J. C., Patil, K. R. & Rocha, I. Systems Biology Perspectives on Minimal and Simpler Cells. *Microbiol Mol Biol Rev* 78, 487–509 (2014).
2. Posfai, G. Emergent Properties of Reduced-Genome *Escherichia coli*. *Science* 312, 1044–1046 (2006).
3. Hutchison, C. A. et al. Design and synthesis of a minimal bacterial genome. *Science* 351, aad6253–aad6253 (2016).
4. Franchini, L. F. & Pollard, K. S. Human evolution: the non-coding revolution. *BMC Biology* 15, (2017).
5. Feschotte, C. & Pritham, E. J. DNA Transposons and the Evolution of Eukaryotic Genomes. *Annu Rev Genet* 41, 331–368 (2007).
6. Keeling, P. J. & Palmer, J. D. Horizontal gene transfer in eukaryotic evolution. *Nat Rev Genet* 9, 605–618 (2008).
7. Wang, T. et al. Identification and characterization of essential genes in the human genome. *Science* 350, 1096–1101 (2015).
8. He, Z. et al. Highly efficient targeted chromosome deletions using CRISPR/Cas9: Highly Efficient Targeted Chromosome Deletions. *Biotechnology and Bioengineering* 112, 1060–1064 (2015).
9. Wang, L. et al. Large genomic fragment deletion and functional gene cassette knock-in via Cas9 protein mediated genome editing in one-cell rodent embryos. *Sci Rep* 5, (2015).
10. Zhang, L. et al. Large Genomic Fragment Deletions and Insertions in Mouse Using CRISPR/Cas9. *PLOS ONE* 10, e0120396 (2015).
11. Essletzbichler, P. et al. Megabase-scale deletion using CRISPR/Cas9 to generate a fully haploid human cell line. *Genome Res* 24, 2059–2065 (2014).
12. Zheng, Q. et al. Precise gene deletion and replacement using the CRISPR/Cas9 system in human cells. *BioTechniques* 57, (2014).
13. Canver, M. C. et al. Characterization of Genomic Deletion Efficiency Mediated by Clustered Regularly Interspaced Palindromic Repeats (CRISPR)/Cas9 Nuclease System in Mammalian Cells. *J. Biol. Chem.* 289, 21312–21324 (2014).

14. Lee, H. J., Kim, E. & Kim, J.-S. Targeted chromosomal deletions in human cells using zinc finger nucleases. *Genome Res.* 20, 81–89 (2010).
15. He, X. et al. Knock-in of large reporter genes in human cells via CRISPR/Cas9-induced homology-dependent and independent DNA repair. *Nucleic Acids Res* 44, e85 (2016).
16. Olbrich, T. et al. A p53-dependent response limits the viability of mammalian haploid cells. *Proceedings of the National Academy of Sciences* 114, 9367–9372 (2017).
17. Dixon, J. R. et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485, 376–380 (2012).
18. Nora, E. P. et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* 485, 381–385 (2012).
19. Lupiáñez, D. G. et al. Disruptions of Topological Chromatin Domains Cause Pathogenic Rewiring of Gene-Enhancer Interactions. *Cell* 161, 1012–1025 (2015).
20. Guo, Y. et al. CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function. *Cell* 162, 900–910 (2015).
21. Haarhuis, J. H. I. et al. The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. *Cell* 169, 693–707.e14 (2017).
22. Wang, Y. et al. The 3D Genome Browser: a web-based browser for visualizing 3D genome organization and long-range chromatin interactions. *Genome Biol* 19, 151 (2018).
23. Suzuki, K. et al. Highly efficient transient gene expression and gene targeting in primate embryonic stem cells with helper-dependent adenoviral vectors. *Proc Natl Acad Sci U S A* 105, 13781–13786 (2008).
24. Frank, S., Skryabin, B. V. & Greber, B. A modified TALEN-based system for robust generation of knock-out human pluripotent stem cell lines and disease models. *BMC Genomics* 14, 773 (2013).
25. Moore, M. M., Parker, L., Huston, J., Harrington-Brock, K. & Dearfield, K. L. Comparison of mutagenicity results for nine compounds evaluated at the hgp_{rt} locus in the standard and suspension CHO assays. *Mutagenesis* 6, 77–85 (1991).
26. Chen, W.-H., Minguéz, P., Lercher, M. J. & Bork, P. OGEE: an online gene essentiality database. *Nucleic Acids Research* 40, D901–D906 (2012).

請求の範囲

- [請求項1] インビトロの方法であって、
- (a) 単離された細胞を含む細胞集団を用意することと、
 - (b) 前記細胞集団中の前記細胞のゲノムDNAに、ゲノムDNA上の2箇所の標的配列を配列特異的に切断することができる配列特異的核酸切断分子をそれぞれ作用させて、ゲノムDNA上の2箇所の標的配列においてそれぞれ切断を生じさせ、これによって、細胞集団中の少なくとも一部の細胞において、前記2箇所の切断部位の間の領域においてDNA領域の欠失を生じさせることと、
 - (c) その後、得られた細胞集団を培養し、細胞の増殖または生存に対する前記DNA領域の欠失の影響を決定することと、
- を含む、
方法。
- [請求項2] 前記(c)において、前記決定が、DNA領域の欠失効率またはその推定値を決定することと、その後の培養後の細胞集団におけるDNA領域の欠失を有する細胞の割合と、決定された欠失効率またはその推定値とを比較することとにより行われる、請求項1に記載の方法。
- [請求項3] 前記(c)において、欠失効率またはその推定値の決定、および、培養後のDNA領域の欠失を有する細胞の割合が、それぞれ細胞集団を含む懸濁液中に含まれる全細胞に対する欠失を有する細胞の割合として決定される、請求項2に記載の方法。
- [請求項4] 上記割合が、懸濁液中に含まれる、欠失を有するゲノムDNAおよび有しないゲノムDNAの計数技術により決定される、請求項3に記載の方法。
- [請求項5] 請求項1～4のいずれか一項に記載の方法であって、
- (d) 対照となる細胞集団と比較して、前記(c)における細胞の生存および／または増殖に対する影響の有無または大きさに基づいて、対照のゲノムDNAと比較して欠失したDNA領域が細胞の生存およ

び／または増殖を制御する遺伝子を含むか否かを決定すること
をさらに含む、方法。

[請求項6] 請求項5に記載の方法であって、対照となる細胞集団が、前記（b）により、より大きなDNA領域、より小さなDNA領域、または異なるDNA領域の欠失を有する細胞を含む細胞集団である、方法。

[請求項7] 請求項5または6に記載の方法であって、対照のゲノムDNAと比較して欠失したDNA領域に存在する遺伝子から、細胞の生存および／または増殖を制御する遺伝子を少なくとも1つ特定することをさらに含む、方法。

[請求項8] 請求項1～7のいずれか一項に記載の方法であって、
（f）欠失したDNA領域に存在する遺伝子から、細胞の生存および／または増殖を制御する遺伝子を少なくとも1つ特定すること
をさらに含む、方法。

[請求項9] （g）制御配列に作動可能に連結した細胞の生存および／または増殖を制御する遺伝子の少なくとも1つを、当該DNA領域の欠失を有するゲノムDNAに異所的に導入することをさらに含む、請求項7または8に記載の方法。

[請求項10] 前記（b）において欠失するDNA領域のサイズが、0.5Mbp以上である、請求項1～9のいずれか一項に記載の方法。

[請求項11] 前記（b）において欠失するDNA領域のサイズが、1Mbp以上である、請求項1～10のいずれか一項に記載の方法。

[請求項12] インビトロの方法であって、
（α）単離された細胞を含む細胞集団を用意することと、ここで単離された細胞は、欠失させる領域中にネガティブ選択用マーカー遺伝子を含み、
（β）前記細胞集団中の前記細胞のゲノムDNAに、ゲノムDNA上の2箇所の標的配列を配列特異的に切断することができる配列特異的核酸切断分子をそれぞれ作用させて、ゲノムDNA上の2箇所の標的

配列においてそれぞれ切断を生じさせ、これによって、細胞集団中の少なくとも一部の細胞において、前記2箇所の切断部位の間の領域においてDNA領域の欠失を生じさせることと、ここで、前記2箇所は、前記ネガティブ選択用マーカー遺伝子を挟む位置に設計され、
(γ) ネガティブ選択マーカー遺伝子を有しない細胞を選択することと、
を含む、方法。

[請求項13] 前記ネガティブ選択マーカー遺伝子が、欠失させる領域中に外来的に挿入されたものである、請求項12に記載の方法。

[請求項14] 前記ネガティブ選択マーカー遺伝子が、ゲノムDNA中の内在的な遺伝子である、請求項12に記載の方法。