(19)日本国特許庁(JP)　(12)公 表 特 許 公 報(A)　(11)特許出願公表番号
特表2005-525082
(P2005-525082A)
(43)公表日　平成17年8月25日(2005.8.25)

(51)Int.Cl.[7]	FI		テーマコード（参考）
C12N 15/09	C12N 15/00	ZNAA	4B024
C12Q 1/68	C12Q 1/68	Z	4B063
G01N 33/53	G01N 33/53	M	
G01N 33/566	G01N 33/566		
G01N 37/00	G01N 37/00	102	

審査請求 未請求　予備審査請求 有　（全 106 頁）　最終頁に続く

(21)出願番号　特願2003-523626(P2003-523626)
(86)(22)出願日　平成14年8月30日(2002.8.30)
(85)翻訳文提出日　平成16年3月18日(2004.3.18)
(86)国際出願番号　PCT/US2002/028471
(87)国際公開番号　WO2003/018775
(87)国際公開日　平成15年3月6日(2003.3.6)
(31)優先権主張番号　60/316,333
(32)優先日　平成13年8月30日(2001.8.30)
(33)優先権主張国　米国(US)
(31)優先権主張番号　2,356,536
(32)優先日　平成13年8月31日(2001.8.31)
(33)優先権主張国　カナダ(CA)
(31)優先権主張番号　60/380,546
(32)優先日　平成14年5月13日(2002.5.13)
(33)優先権主張国　米国(US)

(71)出願人　598038968
エモリー・ユニバーシティ
アメリカ合衆国、３０３２２　ジョージア州、アトランタ、リッジウッド・ドライブ　２００９
(74)代理人　100078282
弁理士　山本　秀策
(74)代理人　100062409
弁理士　安村　高明
(74)代理人　100113413
弁理士　森下　夏樹
(72)発明者　ウォラス，　ダグラス　シー．
アメリカ合衆国　ジョージア　３０３４５，　アトランタ，　エコー　ドライブ　２５０６

最終頁に続く

(54)【発明の名称】ヒトミトコンドリアＤＮＡ多型、ハプログループ、生理学的状態との関連、および遺伝子型決定アレイ

(57)【要約】
本発明は、主要なヒトハプログループを示すヒトｍｔＤＮＡ多型、ならびにこれらのハプログループおよび選択されたハプログループを診断する方法を提供する。本発明は、具体的には、以下の方法を提供する：ヒトのハプログループを診断するための方法であって、ａ）該ヒト由来のミトコンドリア核酸を含むサンプルを提供する工程；およびｂ）ハプログループを示す少なくとも１つのヌクレオチド対立遺伝子の存在または非存在を、該サンプル中で同定する工程、を包含する、方法。

【特許請求の範囲】
【請求項１】
ヒトのハプログループを診断するための方法であって、該方法は、以下の工程：
　ａ）該ヒト由来のミトコンドリア核酸を含むサンプルを提供する工程；および
　ｂ）ハプログループを示す少なくとも１つのヌクレオチド対立遺伝子の存在または非存在を、該サンプル中で同定する工程、
を包含する、方法。
【請求項２】
請求項１に記載の方法であって、前記ハプログループは、ハプログループＬ１であり、前記方法の工程ｂ）は、８２５Ａ、２７５８Ａ、２８８５Ｃ、７１４６Ｇ、８４６８Ｔ、８６５５Ｔ、１０６８８Ａ、１０８１０Ｃ、および１３１０５Ｇからなる群より選択される少なくとも１つのヌクレオチド対立遺伝子を前記サンプル中で同定する工程を包含する、方法。
【請求項３】
請求項１に記載の方法であって、前記ハプログループは、ハプログループＬ２であり、前記方法の工程ｂ）は、２４１６Ｃ、２７５８Ｇ、８２０６Ａ、９２２１Ｇ、１１９４４Ｃ、および１６３９０Ｇからなる群より選択される少なくとも１つのヌクレオチド対立遺伝子を前記サンプル中で同定する工程を包含する、方法。
【請求項４】
請求項１に記載の方法であって、前記ハプログループは、ハプログループＬ３であり、前記方法の工程ｂ）は、１０８１９Ｇ、１４２１２Ｃ、８６１８Ｃ、１００８６Ｃ、１６３６２Ｃ、１０３９８Ａ、および１６１２４Ｃからなる群より選択される少なくとも１つのヌクレオチド対立遺伝子を前記サンプル中で同定する工程を包含する、方法。
【請求項５】
請求項１に記載の方法であって、前記ハプログループは、ハプログループＣであり、前記方法の工程ｂ）は、３５５２Ｃ、４７１５Ｇ、７１９６Ａ、８５８４Ａ、９５４５Ｇ、１３２６３Ｇ、１４３１８Ｃ、および１６３２７Ｔからなる群より選択される少なくとも１つのヌクレオチド対立遺伝子を前記サンプル中で同定する工程を包含する、方法。
【請求項６】
請求項１に記載の方法であって、前記ハプログループは、ハプログループＤであり、前記方法の工程ｂ）は、４８８３Ｔ、５１７８Ａ、８４１４Ｔ、１４６６８Ｔ、および１５４８７Ｔからなる群より選択される少なくとも１つのヌクレオチド対立遺伝子を前記サンプル中で同定する工程を包含する、方法。
【請求項７】
請求項１に記載の方法であって、前記ハプログループは、ハプログループＥであり、前記方法の工程ｂ）は、ヌクレオチド対立遺伝子１６２２７Ｇを前記サンプル中で同定する工程を包含する、方法。
【請求項８】
請求項１に記載の方法であって、前記ハプログループは、ハプログループＧであり、前記方法の工程ｂ）は、４８３３Ｇ、８２００Ｃ、および１６０１７Ｃからなる群より選択される少なくとも１つのヌクレオチド対立遺伝子を前記サンプル中で同定する工程を包含する、方法。
【請求項９】
請求項１に記載の方法であって、前記ハプログループは、ハプログループＺであり、前記方法の工程ｂ）は、１１０７８Ｇ、１６１８５Ｔ、および１６２６０Ｔからなる群より選択される少なくとも１つのヌクレオチド対立遺伝子を前記サンプル中で同定する工程を包含する、方法。
【請求項１０】
請求項１に記載の方法であって、前記ハプログループは、ハプログループＡであり、前記方法の工程ｂ）は、６６３Ｇ、１６２９０Ｔ、および１６３１９Ａからなる群より選択さ

れる少なくとも１つのヌクレオチド対立遺伝子を前記サンプル中で同定する工程を包含する、方法。

【請求項１１】

請求項１に記載の方法であって、前記ハプログループは、ハプログループＩであり、前記方法の工程ｂ）は、４５２９Ｔ、１００３４Ｃ、および１６３９１Ａからなる群より選択される少なくとも１つのヌクレオチド対立遺伝子を前記サンプル中で同定する工程を包含する、方法。

【請求項１２】

請求項１に記載の方法であって、前記ハプログループは、ハプログループＷであり、前記方法の工程ｂ）は、２０４Ｃ、２０７Ａ、１２４３Ｃ、５０４６Ａ、５４６０Ａ、８９９４Ａ、１１９４７Ｇ、１５８８４Ｃ、および１６２９２Ｔからなる群より選択される少なくとも１つのヌクレオチド対立遺伝子を前記サンプル中で同定する工程を包含する、方法。

【請求項１３】

請求項１に記載の方法であって、前記ハプログループは、ハプログループＸであり、前記方法の工程ｂ）は、１７１９Ａ、３５１６Ｇ、６２２１Ｃ、および１４４７０Ｃからなる群より選択される少なくとも１つのヌクレオチド対立遺伝子を前記サンプル中で同定する工程を包含する、方法。

【請求項１４】

請求項１に記載の方法であって、前記ハプログループは、ハプログループＢであり、前記方法の工程ｂ）は、以下：

１）ヌクレオチド対立遺伝子１６１８９Ｃを前記サンプル中で同定する工程；

２）１７１９Ａ、３５１６Ｇ、６２２１Ｃ、１４４７０Ｃ、および１６２７８Ｔからなる群より選択されるヌクレオチド対立遺伝子の非存在を前記サンプル中で同定する工程；および

３）１８８８Ａ、４２１６Ｃ、４９１７Ｇ、８６９７Ａ、１０４６３Ｃ、１１２５１Ｇ、１１４６７Ｇ、１２３０８Ｇ、１２３７２Ａ、１２６３３Ｔ、１３１０４Ｇ、１３３６８Ａ、１４０７０Ｇ、１４９０５Ａ、１５４５２Ａ、１５６０７Ｇ、１５９２８Ａ、１６１２６Ｃ、１６１６３Ｃ、１６１８６Ｔ、１６２４９Ｃ、および１６２９４Ｔからなる群より選択されるヌクレオチド対立遺伝子の非存在を前記サンプル中で同定する工程、

を包含する、方法。

【請求項１５】

請求項１に記載の方法であって、前記ハプログループは、ハプログループＦであり、前記方法の工程ｂ）は、１２４０６Ａおよび１６３０４Ｃからなる群より選択される少なくとも１つのヌクレオチド対立遺伝子を前記サンプル中で同定する工程を包含する、方法。

【請求項１６】

請求項１に記載の方法であって、前記ハプログループは、ハプログループＹであり、前記方法の工程ｂ）は、７９３３Ｇ、８３９２Ａ、１６２３１Ｃ、および１６２６６ＴＣからなる群より選択される少なくとも１つのヌクレオチド対立遺伝子を前記サンプル中で同定する工程を包含する、方法。

【請求項１７】

請求項１に記載の方法であって、前記ハプログループは、ハプログループＵであり、前記方法の工程ｂ）は、３１９７Ｃ、４６４６Ｃ、７７６８Ｇ、９０５５Ａ、１１３３２Ｔ、１３１０４Ｇ、１４０７０Ｇ、１５９０７Ｇ、１６０５１Ｇ、１６１２９Ｃ、１６１７２Ｃ、１６２１９Ｇ、１６２４９Ｃ、１６２７０Ｔ、１６３１１Ｔ、１６３１８Ｔ、１６３４３Ｇ、および１６３５６Ｃからなる群より選択される少なくとも１つのヌクレオチド対立遺伝子を前記サンプル中で同定する工程を包含する、方法。

【請求項１８】

請求項１に記載の方法であって、前記ハプログループは、ハプログループＪであり、前記方法の工程ｂ）は、２９５Ｔ、１２６１２Ｇ、１３７０８Ａ、および１６０６９Ｔからな

る群より選択される少なくとも１つのヌクレオチド対立遺伝子を前記サンプル中で同定する工程を包含する、方法。

【請求項１９】

請求項１に記載の方法であって、前記ハプログループは、ハプログループＴであり、前記方法の工程ｂ）は、１１８１２Ｇ、１２６３３Ｔ、１４２３３Ｇ、１６１６３Ｃ、１６１８６Ｔ、１８８８Ａ、４９１７Ｇ、８６９７Ａ、１０４６３Ｃ、１３３６８Ａ、１４９０５Ａ、１５６０７Ｇ、１５９２８Ａ、および１６２９４Ｔからなる群より選択される少なくとも１つのヌクレオチド対立遺伝子を前記サンプル中で同定する工程を包含する、方法。

【請求項２０】

請求項１に記載の方法であって、前記ハプログループは、ハプログループＶであり、前記方法の工程ｂ）は、７２Ｃ、４５８０Ａ、および１５９０４Ｔからなる群より選択される少なくとも１つのヌクレオチド対立遺伝子を前記サンプル中で同定する工程を包含する、方法。

【請求項２１】

請求項１に記載の方法であって、前記ハプログループは、ハプログループＨであり、前記方法の工程ｂ）は、２７０６Ａおよび７０２８Ｃからなる群より選択される少なくとも１つのヌクレオチド対立遺伝子を前記サンプル中で同定する工程を包含する、方法。

【請求項２２】

請求項１に記載の方法であって、前記ハプログループは、ハプログループＬＯであり、前記方法の工程ｂ）は、４５８６Ｃ、９８１８Ｔ、および８１１３Ａからなる群より選択される少なくとも１つのヌクレオチド対立遺伝子を前記サンプル中で同定する工程を包含する、方法。

【請求項２３】

請求項１に記載の方法であって、前記同定する工程は、既知の位置にて基材に付着された、２つ以上の単離された核酸分子を含むアレイを使用して行われ、各分子は、約７～約３０ヌクレオチドの長さを有し、各分子は、表３の第１列に列挙されるものからなる座の群より選択される座にて少なくとも１つのヌクレオチド対立遺伝子を含む、配列番号１の一部と同一な配列を含む、方法。

【請求項２４】

機械読み取り可能な形態にてコードされるデータセットを含む機械読み取り可能な格納デバイスであって、該データセットは、複数のヌクレオチド対立遺伝子および各対立遺伝子と関連するハプログループ名称を含む、デバイス。

【請求項２５】

請求項２４に記載の格納デバイスを含み、１以上のヌクレオチド対立遺伝子を含むデータセットを入力するための入力手段もまた含む、プログラム格納デバイスであって、該デバイスはまたは、該入力ヌクレオチド対立遺伝子と、関連するハプログループとを関連づけ、結果を表示することにより、ハプログループを診断するためのプログラム工程を包含する、デバイス。

【請求項２６】

進化的に重要な遺伝子を同定するための方法であって、該方法は、以下の工程：

ａ）少なくとも１つの対立遺伝子の核酸配列またはその一部分を含むヌクレオチド配列の第１のセットを、第１の集団から提供する工程；

ｂ）対応する少なくとも１つの対立遺伝子の核酸配列またはその一部分を含むヌクレオチド配列の第２のセットを、第２の種内集団から提供する工程；

ｃ）該第１のセットと該第２のセットとを比較して、データセットを生成する工程を包含する、中立分析（ｎｅｕｔｒａｌｉｔｙ　ａｎａｌｙｓｉｓ）を行う工程；および

ｄ）該データセットを分析して、進化的に重要な遺伝子を同定する工程、

を包含する、方法。

【請求項２７】

請求項２６に記載の方法であって、１を超える前記対立遺伝子は、ミトコンドリアゲノム中に位置する、方法。

【請求項２８】

請求項２６に記載の方法であって、前記集団はヒト集団である、方法。

【請求項２９】

請求項２６に記載の方法であって、前記第１の集団および／または前記第２の集団は、少なくとも１つの亜集団を含み、該亜集団は、マクロハプログループ、ハプログループ、サブハプログループ、および個体からなる群より選択される、方法。

【請求項３０】

請求項２６に記載の方法であって、前記第２のセットのヌクレオチド配列は、配列番号２の一部に同一な少なくとも１００ヌクレオチドを含む、方法。

【請求項３１】

請求項２６に記載の方法であって、前記進化的に重要な遺伝子は、ＮＤ１、ＮＤ２、ＮＤ３、ＮＤ４、ＮＤ５、ＮＤ６、Ｃｙｔｂ、ＣＯＩ、ＣＯＩＩ、ＣＯＩＩＩ、ＡＴＰ６、およびＡＴＰ８からなる群より選択されるミトコンドリア遺伝子である、方法。

【請求項３２】

請求項３１に記載の方法であって、前記進化的に重要なミトコンドリア遺伝子は、ＣＯＩＩＩおよびＡＴＰ６からなる群より選択される、方法。

【請求項３３】

請求項２６に記載の方法であって、該方法はまた、前記第１のヌクレオチド配列と前記第２のヌクレオチド配列との間の配列差異を同定することによって、少なくとも１つの進化的に重要なヌクレオチド対立遺伝子を同定する工程を包含する、方法。

【請求項３４】

請求項３３に記載の方法であって、該方法はまた、前記進化的に重要なヌクレオチド対立遺伝子を含むコドンによってコードされる進化的に重要なアミノ酸対立遺伝子を決定することによって、該進化的に重要なアミノ酸対立遺伝子を同定する工程を包含する、方法。

【請求項３５】

請求項３４に記載の方法であって、該方法はまた、生理学的状態を有する個体を前記第１の集団として使用し、該生理学的状態を有さない個体を前記第２の集団として使用することによって、該生理学的状態に対する素因を示すアミノ酸対立遺伝子を同定する工程を包含する、方法。

【請求項３６】

進化的に重要な遺伝子を同定するための方法であって、該方法は、以下の工程：

ａ）２以上の対応する対立遺伝子を含むヌクレオチド配列のセットを、１つの種の１つの集団から提供する工程；

ｂ）対応する対立遺伝子の該ヌクレオチド配列を比較して、データセットを生成する工程を包含する、中立分析を行う工程；および

ｃ）該データセットを分析して、進化的に重要な遺伝子を同定する工程、

を包含する、方法。

【請求項３７】

請求項３６に記載の方法であって、該方法はまた、進化的に重要な遺伝子の核酸配列を分析して、進化的に重要なヌクレオチド配列を同定することによって、進化的に重要なヌクレオチド対立遺伝子を同定する工程を包含する、方法。

【請求項３８】

請求項３７に記載の方法であって、該方法はまた、前記進化的に重要なヌクレオチド対立遺伝子を含むコドンによってコードされる進化的に重要なアミノ酸対立遺伝子を決定することによって、進化的に重要なアミノ酸対立遺伝子を同定する工程を包含する、方法。

【請求項３９】

選択された生理学的状態に対する素因を有する個体を診断するための方法であって、該方法は、以下の工程：

ａ）個体由来のミトコンドリア核酸分子を含むサンプルを提供する工程；
ｂ）該個体が居住している地理的地域を同定するための情報を提供する工程；
ｃ）該地理的地域に発生しているハプログループのセットを同定する情報を提供する工程；
ｄ）該サンプルから、該個体のハプログループを決定する工程；
ｅ）該個体の該ハプログループと、該地理的地域に発生しているハプログループとを比較する工程；および
ｆ）該個体の該ハプログループが、該地理的地域に発生しているハプログループの該セット内に存在しない場合、該選択された生理学的状態に対する素因を有する個体を診断する工程、
を包含する、方法。

【請求項４０】
請求項３９に記載の方法であって、前記生理学的状態は、エネルギー的不均衡、代謝疾患、異常なエネルギー代謝、異常な温度調節、異常な酸化的リン酸化、異常な電子伝達系、肥満、体脂肪の量、糖尿病、高血圧症、および心臓血管疾患からなる群より選択される、方法。

【請求項４１】
請求項３９に記載の方法であって、該方法はまた、アミノ酸対立遺伝子と、前記生理学的状態とを関連づける工程を包含し、該方法は、前記個体の前記ハプログループを診断するために有用なアミノ酸対立遺伝子を選択する工程を包含し、該アミノ酸対立遺伝子の存在は、該個体が居住している前記地理的地域に発生しているハプログループの前記セットにおいて、１を超えるハプログループを診断するために有用でない、方法。

【請求項４２】
請求項４１に記載の方法であって、前記ハプログループは、ハプログループＧであり、前記アミノ酸対立遺伝子は、ｎｔｌ　４８３３　Ａである、方法。

【請求項４３】
請求項４１に記載の方法であって、前記ハプログループは、ハプログループＴであり、前記アミノ酸対立遺伝子は、ｎｔｌ　４９１７　Ｄ、ｎｔｌ　８７０１　Ｔ、およびｎｔｌ　１５４５２　Ｉからなる群より選択される、方法。

【請求項４４】
請求項４１に記載の方法であって、前記ハプログループは、ハプログループＷであり、前記アミノ酸対立遺伝子は、ｎｔｌ　５０４６　Ｉ、ｎｔｌ　５４６０　Ｔ、ｎｔｌ　８７０１　Ｔ、およびｎｔｌ　１５８８４　Ｐからなる群より選択される、方法。

【請求項４５】
請求項４１に記載の方法であって、前記ハプログループは、ハプログループＤであり、前記アミノ酸対立遺伝子は、ｎｔｌ　５１７８　Ｍおよびｎｔｌ　８４１４　Ｆからなる群より選択される、方法。

【請求項４６】
請求項４１に記載の方法であって、前記ハプログループは、ハプログループＬＯであり、前記アミノ酸対立遺伝子は、ｎｔｌ　５４４２　Ｌ、ｎｔｌ　７１４６　Ａ、ｎｔｌ　９４０２　Ｐ、ｎｔｌ　１３１０５　Ｖ、およびｎｔｌ　１３２７６　Ｖからなる群より選択される、方法。

【請求項４７】
請求項４１に記載の方法であって、前記ハプログループは、ハプログループＬ１であり、前記アミノ酸対立遺伝子は、ｎｔｌ　７１４６　Ａ、ｎｔｌ　７３８９　Ｈ、ｎｔｌ　１３１０５Ｖ、ｎｔｌ　１３７８９　Ｈ、およびｎｔｌ　１４１７８　Ｖからなる群より選択される、方法。

【請求項４８】
請求項４１に記載の方法であって、前記ハプログループは、ハプログループＣであり、前記アミノ酸対立遺伝子は、ｎｔｌ　８５８４　Ｔおよびｎｔｌ　１４３１８　Ｓからなる

群より選択される、方法。

【請求項４９】

請求項４１に記載の方法であって、前記ハプログループは、ハプログループＡ、Ｉ、Ｘ、Ｂ、Ｆ、Ｙ、およびＵからなる群より選択され、前記アミノ酸対立遺伝子は、ｎｔｌ　８７０１Ｔである、方法。

【請求項５０】

請求項４１に記載の方法であって、前記ハプログループは、ハプログループＧであり、前記アミノ酸対立遺伝子は、ｎｔｌ　８７０１　Ｔ、ｎｔｌ　１３７０８　Ｔ、およびｎｔｌ　１５４５２　Ｉからなる群より選択される、方法。

【請求項５１】

請求項４１に記載の方法であって、前記ハプログループは、ハプログループＶおよびＨからなる群より選択され、前記アミノ酸対立遺伝子は、ｎｔｌ　８７０１　Ｔおよびｎｔｌ　１４７６６　Ｔからなる群より選択される、方法。

【請求項５２】

請求項４１に記載の方法であって、前記アミノ酸対立遺伝子をコードするコドン中に含まれるヌクレオチド対立遺伝子は、表１のヌクレオチド対立遺伝子ではない、方法。

【請求項５３】

請求項３９に記載の工程が、機械読み取り可能形態にてコードされるプログラム格納デバイスであって、該デバイスはまた、前記個体が居住している地理的地域を同定するための前記情報および該機械読み取り可能な形態にて前記地理的地域に発生しているハプログループのセットをコードする格納媒体を含む、デバイス。

【請求項５４】

進化的に重要なヒトミトコンドリアヌクレオチド対立遺伝子からなる群より選択されるヌクレオチド対立遺伝子を含む機械読み取り可能な形態にてコードされるデータセットを含む格納デバイスであって、各該対立遺伝子は、該格納デバイス中にて、ヒトにおける生理学的状態を同定するコードされた情報と関連づけられる、デバイス。

【請求項５５】

請求項５４に記載の格納デバイスであって、前記生理学的状態は、エネルギー的不均衡、代謝疾患、異常なエネルギー代謝、異常な温度調節、異常な酸化的リン酸化、異常な電子伝達系、肥満、体脂肪の量、糖尿病、高血圧症、および心臓血管疾患からなる群より選択される、デバイス。

【請求項５６】

請求項５４に記載の格納デバイスであって、該デバイスはまた、前記各ヌクレオチド対立遺伝子と、出身地理的地域とを関連づける情報をコードする工程を包含する、デバイス。

【請求項５７】

請求項５４に記載の格納媒体を含み、個体のハプログループおよび該個体の地理的地域を入力するための入力手段もまた含む、プログラム格納デバイスであって、該デバイスは、該個体を、生理的状態に対する素因を有すると診断するためのプログラム工程をさらに含む、媒体。

【請求項５８】

機械読み取り可能な形態にてコードされたデータセットを含む格納デバイスであって、該デバイスは、ｃｏｍｐｒｉｓｉｎｇ　進化的に重要なヒトミトコンドリアアミノ酸対立遺伝子を含み、該対立遺伝子の各々は、ヒトにおける生理学的状態を同定するコードされた情報と、該格納デバイス中で関連づけられる、デバイス。

【請求項５９】

ヒトにおけるＬＨＯＮに対する素因を診断するための方法であって、該方法は、以下の工程：

ａ）該ヒトに由来するサンプルを提供する工程；

ｂ）該サンプル中で、ヌクレオチド対立遺伝子１０６６３Ｃを同定する工程；および

ｃ）該サンプル中で、遺伝子ＮＤ５のアミノ酸４５８位のスレオニンをコードするヌク

レオチド対立遺伝子を同定する工程、
を包含し、該ヌクレオチド対立遺伝子の存在が、ＬＨＯＮに対する素因を示す、方法。
【請求項６０】
ヒトにおけるＬＨＯＮに対する素因を診断するための方法であって、肺胞法は、以下の工程：
ａ）該ヒトに由来するサンプルを提供する工程；
ｂ）該サンプル中で、ヌクレオチド対立遺伝子１０６６３Ｃを同定する工程；および
ｃ）該サンプル中で、２９５Ｔ、１２６１２Ｇ、１３７０８Ａ、および１６０６９Ｔからなる群より選択される少なくとも１つのヌクレオチド対立遺伝子を同定する工程、
を包含し、該ヌクレオチド対立遺伝子の存在が、ＬＨＯＮに対する素因を示す、方法。
【請求項６１】
ヒトにおけるＬＨＯＮに対する素因を診断するための方法であって、肺胞法は、以下の工程：
ａ）該ヒトに由来するサンプルを提供する工程；ならびに
ｂ）該サンプル中で、３６３５Ａおよび４６４０Ｃからなる群より選択されるヌクレオチド対立遺伝子を同定する工程、
を包含し、該ヌクレオチド対立遺伝子の存在が、ＬＨＯＮに対する素因を示す、方法。
【請求項６２】
ヒトにおいて失明する可能性の増大を診断するための方法であって、該方法は、以下の工程：
ａ）該ヒトに由来するサンプルを提供する工程；
ｂ）該サンプル中で、１１７７８Ａ、１４４８４Ｃおよび１０６６３Ｃからなる群より選択されるヌクレオチド対立遺伝子を同定する工程；ならびに
ｃ）該サンプル中で、遺伝子ＮＤ５のアミノ酸４５８位のスレオニンをコードするヌクレオチド対立遺伝子を同定する工程、
を包含し、該ヌクレオチド対立遺伝子の存在は、失明する素因を示す、方法。
【請求項６３】
少なくとも２つの単離された核酸分子を含むライブラリーであって、各分子は、約７～約３０ヌクレオチドの長さを有し、各分子は、表３の第１列に列挙されるものからなる座の群より選択される座に少なくとも１つのヌクレオチド対立遺伝子を含む、配列番号１の一部分と同一な配列を含む、ライブラリー。
【請求項６４】
請求項６３に記載のライブラリーであって、少なくとも１つの分子は、表３の非ＣａｍｂｒｉｄｇｅヒトｍｔＤＮＡヌクレオチド対立遺伝子からなる群より選択されるヌクレオチド対立遺伝子を含む配列を有する、ライブラリー。
【請求項６５】
請求項６３に記載のライブラリーであって、少なくとも１つの分子は、表４の非ＣａｍｂｒｉｄｇｅヒトｍｔＤＮＡヌクレオチド対立遺伝子からなる群より選択されるヌクレオチド対立遺伝子を含む配列を有する、ライブラリー。
【請求項６６】
請求項６３に記載のライブラリーであって、少なくとも１つの分子は、ヒトハプログループおよびマクロハプログループ（表１１）を診断するために有用なヌクレオチド対立遺伝子からなる群より選択されるヌクレオチド対立遺伝子を含む配列を有する、ライブラリー。
【請求項６７】
請求項６３に記載のライブラリーであって、該ライブラリーは、前記核酸分子を全て含む、ライブラリー。
【請求項６８】
２つ以上のスポットを含む核酸アレイであって、各スポットは、所定の位置にて基質に付着される請求項６３に記載の、複数の実質的に同一なライブラリーの単離された核酸分子

を含む、核酸アレイ。

【請求項６９】

請求項６８に記載のアレイであって、少なくとも１つの分子は、表３の非ＣａｍｂｒｉｄｇｅヒトｍｔＤＮＡヌクレオチド対立遺伝子からなる群より選択されるヌクレオチド対立遺伝子を含む配列を有する、ライブラリー。

【請求項７０】

請求項６８に記載のアレイであって、少なくとも１つの分子は、表４の非ＣａｍｂｒｉｄｇｅヒトｍｔＤＮＡヌクレオチド対立遺伝子からなる群より選択されるヌクレオチド対立遺伝子を含む配列を有する、ライブラリー。

【請求項７１】

請求項６８に記載のアレイであって、少なくとも１つの分子は、ヒトハプログループおよびマクロハプログループ（表１１）を診断するために有用なヌクレオチド対立遺伝子におけるヌクレオチド対立遺伝子からなる群より選択されるヌクレオチド対立遺伝子を含む配列を有する、アレイ。

【請求項７２】

請求項６８に記載のアレイであって、前記核酸分子の全てを含む、アレイ。

【請求項７３】

請求項６８に記載のアレイであって、スライドガラスにプリントされている、アレイ。

【請求項７４】

請求項６８に記載のアレイであって、約１０を超えるスポットを含む、アレイ。

【請求項７５】

請求項６８に記載のアレイであって、約２５を超えるスポットを含む、アレイ。

【請求項７６】

請求項６８に記載のアレイであって、前記単離された核酸分子は、長さが約２０ヌクレオチドである、アレイ。

【請求項７７】

核酸アレイを作製する方法であって、該方法は、以下の工程：

ａ）調製した基材を提供する工程；および

ｂ）該基材上に既知の位置にて２以上のスポットをプリントする工程であって、各スポットは、複数の実質的に同一な単離された核酸分子を含み、各分子は、約７～約３０ヌクレオチドの長さを有し、各分子は、配列番号１の一部分と同一な配列を含み、表３の第１列に列挙されたものからなる座の群より選択される座にて少なくとも１つのヌクレオチド対立遺伝子を含む、工程、

を包含する、方法。

【請求項７８】

請求項７７に記載の方法であって、少なくとも１つの分子は、表３の非ＣａｍｂｒｉｄｇｅヒトｍｔＤＮＡヌクレオチド対立遺伝子からなる群より選択されるヌクレオチド対立遺伝子を含む配列を有する、方法。

【請求項７９】

請求項７７の記載の方法であって、少なくとも１つの分子は、表４の非ＣａｍｂｒｉｄｇｅヒトｍｔＤＮＡヌクレオチド対立遺伝子からなる群より選択されるヌクレオチド対立遺伝子を含む配列を有する、方法。

【請求項８０】

請求項７７に記載の方法であって、前記アレイは、前記核酸分子の全てを含む、方法。

【請求項８１】

サンプル中のヌクレオチド対立遺伝子の存在または非存在を決定するための方法であって、該方法は、以下の工程：

ａ）調製したヒトサンプルを提供する工程；

ｂ）請求項６８に記載のアレイを提供する工程；

ｃ）定量的ハイブリダイゼーションを可能にする条件下で、該アレイと、該サンプルと

を接触させる工程；

ｄ）該アレイに対する該サンプルのハイブリダイゼーションパターンを測定する工程；ならびに

ｅ）該ハイブリダイゼーションを分析する工程、

を包含する、方法。

【発明の詳細な説明】

【技術分野】

【０００１】

（関連出願についてのクロス・リファレンス）

本出願は、米国特許出願出願番号2001年8月30日に出願の第60/316333号および出願番号2002年5月13日に出願の第60/380,546号に、そして、カナダの特許出願番号2,356（2001年8月31日に出願の第536号）まで優先権を請求するそれ、本願明細書において開示と一致していなくない範囲を参照することで、これによってそれらの全部に取り入れられる。

【０００２】

（連邦によって後援された調査に関する記載）

本発明は、米国政府（NIH研究費AG13154、HL64017、NS21328およびNS37167）から、資金提供を有する一部においてなされた。米国政府は、特定の権利をその中で有することができる。

【背景技術】

【０００３】

（発明の背景）

人間のミトコンドリアＤＮＡ（mtDNA）は、母らしく継承される。突然変異は、順番に人間の進化の木上の分岐を作成している血統を放射することにたまる。mtDNAのシーケンスを使用して、人間の人口がハプログループに進化的にに割り切れること（ウォレス（D.C.ほか（1999の）Gene 238）：211-230 ;イングマンM.ほか（（2000の）ネイチャー408）：708- 713;Maca-マイヤー、N.（2001年8月）は本部2をBioMedした：13;T.G.シューアほか（Physical Anthropology 108アメリカの（1999の）ジャーナル）：1-39;そして、V.マコーレーほか（HumanGe7letcs 64のアメリカの（1999の）Jour7lal）：232-249).関連したハプログループは、マクロハプログループに結合されることができる。ハプログループは、subハプログループに再分割されることができる。完全なケンブリッジ・ミトコンドリアＤＮＡシーケンスは、MITOMAP、http://www.将軍emory.edu/cgi-ジン/MITOMAP、Genbank継承no. J01415で見つかることができて、SEQ ID NOにおいて提供される：2.また、アンドルーズその他（1999）を参照、「HumanMitochondrial DNA（自然Genetics 23）のためのケンブリッジReference Sequenceの再分析およびRevision：147。

【０００４】

ミトコンドリア生物学に関する刊行物は、以下から成る：シェッフラー、I. E.（1999の）Mitochondria、ワイリー-リース、NY;レスティエンヌP Ed.;ミトコンドリアDiseases：モデル、そして、Methods（シュプリンガー出版）ベルリン;Enzymology（2000）の方法322：断面VMitochondria、そして、Apoptosis（アカデミック・プレス）CA;ミトコンドリアおよびCell Death（1999の）プリンストン大学プレス（NJ）;パパS、FerruciioGおよびTager J Eds.;移動電話Bioenergeticsの境界：分子のBiology、Biochemistry、そして、Physiopatholo（KluwerAcademicなPlenum Publishers）NY;Lemasters、J.、そして、よりKluwerにAcademicなニエミネン（PathogenesisのA.（2001の）Mitochondria）/PlenumPublishers（NY）;MITOMAP（http://www.将軍emory. edu/cgi-ジン/MITOMAP）;ウォレス。D.Cに、（2001の）「変性疾患およびageing"Novartis財団Symposium235のためのミトコンドリア方法：247- 266;ウォレス（D. C.（1997））「Agingする際のミトコンドリアＤＮＡ、そして、Disease"Scientificアメリカの8月277：40-47;ウォレス（D.C.その他）（1998）「ミトコンドリア生物学、変性疾患および老化（BioFactors 7）：187-190;Heddi（A.その他）（1999）「MitochondrialDisease Patients"JBC 274のTissuesのEnergy Gene ExpressionのInductionを調整する：22968-22976 ;ウォレス。D. Cに、（1999の）「Manお

およびMouse"Science 283のミトコンドリアDiseases：1482-1488;サラステ（M.（1999））「フィンで酸性のPhosphorylationデsiecle"Science 283：1488-1493;Kokoszka et.アル。（2001の）「マウスが達しているミトコンドリア機能の年齢関連の低下に結果としてなるSod2（+/-）の増加するミトコンドリア酸性応力は、apoptosis"PNAS98を増やした：2278-2283;ウォレス、D. C.（2001の）Mental RetardationおよびDevelopmentalDisabilities 7：158-166 ;ウォレス。D. Cに、（2001の）Am。J.は、Medした。将軍106：71-93;lournalな（台北）Wei（YHほか中国の（2001の）Medical）64：259-270;そして、ウォレス（D.C.(2001) EuroMit 5つのアブストラクト）。

【０００５】

特定のミトコンドリア突然変異は、生理的状況を伴った（米国特許6,280、966は、2001年8月28日に出た;米国特許6,140（2000年10月31日上の067発行日）;米国特許5,670、320;米国特許5,296、349;米国特許5、185、244;米国特許5,494、794;ウォレス（D.C.（1999の）サイエンス283）：1482-1488 ;ブラウン（Human Genetics Poster # 2332のためのM.D.ほかアメリカの（2001の）協会）;ブラウン、M. D.ほか、（2001の）Humanジュネ。109:33-39;そして、ブラウン（M. D.ほか（2002年1月）Humanジュネ）。110:130-138）（ウォレス）D.C.ほか（1999の）Gene 238：211-230は、LHON mutants.Grossman、L. I.ほか（2001の）MolecularPhylogeneticsおよびEvolution18（1）の分析を記載する：26-36は、変化を酸素のエネルギー代謝のための生化学機械に記載する。カールマン（B.その他.(1999) ActaNeurol.）乏しい。99の(1)：16-25は、ミトコンドリア突然変異および多発性硬化症（MS）を記載する。Wei、Y. H.ほか中国の（2001の）Medicalジャーナル64：259-270は、老化のミトコンドリア理論を支持して最近の結果を記載する。

【０００６】

イワノーワ（R.ほか（1998の）Geronotology44）：349は、記載するミトコンドリア・ハプロタイプ、そして、フランスの人口の長命。タナカ（M.ほか（1998の）Lancet 351）：185-186は、長命およびハプログループを日本の住民に記載する。デBenedicts（G.ほか（1999の）FASEB13）：1532-1536は、ハプログループおよび長命をイタリアの個体群に記載する。バラ、Human Geraetics 9 G.その他.(2001)ヨーロッパ人ジャーナル：701-707は、ハプログループJを百歳以上の人に記載する。ロス（O. A.ほか（2001の）Experimental Gerontology 36（7））：1161-1178は、記載するハプロタイプ、そして、アイルランドの個体群の長命。

【０００７】

ハプログループTは、ヨーロッパの男性の減少する精子運動性を伴った（E.Ruiz-Pesiniほか（Human Genetics 67アメリカの[2000の]ジャーナル）：682-696）、ハプログループHのtRNA'"np 4336の異型は、関係している遅発性アルツハイマーDisease（J. M.Shoffnerほか（[1993の]Genomics 17）：171-184)。

【０００８】

テイラー（生物エネルギー論および生体膜29（2）のR. W.（1997の）J.）：195-205は、ミトコンドリア疾患を治療する方法を記載する。Collombet（J.およびクテル）C.（1998の）MolecularMedicineツデー4 (1)：1-8はミトコンドリア障害のための遺伝子治療を記載する。そして、健康なミトコンドリアを導くために細胞融合を使用することを含む。オーエン、R.およびFlotte（T.R.（2001の）AntioxidantsおよびRedox Signalingしている3つの(3)）：451-460は、方法およびミトコンドリア疾患のための遺伝子治療への制限を議論する。

【０００９】

特定の疾患を伴ったそれを除いて、人間のミトコンドリアＤＮＡシーケンス・バリエーションは、特定の表現型状況を伴わなくて、中性であるとみなされて、人間の系統発生を再建するために用いた（ヘンリー・ジー、「アフリカのエデンの上の統計Cloud」、（1992年2月13日）ネイチャー355：583;マーシャBarinaga、「『アフリカのイブBackers BeatRetreat」（（1992年2月7日）サイエンス）255：687;S.ブレア・ヘッジほか（Mitochondrial DNA Sequencesの人間のOriginsおよびAnalysis）、「（1992年2月7日）サイエンス（2

55）：737-739;アラン・ウィルソン、そして、レベッカ・キャン（「HumansのRecentアフリカのGenesis」）（1992年4月）サイエンティフィックアメリカン、68）。2つの人間のミトコンドリア・ゲノムの塩基対違いの平均数は、9.5から66まであるために推定される（ZevianiM.その他（1998）「分子の医学の再調査：ミトコンドリア障害、「医学77：59-72)。

【００１０】

Ｄループ、この環状線の中のサイトが集中するミトコンドリア・ゲノムおよび最も多形性ヌクレオチドの最も多くの可変領域は、two'hypervariable部分のもの、HVS-IおよびHVS-IIである（ウィルキンソン-Herbots（H. M.その他）（1996）「中でサイト73超可変部人間のミトコンドリア・ゲノムおよびヨーロッパの個体群の起源のlI」（アンHumジュネ60）：499-508).人口に特有の、中性のmtDNA異型はmtDNA制限部位異型を調査することによって識別された、または、塩基配列決定によって、置換のhypervariableなセグメントはループする。14の規制エンドヌクレアーゼを使用している規制分析は、バリエーションのためのmtDNAシーケンスの15-20%のふるい分けを許した（チェンY.S.その他、（1995の）「アフリカの人口におけるmtDNA変化の分析で、全ての人間の大陸に特有のハプログループで古い最大限がJHumジュネ57であることが分かる：133-149).現在まで発表されるmtDNAシーケンス・データのかなりの大多数は、HVS-Iに限られている。Bandelt（H.J.その他）（1995）「正中のnetworks"Genetics 141を使用している人間の住民のミトコンドリア・ポートレイト：743-753)。

【００１１】

mtDNA ハプログループのために使われた符号化および分類法システムは、主にRFLPsによって提供される情報および制御領域のhypervariableな部分に関連する。（Torroni、A.その他（1996）「3人のヨーロッパの住民（遺伝学144）の分析からのヨーロッパのmtDNAsの分類法：1835-1850、そして、リチャーズMBほか、（1998の）「西ヨーロッパ（アンHumジュネ62）のミトコンドリアＤＮＡのPhylogeography：241-260.)。

【００１２】

方法は、突然変異の中性の可能性を試験することに知られている（タジマ（F.(1989) Genetics 123）：585-595 ;Fu、Y.およびLi（W.（1993の）Genetics 133）：693-709;Li（W.その他.(1985)Mol.）生物学Evol.2 (2) :150-174;そして、Nei、M.およびゴジョウボリ（T.（1986の）Mol. Biol. Evol.3 (5)）：418-426).これらの刊行物における方法の全ては、別々のグループからとられるデータセットを比較するために用いる。これらの方法のいずれも、outgroupを表しているデータを含んでいないデータセットを分析するために用いない。

【００１３】

ワイズ（C.A.ほか（1998の）Genetics 148）：409-421は、チンパンジからのNADH Dehydrogenase Subunit 2つの遺伝子と比較して人間のミトコンドリアNADHDehydrogenase Subunit 2つの遺伝子の中性分析を記載する。テンプルトン（A. R.（1996の）Genetics 144）：ヒト科の霊長類のCOXII遺伝子と比較して1263-1270は人間のミトコンドリアCytochromeOxidase II（COXII）遺伝子の中性分析を記載する。メシエ、W.およびスチュワート（C.（1997の）ネイチャー385）：151-154は、霊長類リゾチームの中性分析を記載する。エンドー（T.ほか（1996の）Mol.Biol. Evol. 13（5））：685-690は、シーケンスの大規模な中性分析を記載するDDBJ、EMBL、そして、GenBankデータベース。ヒューズ、A.L.およびNei（M.（1988の）ネイチャー335）：167-170は、MHCクラスの中性分析を記載する私loci.Nachman（M. W.（1996の）Genetics142）：953-963は、チンパンジからのNADH DehydrogenaseサブユニットΓ3つの遺伝子と比較して人間のミトコンドリアNADHDehydrogenaseサブユニット3つの（NADH3）遺伝子の中性分析を記載する。Nachman、M. W.ほか（1994の）会報全米科学アカデミーUSA76：5269-5273は、マウスの3つの種のミトコンドリアNADHデヒドロゲナーゼ・サブユニット3つの遺伝子の中性分析を記載する。ランド（D. M.ほか（1994の）Genetics138）：741-756;バラード、J. W. O.およびKreitman（M.（1994の）Genetics 138）：757-772;そして、カネコ（M.Y.ほか（1993の）ジュネ）。物。61：195-204

、ミトコンドリアNADHデヒドロゲナーゼ・サブユニット5のための中性分析、Cytochrome bおよびATPase6をショウジョウバエの種に記載する。

【００１４】

上述の刊行物において、Ka/Ks分析を含んで、中性テストは、疾患関連の突然変異を識別するために適用されなかった。分析を検証している中性のための人口は、通常の表現型バリエーションの診察によって識別された。中性テストは、遺伝子が選択の下にあるかどうか決定するために実行された。これらの刊行物の誰もphenotype-関連する突然変異を識別する目的で中性分析を解説しない、そして、疑われた表現型関連の突然変異は識別されなかった。

【００１５】

米国特許6,228、586（2001年5月8日、出た）および米国特許6、280、953（2001年8月28日、出た）はポリヌクレオチドを識別する方法およびポリペプチド・シーケンスを人間のおよび／または人間外の霊長類に記載する。そして、それは生理的状態を伴ってもよい。方法は、統計方法を使用している人間で人間外の霊長類シーケンスの比較を使用する。米国特許6,274、319（2001年8月14日、出た）は、ポリヌクレオチドを識別するKa/Ks方法および採り入れられた工場または動物の商業上または美学的に関連した特徴を伴ってもよいポリペプチド・シーケンスを記載する。方法は、進化的にな重要な変化を識別するために採り入れられた有機体およびその野生の祖先から相応する遺伝子の比較を使用する。上述の刊行物において、Ka/Ks分析を含んで、異種間、同一種内でない中性テストがあてはまられるだけであること（比較）、そして、細胞小器官ゲノム以外から、核ゲノムからの遺伝子だけは、分析される。

【００１６】

ペプチドおよびヌクレオチド図書館を建設することのための方法は、技術にとって周知であるｅ。g.米国特許6,156、511および6,130（092）にて説明したように。塩基配列決定方法は、また、技術にとって公知であるｅ。g.米国特許6,087（095）にて説明したように。核酸の配列が、塩基配列決定のために、そして、疾患関連の対立遺伝子を含んでいる例外的な対立遺伝子を識別するために使われた。核酸配列は、記載されていた。eに、g.特許no.で：米国5、837、832、U.S.5,807、522、米国6,007、987、米国6,110、426、WO 99/05324、99/05591、WO 00/58516、WO95/11995、WO 95/35505A1、WO 99/42813（JP10503841T2、GR3030430T3、ES2134481T3、EP804731B1、DE69509925C0、CA2192095AA、AU2862995A1、AU709276B2、AT180570、EP1066506およびAU 2780499）。計算方法は、雑種形成結果を分析することに役立つｅ。g.特許協力条約Publication WO 99/05574および米国特許5,754、524にて説明したように;6228,575;5,593, 839;そして、5,856、101。疾患標識のふるい分けのための方法は、また、技術にとって公知であるｅ。米国特許6,228、586にて説明したように、g.;6,160,104;6,083, 698;6,268, 398 ;6,228, 578 ;そして、6,265、174。

【００１７】

マイクロアレイ技術の開発は、遺伝子突然変異、遺伝子表現または核酸シーケンスに関する情報を得る努力で、同時に核酸調査シーケンスの非常にかなりの数を調べたいという願望から生じた。マイクロアレイ技術は、個人的にヒトゲノム解析計画と関係がある。そして、それは鍵となる目的として核酸塩基配列決定およびゲノム分析の急速方法の発現を有する‖（E.マーシャル（（1995の）サイエンス268）：1270）、sequence-機能関係の解明として同様に（M. Schenaその他、（1996の）会報国立。アカデミ科学USA（93）：10614).マイクロアレイ雑種形成のオリゴヌクレオチド（ASO）が徹底調査する対立遺伝子-特性に対する断片をPCR-amplifiedしたgenotypingしている大規模な一つのヌクレオチド多形性（SNP）において広く使われている（フーバーM.ほか（2002の）AnalyticalBiochemistsy 303：25-33、そして、南部、E. M.(1996) Tr ndsジュネ。12:110-115)。

【００１８】

AffymetrixGeneChip HuSNP Arrayは同時に追跡ほぼ1,500の遺伝子のバリエーションによって、完全なゲノム調査を可能にする。そして、一つのヌクレオチド多形性（SNPs）がゲノムの全体にわたって分散したので、公知である。HuSNPAffymetrix Arrayが、遺伝す

る疾患をマップすることを意図する家族性の結合研究または薬のために使われている。新たに遺伝子の変更を追うことに関しては同様に感染性。genotypingするために、配列はシーケンスの個々のヌクレオチドに問い合わせるために多数の調査に依存する。目標ベースのアイデンティティは目標位置だけにおいて変化する4台の同一の探測機を使用して演繹されることができる。そして、各々4つの可能なベースのうちの1つを含む。あるいは、共通配列の存在は、特定の対立遺伝子を表している１、２の調査を使用して試験されることができる。遺伝子型にヘテロである、または、遺伝的にサンプルを調合されて、多くの調査を有する配列は、冗長な情報を提供するために作成されることができる。

【００１９】

配列（また、DNAマイクロアレイsまたはDNAチップと呼ばれている）は、ガラスに一般に高速ロボット工学によって製造される‖時々ナイロン基板上の、しかし‖どの調査のための（Phimister、B.（1999の）ネイチャーGenetics21s：1-60）、相補型締め具は、周知の識別によって決定するために用いる。単一のDNAチップの実験は、同時に何千もの遺伝子に関する研究者情報を提供することができる。設計におけるいくつかのステップおよびDNA配列実験の実行がある。多くの戦略は、各々のこれらのステップで調査された：1）DNAタイプ;2）Chip製作;3）Sample準備;4）Assay;5）Readout;そして、6）Software（情報科学）。

【００２０】

配列技術のための2枚の主な申込み用紙が、ある：1）遺伝子の表現レベル（多量）のDetermination;そして、2）シーケンス（遺伝子/遺伝子突然変異）のIdentification。そこで、周知の識別を有する配列されたDNAシーケンスの知的所有権に関して、配列技術の2つの異型であるように見える：私が成るフォーマットは、しみがついているガラス使用しているロボットのような固体の表層に固定されて、一組の目標にもさらされるcDNA（500-5（000ベース）長く）を徹底調査する別に、または、混合の。DNAマイクロアレイと呼ばれている伝統的に、この方法」は、広くスタンフォード大学で開発されたことと思われる。（R.エキンズおよびF. W. Chu"マイクロアレイs：それらの起源およびアプリケーション」、Biotechnologyの[1999の] Trends、17：217-218).フォーマットRはオリゴヌクレオチド（20-80-mer oligos）の配列から成る、または、調査がsitu（オンチップの）において従来の合成によって総合したペプチド核酸（PNA）はオンチップ固定することによってあとに続いた。配列はラベルをつけられたサンプルDNA（雑種を作られる）にさらされる、そして、補完的なシーケンスの識別/多量は決定される。この方法、「historically"calledDNAチップはAffymetrix社で開発された。そして、それはGeneChip商標の下でその写真蝕刻法で製作された製品を販売する。多くの会社は、他の元の位置の合成を使用しているかまたは技術をdepositioningしているオリゴヌクレオチドに基づくチップを製造している。

【００２１】

配列上の調査はfluorescentlyにラベルをつけられた目標ポリヌクレオチドによって雑種を作られることができる、そして、交雑する配列は蛍光顕微鏡検査を走査することによって走査されることができる。不適当な組合せ内容の範囲を決定して、多形性を識別して、一般の若干の塩基配列決定情報を提供するアルゴリズムによって、蛍光パターンは、それから分析される（M.チーほか（[1996の]サイエンス274）：610).選択性は、離れて非選択的に吸着された材料をすすぐために低い厳しさ洗浄によってこのシステムで産出される。配列要素からの相対的な結合する信号の次の分析は、塩基対不適当な組合せが存在することができるところを決定する。この方法はそれから厳しさを最大にするために従来の化学方法に依存する、そして、オートメーション化したパターン認識処理は完全に相補型で部分的に相補型締め具を区別するために用いる。

【００２２】

装置（例えば標準の核酸マイクロアレイsまたは遺伝子チップ）は、サンプル冗長のデータ処理アルゴリズムおよび用途を必要とする（i. e。同じ種類の変則の統計学的に有意のデータ解釈および回避のための配列要素の多数）提供する半定量分析装置面に固定され

る目標シーケンスおよびシーケンス間の不適当な組合せの多形性またはレベルの。

【００２３】

配列分析に適当なラベルは、公知技術である。実施例は、two-色蛍光システム（例えばCy3/Cy5およびCy3.5/Cy5）である。5ホスホラミダイト（グレンResearch（スターリング-ヴァージニア））。シアニン染料をカバーしている特許は、以下から成る：米国6,114、350（Sept.5（2000））;米国6,197、956（2001年3月6日）;米国6,204、389（2001年3月20日）およびU.S.6,224（644（2001年5月1日））。配列プリンタおよび読者は、従来技術において利用できる。

【００２４】

配列を使用する方法は、Grigorenko、E. V.編集、（2002の）DNA Arraysに記載されている：技術およびExperimentalStrategies、CRCプレス、NY;ブラーナ、K. E.ほか、（2001年5月）マイクロアレイsおよびRelatedされたTechnologies：GenomicsResearch、CHI、Upperフォールズ、MAの小型化およびAcceleration;そして、ブランカ、M. A.ほか、（2002年2月）DNA マイクロアレイInformatics：鍵となるTechnological TrendsおよびCommercial Opportunities、CHI、Upperフォールズ、MA。

【００２５】

本願明細書において参照される全ての刊行物は、これと共に矛盾していない範囲を参照することで組み込まれる。このBackground Sectionの刊行の記載は、それが従来技術であるという承認を構成しない。

【発明の開示】

【課題を解決するための手段】

【００２６】

（発明の要旨）

人間のミトコンドリアＤＮＡの高ミトコンドリアＤＮＡ突然変異速度は、mtDNAの広範囲にわたる中立不偏の、人口に特有のベース置換の蓄積に結果としてなると考えられた。これらは、人間の個体群が世界の異なる地理的領域にコロニーを作ったにつれて、ほぼ同じ時間的尺度に関して異なった母性血統を放射することに沿って順番に蓄積した。

【００２７】

アフリカのmtDNAsがハプログループLに落して、bpでHpaI制限部位増加によって定めた全ての76%について、3592.77%のアジアのmtDNAsは、bp 10397でbp 10394およびAluIサイト増加でDdeIサイト増加によって定義されるsuper-ハプログループの範囲内で含まれる。本質的に、全ての自国のアメリカのmtDNAsは、4つのハプログループ（AD）に落ちる。ハプログループAは、bp 663のHaeinサイト増加、bp 8281に対するbp 8271間の9つのbp削除によるB、bp 13259のHincIIサイト損失によるCおよびbp5176でAluIサイト損失によって定義されるDによって定義される。10のハプログループは、ヨーロッパの人口のほとんどすべてのmtDNAsを含む。ヨーロッパ人の10-mtDNAハプログループは、hypervariableな部分の符号化領域および塩基配列決定のRFLP分析から、データの組合せを用いて調査を受けることができるI. 99%のヨーロッパのmtDNAsが10のハプログループのうちの1つに落すAbout：H（私）J、K、M、T、U、V、W‖またはX.‖

本発明は、全ての主要人間のハプログループの診断およびそれらのハプログループを診断する方法であって、sub-ハプログループを選んだ人間のmtDNA多形性を提供する。

【００２８】

本発明も、方法を進化的にな重要なミトコンドリアＤＮＡ遺伝子、ヌクレオチド対立遺伝子およびアミノ酸対立遺伝子を識別することに提供する。進化的にな重要な遺伝子および対立遺伝子は、単一の種の１人か２人の住民を使用して識別される。進化的にな重要なヌクレオチド対立遺伝子を識別する方法はそれらの遺伝子の進化的にな重要な遺伝子およびそれから進化的にな重要なヌクレオチド対立遺伝子を識別することを含む、そして、進化的にな重要なアミノ酸対立遺伝子を識別することは全てのnonsynonymousな対立遺伝子によってコード化されるアミノ酸を識別することが必要である。同義の符号化、対立遺伝子が本願明細書において開示して、含まれる進化的にな重要なアミノ酸に対する等価物は

、本発明の進化的にな重要なアミノ酸対立遺伝子をコード化しているヌクレオチド対立遺伝子の中である‖本発明。同義の符号化同じコドンの範囲内である隣接したヌクレオチド場所の対立遺伝子を含む。

【００２９】

本発明も、方法をハプログループおよび進化的にな重要なヌクレオチドおよびアミノ酸対立遺伝子を生理的状況に対する素因と関連させることに提供する。LHONに対する素因を診断する方法および盲目を呈する増加する可能性、centenariaおよび診断されている個人の地理的位置に、依存していない増加する長命を診断する方法は、本願明細書において提供される。エネルギー代謝-関連した生理的状態に対する素因をもつ個人の診断は、個人の地理的領域に依存している。

【００３０】

本発明の方法によって診断可能な生理的状況は、健康な状況および病的状態を含む。ハプログループによって、そして、本発明によって提供される対立遺伝子によって関連する生理的状況は、精力的なアンバランス、代謝疾患、異常なエネルギー代謝、異常な温度規則、異常な酸化的リン酸化、異常な電子伝達、肥満、体脂肪、糖尿病、高血圧および心血管疾患の量を含む。

【００３１】

本発明によって提供されるシーケンスを有する分子は、ライブラリで、そして、配列をgenotypingすることで提供される。本発明は、本発明のgenotypingしている配列を製作し使用する方法を提供する。本発明の配列は、存在および、ハプログループを決定するために、本発明の、そして、診断のためのヌクレオチド対立遺伝子の欠如を決定することに役立つ。

【００３２】

本発明も、機械可読の記憶装置およびプログラム装置をデータおよびプログラムされた方法をハプログループおよび生理的状況を診断することに備えてたくわえることに提供する。

【００３３】

本発明の配列は、存在および、ハプログループを決定するために、本発明の、そして、診断のためのヌクレオチド対立遺伝子の欠如を決定することに役立つ。本発明も、機械可読の記憶装置およびプログラム装置をデータおよびプログラムされた方法をハプログループおよび生理的状況を診断することに備えてたくわえることに提供する。

【００３４】

（発明の詳細な説明）

表1は人間のミトコンドリア・ヌクレオチド対立遺伝子を示す。そして、それは生理的状況を伴った。表1において、柱3（ヌクレオチド場所）、5（生理的状態ヌクレオチド対立遺伝子）および柱2（生理的状態）は、HumanMitochondrial Nucleotide Alleles Knownの一組がPhysiological ConditionsによってAssociatedされると思いこむ。

【００３５】

【表１】

TABLE 1[1]

Human Mitochondrial Alleles Known to be Associated with Physiological Conditions

Gene	Physiological Condition	Nucleotide Locus	Cambridge Nucleotide Allele	Physiological Condition Nucleotide Allele	Cambridge Amino Acid Allele	Physiological Condition Amino Acid Allele
MTND1	*MELAS	3308	T	C	M	T
MTND1	*NIDDM; LHON; PEO	3316	G	A	A	T
MTND1	*LHON	3394	T	C	Y	H
MTND1	*NIDDM	3394	T	C	Y	H
MTND1	*ADPD	3397	A	G	M	V
MTND1	*LHON	3460	G	A	A	T
MTND1	*LHON	3496	G	T	A	S
MTND1	*LHON	3497	C	T	A	V
MTND1	*LHON	4136	A	G	Y	C
MTND1	*LHON	4160	T	C	L	P
MTND1	*LHON	4216	T	C	Y	H
MTND2	*LHON	4917	A	G	D	N
MTND2	*LHON	5244	G	A	G	S
MTND2	*AD	5460	G	A	A	T
MTND2	*AD	5460	G	T	A	S
MTCO1	*Myoglobinuria, Exercise Intolerance	5920	G	A	W	Ter
MTCO1	*Multisystem Disorder	6930	G	A	G	Ter
MTCO1	*LHON	7444	G	A	Ter	K
MTCO2	*Mitochondrial Encephalomyopathy	7587	T	C	M	T
MTCO2	*MM	7671	T	A	M	K
MTCO2	*Multisystem Disorder	7896	G	A	W	Ter
MTCO2	*Lactic Acidosis	8042	AT	2 nt del (AT)	M	Ter
MTATP6	*NARP	8993	T	G	L	R
MTATP6	*NARP / Leigh Disease	8993	T	C	L	P
MTATP6	*LHON	9101	T	C	I	T
MTATP6	*FBSN / Leigh Disease	9176	T	C	L	P
MTATP6	*Leigh Disease	9176	T	G	L	R
MTCO3	*LHON	9438	G	A	G	S
MTCO3	*Leigh-like	9537	C	C ins	Q	frameshift
MTCO3	*LHON	9738	G	T	A	S
MTCO3	*LHON	9804	G	A	A	T
MTCO3	*Mitochondrial Encephalopathy	9952	G	A	W	Ter
MTCO3	*PEM; MELAS	9957	T	C	F	L
MTND3	*ESOC	10191	T	C	S	P
MTND4	*MELAS	11084	A	G	T	A
MTND4	*LHON	11778	G	A	R	H
MTND4	*Exercise Intolerance	11832	G	A	W	Ter
MTND4	*DM	12026	A	G	I	V
MTND5	*MELAS	13513	G	A	D	N

[1] (MITOMAP: A Human Mitochondrial Genome Database. Center for Molecular Medicine, Emory University, Atlanta, GA, USA. http://www.gen.emory.edu/mitomap.html, 2001).

MTND5	*MELAS	13514	A	G	D	G
MTND5	*LHON-like	13528	A	G	T	A
MTND5	*LHON	13708	G	A	A	T
MTND5	*LHON	13730	G	A	G	E
MTND6	*MELAS	14453	G	A	A	V
MTND6	*LDYT	14459	G	A	A	V
MTND6	*LHON	14484	T	C	M	V
MTND6	*LHON	14495	A	G	L	S
MTND6	*LHON	14568	C	T	G	S
MTCYB	*PD / MELAS	14787	TTAA	4 nt del (TTAA)	I	frameshift
MTCYB	*MM	15059	G	A	G	Ter
MTCYB	*Exercise Intolerance	15150	G	A	W	Ter
MTCYB	*Exercise Intolerance	15197	T	C	S	P
MTCYB	*Mitochondrial Encephalomyopathy	15242	G	A	G	Ter
MTCYB	*LHON	15257	G	A	D	N
MTCYB	*Exercise Intolerance	15615	G	A	G	D
MTCYB	*MM	15762	G	A	G	E
MTCYB	*LHON	15812	G	A	V	M

＊定義：

LHONレーバー遺伝視覚神経障害

mmミトコンドリア筋疾患

ADアルツハイマー病

LIMM致死小児ミトコンドリア筋疾患

ADPDアルツハイマー病およびパーキンソン病

MMC母性筋疾患および心筋症

NARP Neurogenicな筋肉虚弱、AtaxiaおよびRetinitisPigmentosa;MELAS関連の心筋症MELAS Mitochondrial Encephalomyopathy、Lactic AcidosisおよびMyoclonicEpilepsyおよびRaggedされた共産主義者Muscle Fibers MHCMがMaternallyするStrokeのようなエピソードLDYTレーバーの遺伝視覚神経障害およびDYsToniaMERRFが継承したリーDisease FICP Fatal Infantile Cardiomyopathy Plusとして、この場所の交互の表現型は、報告されるHypertrophicなCardioMyopathyCPEO Chronicな進歩党員External Ophthalmoplegia KSSカーンズ・セアSyndromeDM筋疾患およびOphthalmoplegiaDEAFを有するDiabetesメリトゥスDMDF Diabetesメリトゥス+ DeaFness CIPO Chronic IntestinalPseudoobstructionが、MaternallyにDEAFnessまたはアミノグリコシドによって誘発されたDEAFness PEM進歩党員脳障害SNHLSensoriNeural Hearing Lossを継承した

13のタンパク質-符号化ミトコンドリア遺伝子は、公知である（MitoMap、http：//www.将軍emory.edu/cgi-bin/MITOMAP）。

【００３６】

【表２】

Table 2
Protein-coding Human MtDNA Genes

Gene	Map Locus[a]	Abbreviation	Location[b]
NADH dehydrogenase 1	MTND1	ND1	3307-4262
NADH dehydrogenase 2	MTND2	ND2	4470-5511
NADH dehydrogenase 3	MTND3	ND3	10059-10404
NADH dehydrogenase 4L	MTND4L	ND4L	10470-10766
NADH dehydrogenase 4	MTND4	ND4	10760-12137
NADH dehydrogenase 5	MTND5	ND5	12337-14148
NADH dehydrogenase 6	MTND6	ND6	14149-14673
Cytochrome b	MTCYB	Cytb	14747-15887
Cytochrome c oxidase I	MTCO1	COI	5904-7445
Cytochrome c oxidase II	MTCO2	COII	7586-8269
Cytochrome c oxidase III	MTCO3	COIII	9207-9990
ATP synthase 6	MTATP6	ATP6	8527-9207
ATP synthase 8	MTATP8	ATP8	8366-8572

[a,b] As defined on MitoMap, http://www.gen.emory.edu/cgi-bin/MITOMAP, which is numbered relative to the Cambridge Sequence (Genbank accession no. J01415 and Andrews et al. (1999), A Reanalysis and Revision of the Cambridge Reference Sequence for Human Mitochondrial DNA, Nature Genetics 23:147.

mtDNAのためのコドン使用は、汎用コードと僅かに異なる。実施例、終了の代わりにトリプトファンのためのUGAコード、イソロイシンの代わりにメチオニンのためのAUAコードおよびAGAのための、そして、AGGはアルギニンのための符号化の代わりにターミネータである。

【００３７】

固体基板の周知の位置上の核酸の配列を作成する方法に対する使用されたherein"printing"refersとして。本発明の配列は、観測することによって印刷されることができるｅ。g.固体基板にまたは調査の統合に調査の配列を中で適用することは、固体基板に配置する。同じ寸法のガラスの小さい部分に対する標準の顕微鏡と同程度使用されたherein"glass slide"refersは、摺動する。ここで使用しているように、能力がある物質によって備えができている基板に、調査を基板に取り付けることの配属媒体として役立って、substrate"refersをprepaxedされて‖例えばポリエステル繊維Lysine.ここで使用しているように、「genotyped.でありえる人間のミトコンドリアＤＮＡを含んでいる組成物に対するsample"refersここで使用しているように、1つのヌクレオチド対立遺伝子（単一のヌクレオチド多形性）のシーケンスが決定されることができるように、雑種形成に対する量的hybridization"refersは適当な状況および使用している適当な材料の下で実行した、例えば分子の雑種形成によって、2以上までその対立遺伝子を含むことは深く探る。そして、全て周知のようにそのヌクレオチド場所で各々異なる対立遺伝子を含んだ。

【００３８】

ここで使用しているように、「生理的条件」"includesは、状況、健康な状況および化粧状況を病気にかからせた。病気にかかった状況は、代謝疾患（例えば糖尿病、高血圧および心血管疾患）を含むが、これに限定されるものではない。健康な状況は、増加する長命のような特徴を含むが、これに限定されるものではない。生理的状況は、化粧状況を含

む。化粧状況は、体脂肪の量のような特徴を含むが、これに限定されるものではない。生理的状況は、異なるコンテキスト（異なる環境の同じ有機体に関してのそのようなもの）の健康状況を変えることができる。人間のためのこの種の異なる環境は異なる文化的な環境である、または、異なる気象的なコンテキストは例えば異なる大陸で発見される。

【００３９】

ここで使用しているように、「一つ以上のヌクレオチド対立遺伝子の中性を決定する分析および／またはシーケンスの少なくとも2つの対立遺伝子を使用している対立遺伝子（s）を含んでいる遺伝子に対する中性analysis"refers。共通に、分析されるシーケンスの対立遺伝子は、2つのグループ（同義でnonsynonymousな）に分けられる。コドンがいずれをコード化するかについて示しているコドン使用テーブル、どのアミノ酸が、この分析において使われる。多くの有機体およびゲノムのためのコドン使用テーブルは、従来技術において利用できる。遺伝子が決定される場合、中性でない、遺伝子は進化の間、それに印加される選択圧を有して、進化的にに重要であるために決定される。遺伝子（nonsynonymousな）のアミノ酸を変える対立遺伝子は、それから、非中性で進化的にに重要であると決定される。

【００４０】

ここで使用しているように‖DNAシーケンス分析の技術にとって公知にさているように同義の違いの割合に対するnonsynonymousな違いの割合の比率に対する魂/Ks"refers。nonsynonymousな違いの割合は、nonsynonymousな置換が起こることができたサイトにつきシーケンスのnonsynonymousなヌクレオチド置換の数である。同義の違いの割合は、同義の置換が起こることができたサイトにつきシーケンスの同義のヌクレオチド置換の数である。あるいは、各々の割合の分母のサイトの数を含むだけである代わりに、各々のサイトで起こることができた他の置換でより麻痺したものは、また、含まれる。類似した定義が分析のKaおよびKsのために使われる限り、いずれの定義も使われることができる。kcKa/Ks.ある。

【００４１】

使われるように、中で結果としてなる突然変異に対するherein"nonsynonymous"refersはコード化されたアミノ酸に変化する。ここで使用しているように、「中で結果としてならない突然変異に対するsynonymous"refersは、コード化されたアミノ酸に変化する。

【００４２】

ここで使用しているように、「人間の進化の木上の周知のように血統を放射することへのハプログループ"refers。ここで使用しているように、「一群の進化的にな関連したハプログループ.に対するマクロハプログループ"refersここで使用しているように、進化的にに対する副ハプログループ"refersは、ハプログループのサブセットを関連づけた。個人のものハプロタイプ彼が帰属するハプログループにある。

【００４３】

ここで使用しているように、「平均より長く生きることへの拡張longevity"or"extended lifespan"refersは、1が帰属する人口のための寿命を予想した。ここで使用しているように、「少なくとも100年である拡張寿命に対するcentenaria"refers。

【００４４】

ここで使用しているように、「彼が生きる地理的領域に、非本来である個人が個人が生きるところ固有である人口のそれと異なるエネルギー代謝にゆだねる異常なエネルギーmetabolism"in。ここで使用しているように、この種の個人が彼が生きるところ固有である人口のそれと異なる温度規制にゆだねる異常な温度regulation"in。ここで使用しているように、この種の個人が彼が生きるところ固有である人口のそれと異なる酸化的リン酸化にゆだねる異常な酸性phosphorylation"in。ここで使用しているように、この種の個人が彼が生きるところ固有である人口のそれと異なる電子伝達にゆだねる異常な電子transport"in。使用されたherein"metabolicdisease"ofとして、この種の個人は、彼が生きるところ固有である人口のそれと異なる代謝に関連する。ここで使用しているように、「この種の個人がエネルギー生成のバランスにゆだねる精力的なimbalance"ofまたは彼が生きると

ころ固有である人口のそれと異なる使用。ここで使用しているように、「この種の個人が、個人の身長のために、20%、人口土地の人によいと個人が生きるところに勧められる平均体重より高い体重にゆだねるobesity"of。ここで使用しているように、「彼が生きるところ固有である人口に推薦されることと関連して、この種の個人が体脂肪の低いか高いパーセンテージにゆだねる体fat"ofの量。

【００４５】

ここで使用しているように、単離された核酸は、それが自然で見つかるコンテキストの外の核酸である。例えば、項はカバーする：自然の力で起こっているゲノムDNA分子の一部のシーケンスを有するが、符号化の両方とも側面に並んでいるというわけではないDNAまたはそれの側面に位置するnoncodingしているシーケンスがそれが自然に起こる有機体のゲノムの分子の中で、分ける(a);結果として生じる分子がいかなる自然の力で起こっているベクトルもまたはゲノムDNAと同一でないように、核酸はベクトルにまたは原核生物のゲノムDNAまたは方法の真核生物に結合した；(b）断片がポリメラーゼによって生じた別々の分子（例えばcDNA、ゲノム断片）は、反応を鎖でつなぐ（c）（PCR）、または、規制断片;そして、ある再結合物ヌクレオチド配列は、複合型遺伝子（i.e）の中で分かれる（d）。（融合タンパク質をコード化している遺伝子）または事実上分からないシーケンスを有する修正された遺伝子。

【００４６】

ここで使用しているように、「人間のミトコンドリア・ゲノムのヌクレオチド位置に対するヌクレオチドlocus"refers。ケンブリッジ・シーケンス配列番号：2が参照シーケンスとして使われる、そして、本願明細書において関連されるミトコンドリア・ゲノムの位置はそのシーケンスと関連して割り当てられる。ここで使用しているように、「複数の場所に対するloci"refers。ここで使用しているように、「異なるベースが異なる個人のその場所で自然に起こる選択されたシーケンスからの選択されたヌクレオチド場所の単一のヌクレオチドに対するヌクレオチドallele"refers。その場所（例えば3796C）にあるヌクレオチド場所番号およびベースとしてヌクレオチド対立遺伝子情報は本願明細書において提供される。そして、それはケンブリッジ・シーケンスの人間のミトコンドリア位置3796で、シトシン(C)があることを意味する。ここで使用しているように、アミノ酸対立遺伝子」は、異なるアミノ酸が異なる個人のその場所で自然に起こるときに、人間のミトコンドリア・ゲノムの選択されたアミノ酸場所にあるアミノ酸に関連する。人間のミトコンドリアの13のタンパク質-符号化遺伝子が、ある。各々の遺伝子のために、コード化されたタンパク質は、一致して始まって番号をつけられるアミノ酸から成る。ND1は、304H、ヒスチジンがND1タンパク質のアミノ酸304であることを意味する。アミノ酸は、コドンによってコード化される。ここで使用しているように、周知のように、タンパク質のアミノ酸をコード化する3つのヌクレオチドの集まりに対するcodon"refers。アミノ酸対立遺伝子は、ヌクレオチド場所の一つ以上によってそれのためのそのコードにゆだねられることができる。例えば、ntlは15884P、コード化されるプロリン（P）がヌクレオチド場所15884を含んでいるコドンによってあることを意味する。

【００４７】

ここで使用しているように、「他の個人の対応する遺伝子と比較して偶然に予想されるより、統計学的にかなりnonsynonymousなヌクレオチド変化を有する遺伝子に、重要なgene"refersを進化的にする。ここで使用しているように、「そのヌクレオチド対立遺伝子を使用して進化的にに重要であると決定された遺伝子または他の個人の対応する遺伝子の等価なヌクレオチド対立遺伝子に位置するヌクレオチド対立遺伝子に、重要なヌクレオチドallele"refersを進化的にする。ここで使用しているように、「1つの種の範囲内のintraspecific"means。ここで使用しているように、「よりかなりの人口の範囲内の人口に対するsubpopulation"refers。部分母集団が、小さくとして唯一の個人としてあることがありえる。ここで使用しているように、「統計学的に有意の数の個人が同じことを有する地理的領域に地理的なregion"refersハプロタイプ。ここで使用しているように‖地理的領域が有することにゆだねるbeing"native"toハプロタイプその地理的領域伴う。ハプロタイ

プ、領域においてまたは人間の進化に関して歴史的に領域に定住した多くの個人の中で起こったそれは、地理的領域と関係している。

【００４８】

ここで使用しているように、「配列分析のためのサンプルとして使用される核酸のコレクションに対するtarget"or"target sample"refers。目標は、配列の調査によって問い合わせられる。「target"or"targetsample"mayは、結合されるいくつかのサンプルの混成である。例えば、実験的な目標サンプルは、組み合わさられることができる異なって、サンプルに制御目標をラベルをつけて、そして、配列に交雑した関連されている結合されたサンプルとして「それの間の配列の調査によるtarget"interrogatedは、実験する。ここで使用しているように、「試験されるinterrogated"means。調査が目標（i.e）に問い合わせることが可能であるように、調査、目標および雑種形成状況は選択される。目標サンプルの補完的なシーケンスに交雑することで。

【００４９】

ここで使用しているように、blindness"refersを呈する増加する可能性、通常見える能力を失うおよび／または若者で通常見える能力を失う通常の確率よりそれ以上は、熟成する。

【００５０】

本願明細書において定められる全てのシーケンスは、同じく補完的なふさを含むはずである二本鎖ポリヌクレオチド与えられたシーケンスを成る。

【００５１】

本発明は、全ての主要人間のハプログループで見つかる人間のmtDNA多形性のリストを提供する。実施例1は、世界中の主要人間のハプログループの代表である100の人間のmtDNAゲノムの上の塩基配列決定から、データを要約する。概要は、900以上の点突発変異および1つの9-塩基対削除を含む。表3（HumanMtDNA Nucleotide Alleles）は、103において識別される対立遺伝子の一覧を示す第一塔の第3の柱（第二カラムおよびヌクレオチド場所（ケンブリッジ・シーケンスの位置）のケンブリッジmtDNAシーケンスの対応する対立遺伝子）のこの種のシーケンス。表3は、自然に異なるハプログループで起こる人間のmtDNAヌクレオチド対立遺伝子の一組の一覧を示す。表3は、疾患を伴うことはすでに公知の対立遺伝子を含まない（i.e。表1の対立遺伝子を含まない）。柱1の対応するヌクレオチド場所と共に、表3の柱3にリストされるヌクレオチド対立遺伝子は、non-ケンブリッジ人間mtDNAヌクレオチド対立遺伝子の一組を形成する。表4は、柱3の48の人間のmtDNAゲノムおよび柱2の対応するケンブリッジ対立遺伝子においてこの発明者によって識別されるヌクレオチド対立遺伝子の一覧を示す。表4の柱1および3は、48のゲノムの非ケンブリッジ人間のmtDNAヌクレオチド対立遺伝子の一組を形成する。

【００５２】

自然に起こって、ケンブリッジ・ヌクレオチド対立遺伝子を含む表3にリストされるヌクレオチド対立遺伝子は、異常な生理的状況を伴う対立遺伝子を識別することに役立つ。選択された生理的状況と関連する新規な対立遺伝子を識別する方法を実行するときに、これらのヌクレオチド対立遺伝子は分析ステップの間、無視されることができる。

【００５３】

下記のように、影響を受ける個人が異常な生理的状態にかかるときに、それらのハプログループのために固有でない地理的領域において、それらがあるので、表3の特定の対立遺伝子は精力的なアンバランスのようなエネルギー代謝、代謝疾患、異常なエネルギー代謝、異常な温度規則、異常な酸化的リン酸化、異常な電子伝達、肥満、体脂肪の量、糖尿病、高血圧および心血管疾患に関連した生理的状況を識別することに役立つ。

【００５４】

ケンブリッジ・ヌクレオチド対立遺伝子を含んで、表3にリストされるヌクレオチド対立遺伝子は、また、人間のハプログループのmtDNAシーケンスおよび診断を識別することに役立つ。実施例2は、2つのチンパンジmtDNAシーケンスとともに103人の個人およびケンブリッジ・シーケンスのシーケンス・データの系統発生の分析を要約する。結果は、分岐

図の図1に示される。最も最近の普通の祖先（MRCA）以来時間の算出は、表5に示される。104人の個人は周知のハプログループから選択された、そして、対応するハプログループは図にラベルをつけられる。周知のように、人間のハプログループに図1および地理的地方土地の人をもつ104人の個人のシーケンス・データを結合することは図2（実施例3）に結果としてなる。そして、それは人間のmtDNA移動を追う。各々のハプログループを表しているmtDNAゲノムいくつかのシーケンスの分析は、どの対立遺伝子がどの対立遺伝子が一つ以上のハプログループの範囲内であらゆる個人に存在するか同様にハプログループの範囲内で分離しているかについて示した。各々のハプログループの範囲内であらゆる個人に存在する対立遺伝子は、図3（実施例4）に示される。左に、副ハプログループおよびハプログループは、一覧を示す。Macroハプログループは、括弧に示される。各々のグループ（副haploまたはhaplo）の全てのメンバーに存在するヌクレオチド場所および対立遺伝子は、一覧を示す。垂直棒は、右に対する対立遺伝子の全てが左にハプログループおよび／または副ハプログループの全てに存在することを示す。図3は、分岐図として描画される。例えば、図3はmacroハプログループ(R)個人が全て12705Cおよび16223Cを含む、そして、他のいかなる個人もこれらの対立遺伝子を有することは公知でないことを証明する。そして、(R)が診断されることができるマクロハプログループがmtDNA（12705Cか16223Cの存在）を含んでいるサンプルにおいて、したがって、同一化する。同様に、マクロハプログループ（N）は、8701A、9540Tまたは10873Tの存在を識別することによって診断されることができる。

【００５５】

図3のデータの分析は、ハプログループ（実施例5）を診断することに役立つ対立遺伝子の集合を示した。これらの対立遺伝子は、表6および7のハプログループによって、そして、表8および9の副ハプログループによって一覧を示す。ハプログループの全ておよび図3の副ハプログループを診断することに役立つ対立遺伝子の一組は、表10にリストされる。表10は、ヌクレオチド場所を柱1および柱2のハプログループを診断することに役立つヌクレオチド対立遺伝子にリストする。表10は、ケンブリッジ・シーケンスから若干の対立遺伝子を含む。ハプログループを診断する多くの等価な方法が、ある。1だけまたは2、3の場所だけを試験することを必要とするハプログループを診断する方法は、実施例5にリストされる。1つの特定の対立遺伝子だけの存在は通常ハプログループを診断することに充分である、しかし、しばしば、どの場所が試験される必要はあるかはわかっていない。表10にリストされる各々のヌクレオチド場所で、対立遺伝子を決定することによって、未知のサンプルのハプログループは、診断されることができる。あるいは、マクロハプログループは診断されることができるかまたは最初に除外されることができる。そして、それによって残留する、可能なハプログループとを区別するために試験されることを必要とする場所の数を減少させる。1だけまたは2、3の場所だけを試験することを必要とする方法によって、マクロハプログループを診断することに役立つ対立遺伝子は、表11に含まれる。本発明によって提供されるデータの更なる分析は、対立遺伝子のどの集合が追加的な下位のハプログループて追加的なmacro-ハプログループを識別するかについて示す。

【００５６】

サンプルのハプログループを診断することは、捜査および法医学分析に役立つ。選択された生理的状態と関連する、記載されている上記としての、そして、実施例6の新規な対立遺伝子を識別するときに、サンプルを特定のハプログループに帰属して、どの対立遺伝子が選択された生理的状態およびコンテキストにかかわらなかったかについて知っていると確認することは役立つ。実施例6に示すように、新規な対立遺伝子によって特定のハプログループの遺伝子のコンテキストだけの生理的状態が生じるときに、サンプルのハプログループを診断することはまた、選択された生理的状態と関連する新規な対立遺伝子を識別することに役立つ。実施例6において、ロシアで見つかるハプログループと関連する対立遺伝子のリストが、2つのロシアのLHON系統のシーケンス分析において使われた。表3にリストされる対立遺伝子を除去することによって、LHONを伴う2つの新しい突然変異は、識別された。これらの新規な複合体私、突然変異（3635Aおよび4640C）はレーバーHeredi

taryOptic Neuropathy（LHON）に対する素因を診断することに役立つ。
【００５７】
ハプログループ J.と関連するハプログループJがハプログループ J（ND5の458T）を診断することに役立つnonsynonymousな違いによって定義される場合にだけ、私、それがLHONに対する素因に生じさせるように見える複合体の10663C、実施例7は新しい主たるLHON突然変異の識別を示す。本発明は、LHONに対するおよび／または、人、ヌクレオチド対立遺伝子または10663Cの同義のヌクレオチド対立遺伝子からのmtDNAを含んでいて、更にハプログループJ（例えばND5の458T）の対立遺伝子診断を識別しているサンプルで、同一化することによって初期の開始盲目を呈することへの素因をもつ人を診断する方法を提供する。ND5が458Tそうな、ミスセンス突然変異全てのハプログループJ個人（突然変異が直接LHON.を引き起こして、関係していることができるこの事項）のND1は、304H、もう一方であるミスセンス突然変異、全てのハプログループJ個人に存在して、また、直接LHON.を引き起こすことに関係していることができる458Tは、また、ハプログループ T個人に存在する。ハプログループ Jは、また、centenariaおよび拡張寿命に対する素因を伴う。ND5およびND1は、458T、304H、また、直接centenariaおよび拡張寿命に傾向を引き起こすことに関係していることができる。
【００５８】
実施例8は、大陸間のmtDNAシーケンス放射線における人口統計学的要因の重要性を示す。ハプログループは、結合されて、統計分析のためのさまざまな人口に分かれた。
【００５９】
以前従来技術において、人間のmtDNA（例えば表3にリストされるヌクレオチド対立遺伝子）の多形性が全てのコンテキストにおいて中性で、生理的状況を伴うことはありえなかったと思われた。大陸間の移動と関連する人間のmtDNA多様性の違いがランダムな遺伝的浮動によったと思った‖（ｅ。g.創設者効果が、迅速な人口展開によってあとに続いた）。本発明では、これらの人間のmtDNA多形性の生物製剤および臨床重要性は、開示される。表3にリストされるヌクレオチド対立遺伝子の中性は、中性分析（実施例9-12）を使用して試験された。
【００６０】
表3のいくつかのヌクレオチド場所は、ミトコンドリア・タンパク質-符号化遺伝子（表2）に位置する。それらの場所の中で、いくつかの識別されたヌクレオチド対立遺伝子は、ヌクレオチド場所が住むコドンによってコード化されるタンパク質を変える。周知のように、これはミトコンドリア・コドン使用テーブルを使用して決定される。アミノ酸を変えるヌクレオチド対立遺伝子は、呼ばれるミスセンス突然変異、ミスセンス多形性またはnonsynomymousな違い。ミスセンス、多形性は比較されたシーケンスと関連するタンパク質シーケンス、しかし、それらがまだ中性でもよいことを変えるそれらがコード化されたタンパク質の機能に影響を及ぼさない。影響を受けるタンパク質の生化学研究を実行せずに、統計分析は多形性が中性かどうか、進化が対立遺伝子をコード化することの選択を賦課したかどうか、そして、その選択は陽性かどうか決定するために実行されることができる。本発明は、表3の多形性の統計分析の結果を提供して、対立遺伝子が中性でなくてしたがって、進化的にに重要であるリストを提供する。
【００６１】
ヌクレオチド対立遺伝子の中性テストは、最初にそれらのヌクレオチド対立遺伝子を含んでいる遺伝子の中性テストを必要とする。実施例9にて説明したように、2人の同一種内の住民から2セットのallelicな遺伝子を比較することによる一つ以上の遺伝子の中性テストは、実行された。ハプログループは、比較に人口を作るために結合された。実施例9において、地理的地域原産のハプログループを表して、mtDNAゲノムの全ての符号化領域からのヌクレオチド対立遺伝子は、第1の人口およびシーケンスの第1の一組を作るために化合した。異なる地理的地域原産のハプログループから、mtDNAゲノムの全ての符号化領域のヌクレオチド対立遺伝子は、第２の人口およびシーケンスの第２の一組を作るために化合した。ヌクレオチド対立遺伝子は、それらが同義で非同義の違いをコード化することに

分けられた。各々の遺伝子（対立遺伝子を含んでいる人口によって切り離される）のためのKa/Ksの比率は、表12に示される。1つの種の1人の住民から1セットの少なくとも一つの遺伝子の少なくとも2つのヌクレオチド対立遺伝子を比較することによる遺伝子の中性テストは、実施例10において実行された。実施例10において、地球上の全ての地理的地方のハプログループの中で、mtDNAゲノムの全ての符号化領域のシーケンスは、1つの人口になるために結合されて、分析のためのシーケンスの中でセットされた。図4は、1つの種（104の人間のシーケンス）だけの1人の住民から、1セットのシーケンスの比較の結果を示す。実施例11は、人間的な2人の住民（人間対P.paniscus）対P.穴居人（人間対8つの他の霊長類種）の間でシーケンスの一組の比較を含む、そして、人間対13の哺乳類の種。

【００６２】

進化的にな重要な遺伝子を識別するために、2セットのヌクレオチド配列（異なる人口からの各々の集合）は、各々と比較してある。遺伝子または一つ以上の全部の遺伝子の一部を表しているヌクレオチド配列は、役立つ。シーケンスの集合は、中性分析によって各々と比較してある。各々の集合からのシーケンスの違いは、同義であるかnonsynonymousな違いであると決定される。nonsynonymousな違いの割合は、同義の違い（魂/Ks）の割合と比較してある。分析の結果はデータセットにおいて編集される、そして、データセットは一つ以上の進化的にな重要な遺伝子を識別するために周知のように分析される。nonsynonymousな違いが同義の違いより可能性によって予想されるより、しばしば有意に起こるときに、遺伝子または遺伝子の一部は進化的にに重要であると決定される。同義の違いがnonsynonymousな違いより可能性によって予想されるより、しばしば有意に起こるときに、遺伝子または遺伝子の一部は節約されると決定される。比率が偶然に予想通りのときに、選択または進化の重要性の証拠がない。

【００６３】

進化的にな重要な遺伝子がヌクレオチド配列（1人の住民だけから）の1セットだけまた、分析されることができることを識別するためにｅ。g.1つの大陸上の人間生活のヌクレオチド配列典型。シーケンスの1つの一組だけが分析されるときに、一組が少なくとも2つの対応するヌクレオチド対立遺伝子を含まなければならなくて、（ｉ.ｅ。シーケンス多形性が、なければならない）。対応するシーケンスは、少なくとも2人の個人からの同じ遺伝子または遺伝子部のシーケンスである。人口の範囲内の異なる個人からのシーケンスは、各々に関して多形性を含まなければならない。各々と関連するシーケンスの違いは、同義であるかnonsynonymousであると決定される。中性分析は、データセットを生成するために実行される。データセットは、進化的にな重要な遺伝子を識別するために分析される。分析が分析遺伝子のいずれも進化的にに重要でないと確定する場合、ヌクレオチド配列の一組が増加することができて、例えば増加することによって、一つ以上の遺伝子が拡大した人口において進化的にに重要かどうか決定するために、シーケンスがある人口の容積は、由来した。

【００６４】

データがKcにKa/Ksを操作することによって更に分析されることを除いては、実施例12は実施例9と類似している。実施例9-12は、1つのmtDNA遺伝子以外の全てが中性であるというわけではなくて、したがって、進化的にに重要なことを証明する。遺伝子は向かう、公知技術の有意性試験によって中性でない。選択された人口を比較するときに、若干の遺伝子は進化的にに重要なだけである。例えば、ND4はアメリカインディアン・シーケンスをアフリカのシーケンスと比較するときに、重要なことを証明された、そして。そのとき、全ての人間のシーケンスを各々と比較すること、しかし、ヨーロッパ人をアフリカのシーケンスと比較して。ND4Lは、現在の分析によって進化的にに重要なことを示さない唯一のmtDNA遺伝子である。ND4Lは、一つ以上の異なる人口を使用しているかまたは遺伝子シーケンスの一部だけを使用している本発明の方法によって、進化的にに重要なことを証明されるかもしれない。実施例9-12において、各々の遺伝子の全てのシーケンスが分析のために使われたこと、しかし、遺伝子の部分は、また、本発明の方法に役立つ。有意性試験は、あまりに少ない遺伝子分がnon-中性を決定するために用いるのを防止する。

【００６５】
進化的にな重要な遺伝子を識別した後に、進化的にな重要なヌクレオチド対立遺伝子は、識別されることができる。進化的にな重要なヌクレオチド対立遺伝子を識別するために、１人か２人の住民を使用して、進化的にな重要な遺伝子を識別するためのステップは、進化的にな重要なヌクレオチド対立遺伝子を決定するためにシーケンス・データセットを分析するステップの追加によって実行される。進化的にな重要なヌクレオチド対立遺伝子は、進化的にな重要な遺伝子のallelicなアミノ酸をincodingしているシーケンスまたは遺伝子の一部分の一部である。実施例13および14は、進化の重要なヌクレオチド対立遺伝子の識別および実施例9-12において識別される進化的にな重要な遺伝子の進化的にな重要なアミノ酸対立遺伝子を示す。進化的にな重要なアミノ酸対立遺伝子は、進化的にな重要なヌクレオチド対立遺伝子を含んでいるコドンによってコード化されるアミノ酸である。これらの実施例において、進化的にな重要なヌクレオチド対立遺伝子を含んでいる全てのコドンおよびそのコドンによってコード化されるアミノ酸が決定されることができるために、表3にリストされない場所のヌクレオチドはケンブリッジ・シーケンスと同一である。進化的にな重要なアミノ酸対立遺伝子と同じアミノ酸をコード化しているコドンの一部である全てのヌクレオチド対立遺伝子は、本願明細書において同一化してまたは本発明の方法によって同一化して、また、進化的にに重要で、本発明の範囲内のことを目的とする。場所のような、進化的にな重要なアミノ酸対立遺伝子は、2つの隣接したヌクレオチドで複数のヌクレオチド対立遺伝子を含むことができる。人間のミトコンドリア・シーケンス（本発明の方法によって識別される）の進化的にな重要なヌクレオチド対立遺伝子および進化的にな重要なアミノ酸対立遺伝子は、表14にリストされる。柱1において、表14が対立遺伝子を含んでいる遺伝子の一覧を示すこと、柱2はヌクレオチド対立遺伝子の場所を示す、柱3はそのヌクレオチド場所でケンブリッジ・ヌクレオチド対立遺伝子の一覧を示す、柱4は本発明（アミノ酸がケンブリッジ・ヌクレオチド対立遺伝子（他のケンブリッジ・ヌクレオチドがコドンの他のヌクレオチド場所に存在するときに）を含んでいるコドンによって、コード化した柱5つのリスト）のnon-ケンブリッジ対立遺伝子の一覧を示す、そして、柱6は非ケンブリッジ対立遺伝子（ケンブリッジ・ヌクレオチドがコドンの他のヌクレオチド場所に存在するときに）を含んでいるコドンによってコード化されるアミノ酸の一覧を示す。柱2、3および4は、進化的にな重要な人間のミトコンドリア・ヌクレオチド対立遺伝子の一組を作る。柱2、5および6は、進化的にな重要な人間のミトコンドリア・アミノ酸対立遺伝子の一組を作る。表14は、一覧を示す対立遺伝子のヌクレオチド場所を示す。柱5および6にリストされるアミノ酸対立遺伝子のために、関連した場所は、柱2にリストされるヌクレオチド場所を含んでいるコドンをコード化する際の全ての3つのヌクレオチド場所である。

【００６６】
進化的にな重要なアミノ酸対立遺伝子を識別するために、１人か２人の住民を使用して、進化的にな重要な遺伝子を識別するためのステップは、2つのステップの追加によって実行される：1）進化的にな重要なヌクレオチド対立遺伝子を決定するデータセットを分析する;そして、2）コード化されたアミノ酸対立遺伝子を決定する。進化的にな重要なアミノ酸対立遺伝子は、それが比較された対応するアミノ酸対立遺伝子と関連して、nonsynonymousな違いを表して、異なるアミノ酸である。そこにおいて、遺伝子は対応する一つ以上人口において進化的にに重要であると決定された。

【００６７】
本発明において、偶然に予想されるより、アミノ酸置換突然変異（nonsynonymousな違い）が人間のmtDNAsにおいて非常に一般的である、そして、それらのほとんどが進化的にに重要なことを証明される。本発明は、これらの対立遺伝子が選択によって固定したことを証明する。ミトコンドリア遺伝子は、体温を維持するためにエネルギーを生成するために、そして、生成熱源のために原因となるタンパク質をコード化する。人間が世界の異なる一部に移ったので、それらはダイエットおよび気候の変化に遭遇した。mtDNAの高突然変異速度および細胞エネルギー論のミトコンドリア・タンパク質の中心的な役割は、mtDN

Aに気象的で食事の状況を変化させることへの急速な哺乳類の適応ができるようにする理想的なシステムを作る。人間のmtDNA遺伝子の一つであるとわかった増加するアミノ酸シーケンス変動性は、自然な選択が電子伝達チェーン（ETC）およびATP合成間のカップリング効率を変えたmtDNA対立遺伝子を支持したという事実によると、勾配（AT）がミトコンドリア内側膜陽子によって決定した。ETCおよびATP合成間のカップリング効率はATPシンターゼの陽子チャネルによって、相当な範囲に調停される。そして、それはmtDNA-コード化されたATP6タンパク質および核DNAコード化されたATP9タンパク質から成る。ATP6遺伝子（それはより漏れるATPシンターゼ陽子チャネルをつくる）の突然変異は、ATP製造を減らすが、消費される各々のカロリーのための熱発生量を増やした。エネルギー・バランスのこの種の変化は、温暖であるか北極気候において有益であるが、熱帯気候において有害だった。ダイエットおよび気候の遭遇された変化へのより良い適応を可能にしているmtDNA対立遺伝子を得ている人間はより高い遺伝子の適合性を経験した、そして、それらの対立遺伝子は選択された。特に、それらが適応効果があったので、人間が北極（図2）に、EurAsian温帯に夢中で動いているアフリカの熱帯地方から移動したので、これらの対立遺伝子は遺伝的に決められた。母らしく継承されたmtDNAsの組換えの不足は、迅速な分離、表現および有利なmtDNA対立遺伝子の適応より抜きを支持した。中で違いのなし明瞭な非ランダム-大陸間の変化が選択も影響したことを証明する同義のmtDNA対同義の大陸間の植民地化。遺伝子のランダムがヒッチハイクをして、例えば中で、同義の対立遺伝子は、それから定義可能な大陸に特有のハプログループに結果としてなった。

【００６８】

我々の祖先が異なる環境条件へ移動したので、現代のmtDNAバリエーションは適合によって成形された。個人が異なる環境に居住を定めるときに、1つの気象的で食事の環境において有利である異型はmaladaptiveである。本発明の方法は、mtDNAヌクレオチド対立遺伝子をハプログループと関連させせて、従来技術において個人を傾向を有すると診断することは公知であるように、このデータを自国のハプログループ地理的地方と結合するために遅発性臨床的な障害（例えば肥満、糖尿病、高血圧およびそれらのmtDNA対立遺伝子に関して不利である気象的で食事の環境において、それらの個人が生きる心血管疾患）。局所のmtDNA対立遺伝子を有する人間が対立遺伝子が選ばれたものと異なる熱および／または食事の環境へ移動するときに、それらはそれらの環境によって精力的にimbalancedされて、その結果代謝疾患（例えば糖尿病、高血圧、心血管疾患および代謝およびミトコンドリア機能と関係している技術にとって公知の他の疾患）を有することの素因を作られる。上述の遅発性、臨床障害は、急速に世界中の我々のグローバルにモバイル社会の部材の流行になっている。本発明は、制限されなくて、例えば生理的状態に対する素因をもつ人間を診断する方法を提供するために（精力的なアンバランス）代謝疾患、異常なエネルギー代謝、異常な温度規則、異常な酸化的リン酸化、異常な電子伝達、肥満、体脂肪の量、糖尿病、高血圧および心血管疾患。方法はサンプルの、そして、したがって、個人のハプログループを決定するために地理的領域の個人から、ミトコンドリア核酸を含んでいるサンプルを試験することを含む。そして、個人のハプログループがその地理的地域原産のハプログループの一組において、ない場合、その地理的地域原産のことは公知のハプログループの一組に対する個人のハプログループを比較して、上述の状況に対する素因をもつ個々の人間を診断する。本発明は上述の方法によって診断される上述の状況のうちの1つの処置を可能にする。そして、診断された人間を人間のハプログループ原産の領域（s）として、類似した気候の中である地理的領域へ再配置することから成っておよび／またはより密接に歴史的に人間のハプログループ原産の領域（s）において利用できるダイエットにマッチするために診断された人間のダイエットを変える。

【００６９】

生理的状態に対する素因を診断する上記の方法は、また、アミノ酸対立遺伝子を生理的状態と関連させることに役立つ。診断された個人のハプログループにおいて、そして、個人の地理学的位置原産のハプログループ以外において存在する進化的にな重要なアミノ酸対立遺伝子は、本発明の方法によって生理的状態を伴う。アミノ酸対立遺伝子、そして、

対応するヌクレオチド対立遺伝子（ハプログループを診断することに役立つ）、そして、それらが役立つハプログループ、診断することは表15にリストされる。表15および同義的に符号化ヌクレオチド対立遺伝子において一覧を示すアミノ酸対立遺伝子および対応するヌクレオチド対立遺伝子は、上述の生理的状況を伴う。表15は、ハプログループを診断することに役立つアミノ酸対立遺伝子の一組の一覧を示す。表15の柱1は遺伝子の一覧を示す、柱2はヌクレオチド場所の一覧を示す、柱3は役立つヌクレオチド対立遺伝子（ケンブリッジ・ヌクレオチドがコドンをコード化する他のヌクレオチド場所に存在する役立つヌクレオチド対立遺伝子によって、役立つアミノ酸対立遺伝子がコード化した柱4つのリスト）の一覧を示す、そして、対応する対立遺伝子を含む括弧で、柱5はハプログループまたは副ハプログループの一覧を示す。アミノ酸対立遺伝子（柱4）は、ヌクレオチド場所（柱2）を含んでいるコドンによって識別されることができる。ntlがヌクレオチド場所（ntl）を含んでいるコドンを示す所で、ND1遺伝子のプロリンが3796Pntlと確認されて、例えば3796。表15の柱5にリストされるハプログループのうちの1つの個人が上述の方法によって上述の生理的状況のうちの1つによって診断されるときに、生理的状態は表15にリストされる対立遺伝子のうちの1つの存在を伴う。個人のハプログループがハプログループGであるときに、生理的状態を引き起こしてそうなアミノ酸対立遺伝子がntl 4833のA. Whenであって、個人のハプログループはハプログループ Tである、アミノ酸対立遺伝子は8701Tntl 14917のD、ntlおよびntl 15452からなるグループから選択されるI. When、ハプログループはハプログループ Wである、アミノ酸対立遺伝子はntl5046からなるグループから選択される私、5460T ntl、8701T ntl、そして、ntl 15884 P. When ハプログループがハプログループDである、アミノ酸対立遺伝子は5178M ntlおよびntl 8414のF. Whenからなるグループから選択される。そして、ハプログループはハプログループL0である、アミノ酸対立遺伝子は5442L 9402P 13105V ntl、ntl 7146のA、ntl、ntlおよびntl 13276のV. Whenからなるグループから選択される。そして、ハプログループはハプログループL1である、アミノ酸対立遺伝子は7389H 13105V 13789H ntl 7146のA、ntl、ntl、ntlおよびntl 14178のV. Whenからなるグループから選択される。そして、ハプログループはアミノ酸対立遺伝子がグループから選択されるハプログループCであるntlハプログループがグループから選択される8584Tおよびntl 14318のS. When‖ハプログループ A（私）X、B、F、Y、そして、U、アミノ酸対立遺伝子はntl8701のT. Whenである。そして、ハプログループはアミノ酸対立遺伝子がグループから選択されるハプログループ Jである8701T ntl、13708Tntl、そして、ntl 15452 I. When、ハプログループはハプログループ VおよびHからなるグループから選択されるハプログループである。そして、アミノ酸対立遺伝子は8701T14766T ntlおよびntlからなるグループから選択される。

【００７０】

進化的にな重要なヌクレオチドおよびアミノ酸対立遺伝子も、エネルギー代謝-related..生理的状態（例えば精力的なアンバランス、代謝疾患、異常なエネルギー代謝、異常な温度規則、異常な酸化的リン酸化、異常な電子伝達、肥満、centenaria、糖尿病、高血圧および心血管疾患）に対する素因を診断して、役立つATP9をコード化されるnuclear-の中に存在する。これらの対立遺伝子は、本発明の方法によって識別されることができる。

【００７１】

進化的にな重要なアミノ酸対立遺伝子および対応するヌクレオチド対立遺伝子は、傾向が本発明の方法によって診断可能である生理的条件を引き起こしている対立遺伝子の候補である。本発明（表19）の方法によって識別される進化的にな重要なアミノ酸およびヌクレオチド対立遺伝子は、対応する生理的状況を治療するために遺伝子治療およびミトコンドリア置換療法に役立つ。進化的にな重要な遺伝子、アミノ酸対立遺伝子および本発明の方法によって識別されるヌクレオチド対立遺伝子は、従来の治療のための、そして、通信治療的なエージェントを設計するための目標を識別することに役立つ。本発明の方法によって識別される進化的にな重要な遺伝子およびアミノ酸およびヌクレオチド変化は、対応する人間の生理的状況の生成動物のモデルに役立つ。

【００７２】

技術にとって公知にさているように、個人はいかなる与えられたヌクレオチド場所でも複数のミトコンドリアＤＮＡ対立遺伝子を含むことができる。1つの細胞は多くのミトコンドリアを含む、そして、1つの有機体の範囲内の1つの細胞であるか異なる細胞は遺伝的に異なるミトコンドリアを含むことができる。Heteroplasmyは、複数種類の個人またはサンプルのミトコンドリアの発生である。heteroplasmyの程度を変化させることは、本願明細書において記載されている生理的状況の程度を変化させることと関係している。Heteroplasmyは技術にとって公知の手段によって識別されることができる、そして、特定のヌクレオチド対立遺伝子と関連する生理的状態の厳しさは個人の範囲内でこの種の関連する対立遺伝子のパーセンテージによって変化すると思われる。

【００７３】

本発明の方法は一覧を示す実施例の人間のミトコンドリア・ゲノムを分析するために用いる、しかし、方法はまた、他のゲノムおよび他の種を分析することに役立つ。本発明の方法は、核および葉緑体ゲノムにおいて例えばミトコンドリア・ゲノムに加えて他のゲノムの進化的にな重要なタンパク質-符号化遺伝子および対応してコード化された突然変異を識別することに役立つ。人口（図1）として人間のハプログループを使用して、本発明の方法は、人間の核遺伝子の進化的にな重要なタンパク質-符号化遺伝子および対応する進化的にな重要な対立遺伝子を識別することに役立つ。本発明の方法は、また、多くの種の進化的にな重要なタンパク質-符号化遺伝子および対応する対立遺伝子を識別することに役立つ。例えば、本発明の方法は、牛肉または酪農牛またはブタ線の種類に適用できる。穀物線は、表現型および／または分子の標識によって割り切れるにヘテロ強勢である本発明の方法の役立つ人口であるグループ。穀物を使用するヘテロ強勢である人口として集める、本発明の方法は、核兵器、葉緑体および穀物のミトコンドリア・ゲノムの進化的にな重要なタンパク質-符号化遺伝子および対応する突然変異を識別することに役立つ。

【００７４】

本発明は、ライブラリの本発明の新規なヌクレオチド対立遺伝子を含んでいる単離された核酸分子を提供する。ライブラリは、最少の二つでこの種の分子を含む。好ましくは、分子はユニークなシーケンスを有する、分子は約7から約30のヌクレオチドまで概して長さを有する。「本願明細書において用いられるAbout"asは、約10%以内に意味する（ｅ。ｇ.（「約30のnucleotides"means27-33のヌクレオチド））。しかし、約50のヌクレオチドのような長く、分子はより長くてもよい。本発明のライブラリは、本発明の少なくとも一つの非ケンブリッジ・ヌクレオチド対立遺伝子を各々含んでいる少なくとも2つの単離された核酸分子を含む。本発明のライブラリは、少なくとも10の、25の、50の、100,500の以上単離された核酸分子（いずれが本発明のヌクレオチド対立遺伝子を含むか、少なくとも1つ）を含むことができる。本発明のライブラリは、本発明のヌクレオチド対立遺伝子の全てに、少なくとも2を有する分子を含むことができる以下から成る同義の符号化進化的にな重要なアミノ酸対立遺伝子の。本発明のヌクレオチド対立遺伝子は、ヌクレオチド場所（人間のミトコンドリア・ゲノムのヌクレオチド場所）によって、そして、Aジーシー T（またはU）ヌクレオチドによって定義される。本発明のライブラリで、単離された核酸分子は本発明のヌクレオチド対立遺伝子を含むと確認されることができる。－その理由は、次のことにある。本発明のヌクレオチド対立遺伝子はミトコンドリア・ゲノムのそのコンテキストによって、少なくとも一つの側に囲まれている。統計学的に、人間のミトコンドリア・ゲノムにおいてユニークであるために、この種の分子は、長く少なくとも約7つのヌクレオチドであることを必要とする。統計学的に、全ヒトゲノムにおいてユニークであるために、ミトコンドリア・ゲノムを含んで、この種の分子は、長く少なくとも約15のヌクレオチドであることを必要とする。本発明の単離された核酸分子の実施例は、以下のヌクレオチド対立遺伝子を含んでいる分子である：1）人間のmtDNAヌクレオチド場所168-170、場所の171A非ケンブリッジ対立遺伝子および人間のmtDNAヌクレオチド場所172-174のケンブリッジ対立遺伝子でケンブリッジ対立遺伝子;そして、2）11940-11946、11947Gの非ケンブリッジ対立遺伝子および11948-11954のケンブリッジ対立遺伝子でケンブリッジ対立遺伝子。本発明の単離された核酸分子は、本発明の複数のヌクレオチド対立遺伝子

を含むことができる。本発明のヌクレオチド対立遺伝子が、単離された核酸分子のいかなる位置にもあってもよい。しばしば、単離された核酸分子の中央にまたは分子の3'endに関連したヌクレオチド対立遺伝子を有することは、役立つ。本発明の単離された核酸分子は、質問することに役立つ有無を決定するの、公知技術のいかなる方法も用いて人間からミトコンドリア核酸を含んでいるサンプルのミトコンドリア・ゲノムの対応するヌクレオチド場所のヌクレオチド対立遺伝子。ヌクレオチド対立遺伝子の欠如の存在を決定することのための方法は、明確なallele-を含むPCR、そして、核酸配列雑種形成または塩基配列決定。

【００７５】

本発明の対立遺伝子およびライブラリは、核酸配列のための調査を設計することに役立つ。本発明は、2つ以上の核酸分子または点（複数のかなり同一の単離された核酸分子から成る各々の点）を有する核酸配列を提供する、本発明の対立遺伝子のシーケンスを有する各々の分子。本発明の配列上の分子は、長く通常約30のヌクレオチドに対する約7である。配列は、対立遺伝子の有無を検出することに役立つ。本発明の配列は、また、塩基配列決定人間mtDNAに役立つ。さまざまな表およびこのテーブルの部分において一覧を示すにつれて、対立遺伝子は人間のmtDNAヌクレオチド対立遺伝子を含んでいるヌクレオチド対立遺伝子、非ケンブリッジ人間のmtDNAヌクレオチド対立遺伝子、48のゲノムおよびケンブリッジ・シーケンスの人間のmtDNAヌクレオチド対立遺伝子、48のゲノムの非ケンブリッジ人間のmtDNAヌクレオチド対立遺伝子、人間のハプログループおよびマクロハプログループを診断することに役立つヌクレオチド対立遺伝子、人間のハプログループを診断することに役立つヌクレオチド対立遺伝子および進化的にな重要な人間のミトコンドリア・ヌクレオチド対立遺伝子の一組から選ばれることができる。本発明の配列は、対立遺伝子の上述の集合のうちの1つの対立遺伝子の全てに問い合わせることができる分子を含むことができる。シーケンス多形性（例えば、それは本発明によって提供される）を検出することに役立つgenotypingしている配列は、PerfectMatch調査（PM）を含んでいる配列をgenotypingしているAffymetrix（サンタクララ、CA、USA）および対応するMismatch調査と類似している（mM）。PM調査は選択されたヌクレオチド場所で非ケンブリッジ対立遺伝子から成ることができた、そして、対応するMM調査は選択されたヌクレオチド場所で対応するケンブリッジ対立遺伝子から成ることができた。本発明の配列は、人間のmtDNAのための塩基配列決定配列を含む。

【００７６】

ここで使用しているように、単離された核酸分子の順序集合またはかなり同一の単離された核酸分子の複数からなる点に対するarray"refers。好ましくは、分子は基板に付着する。各々（基板上の）の位置が公知であるために、点または分子は命じられる、そして、各々のアイデンティティは公知である。マイクロスケール上の配列は、マイクロアレイsと呼ばれていることができる。固体基板（例えばガラスまたは他のセラミック・スライド）上のMicroaraysは、遺伝子チップまたはチップと呼ばれていることができる。

【００７７】

配列は、好ましくは固体基板に印刷される。印刷する前に、技術にとって公知にさているように、ガラス・スライドのような基板は結合することに役立つ表層を提供する準備ができている。配列は、いかなる印刷技術もおよび公知技術の機械を使用して印刷されることができる。印刷することは、調査を基板に配置して、調査を基板に取り付けて、予防する基板をブロックすることが必要である非特異的である雑種形成。点は、周知の場所で印刷される。配列は、ガラス顕微鏡スライドに印刷されることができる。あるいは、調査は用意された固体基板（Affymetrix、サンタクララ、CA、USA）の周知の位置において総合されることができる。

【００７８】

本発明の配列はわずか2つの点を含むことができる、または、約10以上の点、約25以上の点、約100以上の点、約1000以上の点、約65以上、000点または上に向かう乃至約数百千はしみがつく。

【００７９】
　マイクロアレイsを使用することは目標の拡大がそばに例えば重要なシーケンスの中で配列する（同じシーケンスの多数のコピーの生成）ことが必要であることができる‖PCR、または、転写を逆転させせる。核酸が複製されるにつれて、それは電球のように光を発する蛍光ラベルを付け加えられる。ラベルをつけられた核酸は、マイクロアレイに導入されて、期間の間反応することができる。この核酸は、支持するかまたは雑種を作る、調査が十分に相補的である配列上の調査を有するためにラベルをつけて、拡大する、サンプル核酸。余分の核酸は配列の中で洗い落とされる。そして、調査に結合した核酸だけを残す。蛍光スキャナを有する配列のイメージを得て、ソフトウェアを雑種を作られた配列イメージを分析するために用いることによって、それが決定されることができること、そして、どんな範囲に、遺伝子は入れ換えられるシーケンスがそうであるのであるにせよ、または、配列上の特定の場所で蛍光強度を比較することによって、現れる。信号の強度は、シーケンスがどんな範囲にあるかについて指し示す。表現配列において、高いまばゆい信号は遺伝子の多くのコピーがサンプルに存在することを示す、そして、より低いまばゆい信号は遺伝子がより活発でないことを示す。適当な雑種形成状況および調査を選ぶことによって、この技術は、一つのヌクレオチド多形性（SNPs）を検出するために、そして、塩基配列決定のために役立つ。マイクロアレイsを設計して、使用する方法は、連続的に改善されている（Relogio（A.その他.(2002)Nuc.）酸。物。30（11）：e51;イワサキ（Hほか（2002の）DNA Res. 9（2））：59-62 ;そして、Lindroos（K.ほか（2002の）Nuc.Acids. Res. 30（14））：E70）。

【００８０】
　本発明の配列は、公知技術のいかなる配列合成方法（例えば観測している技術または強力な位相合成）にもよってなされることができる。好ましくは、本発明の配列は、フォトリソグラフィおよび組合せの化学の組合せを使用している強力な位相合成によって合成される。調査選択および配列設計のいくつかのキー素子は、全ての配列の生産に共通である。例えば、調査雑種形成を最適化する戦略は、一定不変に調査精選品の加工に含まれる。特定のpH、塩および温度条件の雑種形成は、温度を溶解させている利益への取得によって、そして、所望の雑種形成挙動と相関する経験的規則を用いて最適化されることができる。コンピュータ・モデルが、調査雑種形成の強度およびconcentration-依存を予測するために使われることができる。

【００８１】
　特定の多形性を検出することは、2台の探測機を使用して達成されることができる。1台の探測機は完全に目標シーケンスと相補的に設計されている、そして、その中心の単一のベース不適当な組合せを除いて同一であるパートナー調査は発生する。Affymetrixシステムにおいて、これらの調査一組は、PerfectMatch調査（PM）およびMismatch調査（mM）と呼ばれている。それらは、生じられる信号の計量および減算を考慮に入れる非特異的である十字-雑種形成。雑種形成の違いはパートナーの間で信号を送る、それらの強度比率と同様に、特定の目標多量の、そして、従ってシーケンスのインジケータとして役立つ。配列は、シーケンスの個々のヌクレオチドに問い合わせるために多数の調査に依存することができる。目標ベースのアイデンティティは目標位置だけにおいて変化する4台の同一の探測機を使用して演繹されることができる。そして、各々4つの可能なベースのうちの1つを含む。あるいは、共通配列の存在は、特定の対立遺伝子を表している１、２の調査を使用して試験されることができる。遺伝子型にヘテロである、または、遺伝的にサンプルを調合されて、多くの調査を有する配列は、明確にgenotypingすることに結果としてなって、冗長な情報を提供するために作成されることができる。

【００８２】
　アニール化されることは起こるまで、固体基板および目標（サンプルのヌクレオチド配列）に取り付けられる探測機は雑種形成緩衝液に組み込まれて、適当な温度で保持した。その後で、外来の材料から自由な基板は洗われる。そして、放射能写真、計数している液体火花および／または蛍光によって例えば公知技術の方法によって検出および計量を考慮

に入れている固定された調査分子に結合される目標上の核酸を残す。改良が雑種形成および検出技術でなされるにつれて、それらは従来技術において当業者によって直ちに適用されることができる。周知のように、調査分子および目標分子が2つの分子間の強いnon-共有結合を形成することによって交雑する場合、アニール化されて、ステップを洗うことは高い厳しさの状態の下で、実行される場合、調査および目標核酸が本質的に同一であるかほぼ完全に相補的であると合理的に仮定されることができる。検出可能なラベルは、手段を雑種形成が起こったかどうかについて決定することに提供する。

【００８３】

雑種形成調査としてオリゴヌクレオチドまたはポリヌクレオチドを使用するときに、調査はラベルをつけられることができる。本発明の配列において、目標は技術にとって公知の手段によってその代わりにラベルをつけられることができる。目標は、放射性であるか非放射性ラベルについてのラベルがついていることができる。目標は、好ましくは蛍光ラベルを含む。

【００８４】

雑種形成の厳しさのさまざまな程度は、使用されることができる。状況は、より厳しい、より大きい‖二重形成のために必要である相補性。厳しさは、温度、調査集中、調査長、イオン強度、時間、などによって制御されることができる。雑種形成実験はしばしば適度にの下にへ技術による状況が従来技術においてよく知っている高い厳しさに導かれる。そして、ケラー、G. H.およびM. M. Manak（1987の）DNA Probeの実施例のために、そのことはストックトン・プレス、ニューヨーク、NYを記載した。（pp.169-170）、本願明細書に引用したものとする。しかし、周知のように、塩基配列決定配列は、概して下部の雑種形成厳しさを使用する。

【００８５】

雑種形成の条件が公知である高い厳しさに、技術を加減する。汚点の高い厳しさ条件の実施例は、5X SSC/5Xの68のCで交雑しているデンハート液/0。0.2X SSC/0の1%の特別割引販売および洗濯物。1！室温で、o SDS。適度な厳しさの条件の実施例は、5XSSC/5Xの68のCで交雑しているデンハート液/0。3X SSC.の42のCの1%の特別割引販売および洗濯物温度および塩濃度のパラメータは、調査間のシーケンス識別の所望の面を達成して、核酸を目標とするために変化することができる。確認する。eに、g.Sambrookその他（1989）は、下に参照する、または、オースベルその他（1995）MolecularBiology、ジョン・ワイリー、NY、雑種形成状況の更なるガイダンスのために、NYのCurrent Protocols。

【００８６】

溶融する温度は、以下の公式によって記載されている（Beltz（G. A.その他）[1983] Methods ofEnzymology、R.呉語、L.グロスマン、そして、K.Moldave [編集]アカデミック・プレス（ニューヨーク100）：266-285)。

Tm=81.5o C +16.6 Log [Na+] +041（+G+C）-0。61（%ホルムアミド）-塩基対のデュプレックスの600/length。

【００８７】

洗浄は、次のように概して行われることができる：IX SSPE、0.1%の特別割引販売（低い厳しさ洗浄）および0.2X SSPE（0.1%の特別割引販売（適度な厳しさ洗浄））の15分の間のTM-20oCの一度の15分の間の室温の二度。

【００８８】

本発明に役立つ核酸は、Polymerase Chain Reactionによってつくられることができる（PCR）拡大。PCR、製品はアガロース・ゲル電気泳動によって確認されることができる。PCR核酸シーケンスの反復、酵素の、満たされた合成にある。この手順は、周知でこの技術に熟練しているそれらによって共通して使う（マリス、米国特許数字4,683、195,4、683,202および4、800、159を参照;佐伯ほか[1985の]サイエンス230：1350-1354).PCR酵素で目標シーケンスの対向するふさに交雑する2冊のオリゴヌクレオチド入門書が側面に並んでいる関心のDNA断片を拡大する用いる。下塗りは、各々の方へ指している3'endsによって正しい位置に置かれる。それらの補完的なシーケンスに下塗りの中でアニール化されて

、テンプレートの熱変性のサイクルを繰り返す、そして、部分の拡大のポリメラーゼ結果が5'endsによって定義したDNAを有する鍛えられた下塗りの拡張‖PCR下塗り。各々の下塗りの拡張製品が他の下塗りのためのテンプレートとして役立つことができるので、各サイクルは本質的に前のサイクルにおいて生産されるDNAテンプレートの量を２倍にする。２、３時間のいくつかの100万-折り目まで、これは特定の目標断片の指数蓄積に結果としてなる。を用いて耐熱である分離されるDNAポリメラーゼ（例えばTaqポリメラーゼ）から高温細菌aquaticusにThermus、拡大プロセスは、完全に自動化されることができる。使われることができる他の酵素は、当業者にとって公知である。

【００８９】

本発明のポリヌクレオチド・シーケンスは、頭を切って短くされることができておよび／または突然変異した最初のノーカット・シーケンスの断片および／または変異体がノーカット・シーケンスの所望の特徴を保持することができる結果になる確かな。より大きい核酸分子から断片を生成することに適している多種多様な規制酵素は、周知である。加えて、Bal31エキソヌクレアーゼがDNAの時間に制御された限られた消化のために、便利に使われることができることは、周知である。CloningしているManiatis（1982の）Molecular参照：Laboratoryマニュアル、コールドスプリングハーバーLaboratory、ニューヨーク、ページ135-139（ここにて組み込まれる）。また、Weiほか（1983の）J.Biol.Chemを参照。258:13006-13512.Bal31エキソヌクレアーゼ（共通に、as"erase-a-base"proceduresに関連した）を用いて、通常熟練した職人がヌクレオチドを取り除くことができることからどちらか、または、広い範囲の機能的に主題ヌクレオチド配列に等しい断片を生成する主題核酸の両方の端。当業者は、従来技術において、このように、最初の分子に沿って全て場所から制御、様々な長さの何百もの断片を生成することができる。本願明細書において教示されるにつれて、通常熟練した職人は通常試験することができてまたはそれらの特徴のための生成された断片を隠すことができて、断片の有用性を決めることができる。変異体シーケンスが容易に産出されることができることは、また、周知である‖部位特異的変異誘発。例えば、ラリォノフ、O.A.およびニキフォロフは知る ― V. G. Genetika 18の（1982の）(3)：349-59;そして、Shortle、D.ほか、（1981の）アンヌ。Rev.ジュネ。15:265-94（両方とも本願明細書に引用したものとする）。熟練した職人は、通常削除、挿入または置換-タイプ突然変異を起こすことができて、野生のタイプ・シーケンスの所望の特徴またはそれの断片を含むそれらの結果として生じる変異体を識別することができる。

【００９０】

2つの核酸のパーセント・シーケンス・アイデンティティは、カーリンてアルチュール（1990の）会報全米科学アカデミーUSA87のアルゴリズムを使用して決定されることができる：2264-2268（カーリンてアルチュール（1993の）会報全米科学アカデミーUSA 90に記載の修正される）：5873-5877.この種のアルゴリズムは、アルチュールほか（1990の）J.MolのNBLASTおよびXBLASTプログラムに組み込まれる。Biol.215：402-410.所望のパーセント・シーケンス識別を有するヌクレオチド配列を得るために、ヌクレオチドが検索するBLASTは、NBLASTプログラム、得点=100、wordlength = 12によって実行される。比較のためにすき間を作られた配置を得るために、GappedされたBLASTが、アルチュールほか（1997の）Nuclにて説明したように、使われる。酸。物。25：3389-3402.BLASTおよびGappedされたBLASTプログラムを利用するときに、それぞれのプログラム（NBLASTおよびXBLAST）のデフォルト・パラメータが使われる。http://www.ncbi. nih.政府を参照。

【００９１】

クローニング、DNA隔離、拡大および、DNAリガーゼを含んでいる酵素の反応、DNAポリメラーゼ、規制エンドヌクレアーゼなどのために、浄化のための標準の技術、そして、役立つさまざまな分離技術は、本願明細書において公知で、当業者によって共通に使用されるそれらである。多くの標準の技術は、Sambrookほか（1989の）MolecularCloning（Second Edition）に記載されているコールドスプリングハーバーLaboratory、Plainview、ニューヨーク;ClonningしているManiaisほか（1982の）Molecular、コールドスプリングハーバーLaboratory、Plainview、ニューヨーク;烏江（編集）（1993）メソジストEnzymol.21

8（パート１）;烏江（編集）（1979）メソジストEnzymol.68;呉語ほか（編集）（1983）メソジストEnzymol.100および101;グロスマン、そして、Moldave（編集）メソジストEnzymol.65;MolecularGeneticsのミラー（編集）（1972の）Experiments、コールドスプリングハーバーLaboratory、コールドスプリングハーバー（ニューヨーク）;GeneManipulation、カリフォルニア大学プレス、バークレーの古いPrimrose（1981の）Principles;Molecular Biologyのシュレイフてウェンシンク（1982の）PracticalMethods;第I巻およびIIをCloningしているグラヴァー（編集）（1985の）DNA、IRLプレス、オックスフォード、英国;故郷、そして、ヒギンズ（編集）（1985のNucleicなAcidHybridization、IRLプレス、オックスフォード、英国;SetlowおよびEngineeringしてGeneticなHollaender（1979）：原理およびMethods（Vols.）1-4、Plenumプレス、ニューヨーク;そして、MolecularBiology、グリーン/ワイリー、ニューヨーク、NYのオースベルその他.(1992) Current Protocols。使用される所で、省略形および命名法はフィールドにおいて標準であると考えられて、本願明細書において引用される人々のような専門のジャーナルにおいて共通して使う。

【００９２】

本発明は、データを有する機械可読の記憶装置およびプログラム記憶装置および方法をハプログループおよび生理的状況を診断することに提供する。本発明によって提供される1台のプログラム記憶装置は、プログラムステップを含む：サンプルの核酸からヌクレオチド配列データを使用している個人から、サンプルのハプログループを決定しているa);ハプログループを個人の地理的領域を識別している情報と関連させているb);ハプログループを比較しているc)および個人の地理的地域原産のハプログループの一組に対するサンプルの地理的領域;そして、個人のハプログループが個人の地理的地域原産のハプログループの一組の範囲内で、ない場合、エネルギーに対する素因をもつ個人を診断しているd)は生理的状態をmetabolism-relatedした;機械読める書式においてコード化されている全ての前記プログラムステップおよび機械読める書式においてコード化される全ての前記情報。本発明もデータセットを提供する。そして、人間の生理的状態を識別しているコード化された情報と関連する各々の対立遺伝子については、表19にリストされるヌクレオチド対立遺伝子を含んで、機械可読の形式においてコード化される。これらの生理的状況は、精力的なアンバランス、代謝疾患、異常なエネルギー代謝、異常な温度規則、異常な酸化的リン酸化、異常な電子伝達、肥満、体脂肪、糖尿病、高血圧および心血管疾患の量を含むenergy-代謝-関連した状況である。この記憶装置は、また、各々の対立遺伝子を一つ以上の自国の地理的地方と関連させている情報を含むことができる。本発明によって提供されるプログラム記憶装置は、個人のハプログループおよびその個人の地理的領域を入力するための入力手段を含んで、対立遺伝子を自国の地理的地方と関連させている情報および生理的状態に対する素因をもつ個人を診断するためのプログラムステップを含む。本発明によって提供される機械読める書式のデータセットを含んでいる記憶装置は、人間の生理的状態と関連する各々の対立遺伝子については、表19にリストされるアミノ酸対立遺伝子から成るコード化された情報を含むことができる。

【００９３】

それは、従来技術において通常の技術のそれらによって認められるその人口、部分母集団、細胞小器官、そして、アミンの酸性のおよびヌクレオチド配列比較方法（中性試験方法）ヌクレオチド塩基配列決定方法、コドン、サンプル、サンプル収集技術、サンプル準備技術、調査、調査生成技術、ミトコンドリア生物学に関係している遺伝子、雑種形成技術、技術を印刷している配列、生理的状況、細胞系、変異体圧力、有機体、組織、固体基板、機械可読の記憶装置、プログラム装置および特に本願明細書において開示されるそれら以外のデータ分析の方法は、従来技術において利用できて、本発明の実行において使用されることができる。全ての技術-周知の機能的な等価物は、本発明の範囲内で含まれることを目的とする。

【００９４】

ここで請求されるように、以下の実施例は説明の便宜上提供されて、本発明の範囲を制限することを目的としない。組成物および方法のいかなるバリエーションも、それを例証

した熟練した職人に起こる本発明の範囲内になる目的とする。
【実施例】
【００９５】
（実施例1）
本発明は、全ての主要人間のハプログループで見つかる人間のmtDNA多形性を提供する。表3は、mtDNAケンブリッジ・シーケンスと比較すると、103人の個人の完全なmtDNAシーケンスにおいて識別される自然に起こっているヌクレオチド対立遺伝子を示す。一覧を示さない全てのヌクレオチド配列は、ケンブリッジ・シーケンスと同一である。疾患状況（例えば表1にリストされるそれら）を伴うことはすでに公知のヌクレオチド対立遺伝子は、表3にリストされない。若干の削除または再配置多形性は、また、除外された。一覧を示す全ての多形性は、位置8271-8279の9-アデニン・ヌクレオチド削除を除いてヌクレオチド置換である。
【００９６】

【表３】

Table 3
Human MtDNA Nucleotide Alleles

nucleotide locus	Cambridge alleles	non-Cambridge alleles
64	C	T
72	T	C
73	A	G
89	T	C
93	A	G
95	A	C
114	C	T
143	G	A
146	T	C
150	C	T
151	C	T
152	T	C
153	A	G
171	G	A
180	T	C
182	C	T
183	A	G
185	G	A
185	G	T
186	C	A
189	A	C
189	A	G
194	C	T
195	T	A
195	T	C
198	C	T
199	T	C
200	A	G
204	T	C
207	G	A
208	T	C
210	A	G
212	T	C
215	A	G
217	T	C
225	G	A
227	A	G
228	G	A
235	A	G
236	T	C
247	G	A
250	T	C
252	T	C
263	A	G
291	A	G
295	C	T
297	A	G
316	G	A
317	C	A
317	C	G
320	C	T
325	C	T
340	C	T
357	A	G
373	A	G
400	T	G
408	T	A
418	C	T
456	C	T
462	C	T
465	C	T
467	C	T
471	T	C
480	T	C
482	T	C
489	T	C
493	A	G
499	G	A
508	A	G
593	T	C
597	C	T
663	A	G
678	T	C
680	T	C
709	G	A
710	T	C

721	T	C
750	A	G
769	G	A
825	T	A
827	A	G
850	T	C
921	T	C
930	G	A
961	T	C
961	T	G
1018	G	A
1041	A	G
1048	C	T
1119	T	C
1189	T	C
1243	T	C
1290	C	T
1382	A	C
1406	T	C
1415	G	A
1420	T	C
1438	A	G
1442	G	A
1503	G	A
1598	G	A
1700	T	C
1703	C	T
1706	C	T
1709	G	A
1715	C	T
1719	G	A
1736	A	G
1738	T	C
1780	T	C
1811	A	G
1888	G	A
1927	G	A
2000	C	T
2060	A	G
2092	C	T

2245	A	C
2245	A	G
2263	C	A
2308	A	G
2332	C	T
2352	T	C
2358	A	G
2380	C	T
2416	T	C
2483	T	C
2581	A	G
2639	C	T
2650	C	T
2706	A	G
2755	A	G
2758	G	A
2768	A	G
2789	C	T
2792	A	G
2834	C	T
2836	C	A
2857	T	C
2863	T	C
2885	T	C
3010	G	A
3083	T	C
3197	T	C
3200	T	A
3202	T	C
3204	C	T
3206	C	T
3221	A	G
3290	T	C
3308	T	C
3316	G	A
3372	T	C
3394	T	C
3438	G	A
3450	C	T
3480	A	G

3505	A	G
3513	C	T
3516	C	A
3516	C	G
3547	A	G
3549	C	T
3552	T	A
3552	T	C
3565	A	G
3594	C	T
3644	T	C
3666	G	A
3693	G	A
3699	C	T
3720	A	G
3756	A	G
3796	A	G
3796	A	T
3796	A	C
3808	A	G
3816	A	G
3834	G	A
3843	A	G
3847	T	C
3866	T	C
3918	G	A
3921	C	A
3927	A	G
3970	C	T
3981	A	G
4025	C	T
4040	C	T
4044	A	G
4048	G	A
4086	C	T
4104	A	G
4117	T	C
4122	A	G
4123	A	G
4158	A	G

4203	A	G
4216	T	C
4221	C	T
4225	A	G
4232	T	C
4248	T	C
4312	C	T
4336	T	C
4370	T	C
4388	A	G
4454	T	A
4491	G	A
4506	A	G
4508	C	T
4512	G	A
4529	A	C
4529	A	T
4541	G	A
4580	G	A
4586	T	C
4596	G	A
4646	T	C
4655	G	A
4688	T	C
4695	T	C
4715	A	G
4742	T	C
4767	A	G
4769	A	G
4820	G	A
4824	A	G
4833	A	G
4841	G	A
4883	C	T
4907	T	C
4917	A	G
4960	C	T
4977	T	C
4994	A	G
5004	T	C

5027	C	T
5036	A	G
5043	G	T
5046	G	A
5063	T	C
5096	T	C
5108	T	C
5147	G	A
5153	A	G
5178	C	A
5231	G	A
5237	G	A
5255	C	T
5262	G	A
5263	C	T
5285	A	G
5300	C	T
5330	C	A
5331	C	A
5390	A	G
5393	T	C
5417	G	A
5426	T	C
5442	T	C
5460	G	A
5465	T	C
5471	G	A
5492	T	C
5495	T	C
5580	T	C
5581	A	G
5601	C	T
5603	C	T
5606	C	T
5633	C	T
5655	T	C
5711	A	G
5773	G	A
5811	A	G
5814	T	C

5821	G	A
5826	T	C
5843	A	G
5951	A	G
5984	A	G
5987	C	T
6026	G	A
6029	C	T
6045	C	T
6071	T	C
6077	C	T
6104	C	T
6150	G	A
6152	T	C
6164	C	T
6167	T	C
6182	G	A
6185	T	C
6185	T	C
6221	T	C
6227	T	C
6253	T	C
6257	G	A
6324	G	A
6366	G	A
6371	C	T
6392	T	C
6473	C	T
6491	C	A
6524	T	C
6548	C	T
6587	C	T
6607	T	C
6680	T	C
6713	C	T
6719	T	C
6734	G	A
6752	A	G
6770	A	G
6776	T	C

6815	T	C
6827	T	C
6875	C	A
6938	C	T
6962	G	A
6989	A	G
7028	C	T
7052	A	G
7055	A	G
7058	T	A
7076	A	G
7146	A	G
7154	A	G
7175	T	C
7196	C	A
7202	A	G
7226	G	A
7256	C	T
7257	A	G
7271	A	G
7274	C	T
7319	T	C
7337	G	A
7347	G	A
7389	T	C
7403	A	G
7424	A	G
7444	G	A
7476	C	T
7493	C	T
7521	G	A
7561	T	C
7571	A	G
7600	G	A
7624	T	A
7645	T	C
7648	C	T
7660	T	C
7664	G	A
7673	A	G

7675	C	T
7693	C	T
7694	C	T
7697	G	A
7744	T	C
7765	A	G
7768	A	G
7771	A	G
7858	C	T
7861	T	C
7864	C	T
7867	C	T
7933	A	G
7948	C	T
7999	T	C
8014	A	G
8020	G	A
8027	G	A
8080	C	T
8087	T	C
8113	C	A
8142	C	T
8149	A	G
8152	G	A
8155	G	A
8185	T	C
8200	T	C
8206	G	A
8248	A	G
8251	G	A
8260	T	C
8269	G	A
8271-8279	A	DEL
8286	T	C
8292	G	A
8298	T	C
8344	A	G
8387	G	A
8389	A	G
8392	G	A

8404	T	C
8414	C	T
8428	C	T
8448	T	C
8460	A	G
8468	C	T
8472	C	T
8473	T	C
8485	G	A
8545	G	A
8553	C	T
8563	A	G
8566	A	G
8577	A	G
8584	G	A
8618	T	C
8655	C	T
8697	G	A
8701	A	G
8703	C	T
8705	T	C
8709	C	T
8721	A	G
8733	T	C
8764	G	A
8781	C	A
8784	A	G
8790	G	A
8793	T	C
8794	C	T
8805	A	G
8836	A	G
8838	G	A
8856	G	A
8860	A	G
8875	T	C
8877	T	C
8911	T	C
8913	A	G
8928	T	C

8943	C	T
8962	A	G
8994	G	A
9042	C	T
9053	G	A
9055	G	A
9072	A	G
9077	T	C
9090	T	C
9093	A	C
9103	T	C
9114	A	G
9120	A	G
9123	G	A
9123	G	A
9136	A	G
9151	A	G
9156	A	G
9174	T	C
9221	A	G
9237	G	A
9242	A	G
9248	C	T
9263	A	G
9272	C	T
9296	C	T
9311	T	C
9325	T	C
9335	C	T
9347	A	G
9355	A	G
9356	C	T
9377	A	G
9402	A	C
9449	C	T
9456	A	G
9477	G	A
9509	T	C
9536	C	T
9540	T	C

9545	A	G
9548	G	A
9554	G	A
9559	C	G
9575	G	A
9591	G	A
9599	C	T
9632	A	G
9647	T	C
9667	A	G
9682	T	C
9698	T	C
9755	G	A
9818	C	T
9822	C	A
9824	T	A
9911	C	T
9932	G	A
9950	T	C
9957	T	C
9966	G	A
9977	T	C
10034	T	C
10086	A	G
10086	A	C
10115	T	C
10118	T	C
10142	C	T
10151	A	G
10152	G	C
10172	G	A
10182	G	C
10197	G	A
10238	T	C
10253	T	C
10256	T	C
10310	G	A
10313	A	G
10321	T	C
10325	G	A

10358	A	G
10370	T	C
10398	A	G
10400	C	T
10410	T	C
10414	G	T
10427	G	A
10463	T	C
10499	A	G
10505	T	C
10550	A	G
10586	G	A
10589	G	A
10609	T	C
10637	C	T
10640	T	C
10646	G	A
10659	C	T
10664	C	T
10667	T	C
10688	G	A
10736	C	T
10790	T	C
10792	A	G
10793	C	T
10804	A	G
10810	T	C
10819	A	G
10828	T	C
10873	T	C
10876	A	G
10894	C	T
10915	T	C
10920	C	T
10939	C	T
10966	T	C
10984	C	G
11002	A	G
11016	G	A
11017	T	C

11023	A	G
11078	A	G
11092	A	G
11147	T	C
11150	G	A
11167	A	G
11172	A	G
11176	G	A
11177	C	T
11215	C	T
11251	A	G
11257	C	T
11296	C	T
11299	T	C
11332	C	T
11362	A	G
11365	T	C
11377	G	A
11467	A	G
11476	C	T
11536	C	T
11590	A	G
11611	G	A
11641	A	G
11653	A	G
11654	A	G
11674	C	T
11701	T	C
11719	G	A
11722	T	C
11767	C	T
11812	A	G
11854	T	C
11884	A	G
11887	G	A
11893	A	G
11899	T	C
11909	A	G
11914	G	A
11944	T	C

11947	A	G
11959	A	G
11963	G	A
11969	G	A
12007	G	A
12049	C	T
12070	G	A
12083	T	G
12121	T	C
12134	T	C
12153	C	T
12172	A	G
12175	T	C
12234	A	G
12236	G	A
12239	C	T
12248	A	G
12308	A	G
12346	C	T
12358	A	G
12361	A	G
12372	G	A
12373	A	G
12397	A	G
12406	G	A
12414	T	C
12477	T	C
12501	G	A
12507	A	G
12519	T	C
12528	G	A
12540	A	G
12612	A	G
12630	G	A
12633	C	T
12635	T	C
12669	C	T
12672	A	G
12693	A	G
12705	C	T

12720	A	G
12738	T	C
12768	A	G
12771	G	A
12810	A	G
12822	A	G
12850	A	G
12879	T	C
12882	C	T
12930	A	T
12940	G	A
12948	A	G
12967	A	C
12972	A	G
12999	A	G
13020	T	C
13059	C	T
13068	A	G
13101	A	C
13104	A	G
13105	A	G
13135	G	A
13143	T	C
13145	G	A
13149	A	G
13194	G	A
13197	C	T
13212	C	T
13221	A	G
13263	A	G
13276	A	G
13281	T	C
13368	G	A
13440	C	G
13477	G	A
13485	A	G
13494	C	T
13500	T	C
13500	T	G
13506	C	T

13512	A	G
13563	A	G
13590	G	A
13594	A	G
13602	T	C
13611	A	G
13617	T	C
13641	T	C
13650	C	T
13651	A	G
13660	A	G
13708	G	A
13722	A	G
13734	T	C
13759	G	A
13780	A	G
13789	T	C
13803	A	G
13812	T	C
13818	T	C
13819	T	C
13827	A	G
13880	C	A
13886	T	C
13914	C	A
13924	C	T
13927	A	T
13928	G	C
13958	G	C
13965	T	C
13966	A	G
13980	G	A
14000	T	A
14016	G	A
14020	T	C
14022	A	G
14025	T	C
14034	T	C
14059	A	G
14070	A	T

14070	A	G
14088	T	C
14094	T	C
14097	C	T
14118	A	G
14128	A	G
14148	A	G
14152	A	G
14167	C	T
14178	T	C
14182	T	C
14200	T	C
14203	A	G
14209	A	G
14212	T	C
14215	T	C
14221	T	C
14233	A	G
14272	C	G
14284	C	T
14308	T	C
14311	T	C
14318	T	C
14319	T	C
14371	T	C
14374	T	C
14384	G	C
14455	C	T
14459	G	A
14470	T	C
14484	T	C
14488	T	C
14502	T	C
14560	G	A
14566	A	G
14569	G	A
14571	T	A
14580	A	G
14587	A	G
14605	A	G

14668	C	T
14693	A	G
14766	C	T
14769	A	G
14783	T	C
14793	A	G
14798	T	C
14812	C	T
14836	A	G
14861	G	A
14862	C	T
14905	G	A
14911	C	T
14971	T	C
14974	C	G
14979	T	C
15016	C	T
15034	A	G
15043	G	A
15110	G	A
15113	A	G
15115	T	C
15136	C	T
15172	G	A
15204	T	C
15217	G	A
15218	A	C
15229	T	C
15238	C	G
15244	A	G
15257	G	A
15261	G	A
15301	G	A
15317	G	A
15318	C	T
15323	G	A
15326	A	G
15346	G	A
15358	A	G
15431	G	A

15442	A	G
15452	C	A
15466	G	A
15470	T	C
15487	A	T
15497	G	A
15514	T	C
15519	T	C
15535	C	T
15607	A	G
15626	C	T
15629	T	C
15646	C	T
15661	C	T
15663	T	C
15670	T	C
15724	A	G
15731	G	A
15746	A	G
15766	A	G
15784	T	C
15793	C	T
15803	G	A
15806	G	A
15812	G	A
15824	A	G
15833	C	T
15849	C	T
15884	G	C
15900	T	C
15904	C	T
15907	A	G
15924	A	G
15927	G	A
15928	G	A
15930	G	A
15932	T	C
15939	C	T
15941	T	C
15942	T	C

15968	T	C
16017	T	C
16038	A	G
16051	A	G
16069	C	T
16071	C	T
16075	T	C
16086	T	C
16093	T	C
16108	C	T
16111	C	T
16114	C	A
16124	T	C
16126	T	C
16129	G	A
16129	G	C
16140	T	C
16144	T	C
16145	G	A
16147	C	T
16148	C	T
16153	G	A
16162	A	G
16163	A	C
16166	A	C
16167	C	T
16168	C	T
16169	C	T
16171	A	G
16172	T	C
16175	A	G
16176	C	T
16182	A	C
16183	A	C
16184	C	T
16185	C	T
16186	C	T
16187	C	T
16188	C	A
16188	C	G

16189	T	C
16192	C	T
16193	C	T
16207	A	G
16209	T	C
16212	A	G
16213	G	A
16214	C	T
16217	T	C
16219	A	G
16223	C	T
16224	T	C
16227	A	G
16229	T	C
16230	A	G
16231	T	C
16232	C	T
16234	C	T
16235	A	G
16239	C	T
16241	A	G
16242	C	T
16243	T	C
16245	C	T
16247	A	G
16249	T	C
16254	A	C
16255	G	A
16256	C	T
16257	C	T
16258	A	G
16260	C	T
16261	C	T
16264	C	T
16265	A	C
16266	C	T
16268	C	T
16270	C	T
16271	T	C
16274	G	A

16278	C	T
16284	A	G
16286	C	G
16287	C	T
16288	T	C
16290	C	T
16291	C	T
16292	C	T
16293	A	G
16294	C	T
16296	C	T
16298	T	C
16304	T	C
16309	A	G
16311	T	C
16316	A	G
16317	A	T
16318	A	T
16319	G	A
16320	C	T
16324	T	C
16325	T	C
16326	A	G
16327	C	T
16343	A	G
16344	C	T
16354	C	T
16355	C	T
16356	T	C
16357	T	C
16360	C	T
16362	T	C
16366	C	T
16368	T	C
16390	G	A
16391	G	A
16399	A	G
16438	G	A
16439	C	A
16483	G	A

16519	T	C
16527	C	T

表4は、ケンブリッジ・シーケンスと比較して、48のミトコンドリア・ゲノムにおいて識別されるヌクレオチド対立遺伝子の一覧を示す。

【００９７】

【表４】

Table 4

Human MtDNA Nucleotide Alleles in 48 Genomes

nucleotide locus	Cambridge alleles	non-Cambridge alleles
64	C	T
72	T	C
73	A	G
89	T	C
93	A	G
95	A	C
114	C	T
146	T	C
150	C	T
151	C	T
152	T	C
153	A	G
171	G	A
180	T	C
182	C	T
185	G	A
185	G	T
186	C	A
189	A	C
194	C	T
195	T	C
198	C	T
199	T	C
200	A	G
204	T	C
207	G	A
210	A	G
217	T	C
225	G	A
227	A	G
228	G	A
235	A	G
236	T	C
247	G	A
250	T	C
263	A	G
295	C	T
297	A	G
316	G	A
317	C	G
320	C	T
325	C	T
340	C	T
357	A	G
400	T	G
418	C	T
456	C	T
462	C	T
467	C	T
482	T	C
489	T	C
493	A	G
499	G	A
508	A	G
597	C	T
663	A	G
680	T	C
709	G	A
710	T	C
750	A	G
769	G	A
825	T	A
827	A	G
921	T	C
930	G	A
961	T	C
961	T	G
1018	G	A
1048	C	T
1189	T	C
1243	T	C
1290	C	T
1406	T	C
1415	G	A
1438	A	G
1442	G	A
1598	G	A
1700	T	C

1703	C	T
1706	C	T
1709	G	A
1715	C	T
1719	G	A
1736	A	G
1738	T	C
1780	T	C
1811	A	G
1888	G	A
2092	C	T
2245	A	C
2245	A	G
2308	A	G
2332	C	T
2352	T	C
2358	A	G
2416	T	C
2581	A	G
2639	C	T
2706	A	G
2758	G	A
2768	A	G
2789	C	T
2834	C	T
2857	T	C
2885	T	C
3010	G	A
3083	T	C
3197	T	C
3200	T	A
3202	T	C
3221	A	G
3308	T	C
3316	G	A
3394	T	C
3450	C	T
3480	A	G
3505	A	G
3516	C	A

3516	C	G
3547	A	G
3552	T	A
3552	T	C
3565	A	G
3594	C	T
3644	T	C
3666	G	A
3693	G	A
3720	A	G
3756	A	G
3796	A	G
3796	A	T
3796	A	C
3808	A	G
3816	A	G
3834	G	A
3843	A	G
3847	T	C
3866	T	C
3921	C	A
3970	C	T
3981	A	G
4025	C	T
4040	C	T
4044	A	G
4086	C	T
4104	A	G
4122	A	G
4123	A	G
4158	A	G
4216	T	C
4221	C	T
4225	A	G
4232	T	C
4248	T	C
4312	C	T
4336	T	C
4370	T	C
4454	T	A

4529	A	C
4529	A	T
4580	G	A
4586	T	C
4596	G	A
4646	T	C
4715	A	G
4767	A	G
4769	A	G
4820	G	A
4824	A	G
4833	A	G
4841	G	A
4883	C	T
4907	T	C
4917	A	G
4960	C	T
4977	T	C
5027	C	T
5036	A	G
5043	G	T
5046	G	A
5096	T	C
5108	T	C
5147	G	A
5153	A	G
5178	C	A
5231	G	A
5300	C	T
5331	C	A
5390	A	G
5393	T	C
5417	G	A
5426	T	C
5442	T	C
5460	G	A
5465	T	C
5471	G	A
5495	T	C
5581	A	G

5601	C	T
5603	C	T
5606	C	T
5633	C	T
5711	A	G
5773	G	A
5814	T	C
5951	A	G
5984	A	G
6026	G	A
6029	C	T
6045	C	T
6071	T	C
6152	T	C
6185	T	C
6221	T	C
6227	T	C
6257	G	A
6371	C	T
6392	T	C
6473	C	T
6491	C	A
6607	T	C
6680	T	C
6713	C	T
6734	G	A
6752	A	G
6776	T	C
6815	T	C
6827	T	C
6962	G	A
6989	A	G
7028	C	T
7052	A	G
7055	A	G
7146	A	G
7154	A	G
7175	T	C
7196	C	A
7256	C	T

7271	A	G
7274	C	T
7389	T	C
7424	A	G
7476	C	T
7521	G	A
7561	T	C
7600	G	A
7624	T	A
7664	G	A
7694	C	T
7765	A	G
7771	A	G
7864	C	T
7867	C	T
7933	A	G
7999	T	C
8027	G	A
8080	C	T
8087	T	C
8113	C	A
8142	C	T
8149	A	G
8152	G	A
8155	G	A
8185	T	C
8200	T	C
8206	G	A
8248	A	G
8251	G	A
8260	T	C
8269	G	A
8271-8279	A	DEL
8286	T	C
8298	T	C
8344	A	G
8387	G	A
8389	A	G
8392	G	A
8414	C	T

8428	C	T
8448	T	C
8460	A	G
8468	C	T
8472	C	T
8545	G	A
8553	C	T
8563	A	G
8566	A	G
8584	G	A
8618	T	C
8655	C	T
8697	G	A
8701	A	G
8705	T	C
8709	C	T
8721	A	G
8790	G	A
8794	C	T
8836	A	G
8856	G	A
8860	A	G
8875	T	C
8913	A	G
8962	A	G
8994	G	A
9042	C	T
9053	G	A
9055	G	A
9072	A	G
9077	T	C
9090	T	C
9093	A	C
9114	A	G
9120	A	G
9123	G	A
9151	A	G
9221	A	G
9237	G	A
9325	T	C

9335	C	T
9347	A	G
9355	A	G
9377	A	G
9402	A	C
9449	C	T
9456	A	G
9477	G	A
9540	T	C
9545	A	G
9548	G	A
9559	C	G
9575	G	A
9632	A	G
9682	T	C
9698	T	C
9755	G	A
9818	C	T
9822	C	A
9911	C	T
9932	G	A
9950	T	C
9957	T	C
9966	G	A
10034	T	C
10086	A	G
10086	A	C
10115	T	C
10151	A	G
10152	G	C
10172	G	A
10182	G	C
10238	T	C
10256	T	C
10310	G	A
10321	T	C
10325	G	A
10398	A	G
10400	C	T
10414	G	T
10463	T	C
10550	A	G
10586	G	A
10589	G	A
10609	T	C
10637	C	T
10646	G	A
10659	C	T
10664	C	T
10688	G	A
10790	T	C
10810	T	C
10828	T	C
10873	T	C
10876	A	G
10915	T	C
10966	T	C
10984	C	G
11002	A	G
11078	A	G
11092	A	G
11147	T	C
11167	A	G
11172	A	G
11176	G	A
11177	C	T
11215	C	T
11251	A	G
11257	C	T
11299	T	C
11332	C	T
11362	A	G
11377	G	A
11467	A	G
11476	C	T
11536	C	T
11590	A	G
11641	A	G
11674	C	T
11719	G	A

11767	C	T
11812	A	G
11854	T	C
11899	T	C
11914	G	A
11944	T	C
11947	A	G
11969	G	A
12007	G	A
12083	T	G
12121	T	C
12172	A	G
12234	A	G
12236	G	A
12308	A	G
12358	A	G
12361	A	G
12372	G	A
12373	A	G
12397	A	G
12406	G	A
12414	T	C
12501	G	A
12507	A	G
12519	T	C
12528	G	A
12540	A	G
12612	A	G
12633	C	T
12669	C	T
12672	A	G
12693	A	G
12705	C	T
12720	A	G
12738	T	C
12810	A	G
12822	A	G
12882	C	T
12930	A	T
12948	A	G

12967	A	C
12972	A	G
13020	T	C
13068	A	G
13101	A	C
13104	A	G
13105	A	G
13194	G	A
13263	A	G
13276	A	G
13368	G	A
13440	C	G
13485	A	G
13494	C	T
13500	T	G
13506	C	T
13512	A	G
13563	A	G
13590	G	A
13617	T	C
13650	C	T
13708	G	A
13734	T	C
13759	G	A
13780	A	G
13789	T	C
13803	A	G
13812	T	C
13827	A	G
13880	C	A
13886	T	C
13914	C	A
13924	C	T
13928	G	C
13958	G	C
13966	A	G
14000	T	A
14016	G	A
14034	T	C
14059	A	G

14070	A	G
14088	T	C
14118	A	G
14128	A	G
14148	A	G
14167	C	T
14178	T	C
14200	T	C
14203	A	G
14215	T	C
14221	T	C
14233	A	G
14272	C	G
14284	C	T
14308	T	C
14318	T	C
14374	T	C
14459	G	A
14470	T	C
14484	T	C
14488	T	C
14502	T	C
14560	G	A
14566	A	G
14569	G	A
14668	C	T
14693	A	G
14766	C	T
14783	T	C
14793	A	G
14798	T	C
14836	A	G
14861	G	A
14905	G	A
14911	C	T
14974	C	G
15034	A	G
15043	G	A
15110	G	A
15115	T	C

15136	C	T
15172	G	A
15204	T	C
15217	G	A
15218	A	C
15238	C	G
15257	G	A
15261	G	A
15301	G	A
15317	G	A
15318	C	T
15323	G	A
15326	A	G
15431	G	A
15442	A	G
15452	C	A
15466	G	A
15487	A	T
15497	G	A
15519	T	C
15535	C	T
15607	A	G
15661	C	T
15724	A	G
15766	A	G
15784	T	C
15793	C	T
15806	G	A
15812	G	A
15824	A	G
15833	C	T
15849	C	T
15884	G	C
15900	T	C
15904	C	T
15907	A	G
15924	A	G
15928	G	A
15930	G	A
15939	C	T

15941	T	C
15968	T	C
16017	T	C
16051	A	G
16069	C	T
16086	T	C
16093	T	C
16108	C	T
16111	C	T
16114	C	A
16124	T	C
16126	T	C
16129	G	A
16129	G	C
16145	G	A
16148	C	T
16153	G	A
16162	A	G
16163	A	C
16167	C	T
16168	C	T
16172	T	C
16176	C	G
16182	A	C
16183	A	C
16184	C	T
16185	C	T
16186	C	T
16187	C	T
16188	C	A
16188	C	G
16189	T	C
16192	C	T
16193	C	T
16212	A	G
16213	G	A
16214	C	T
16217	T	C
16219	A	G
16223	C	T
16224	T	C
16227	A	G
16229	T	C
16230	A	G
16231	T	C
16232	C	T
16234	C	T
16235	A	G
16239	C	T
16243	T	C
16245	C	T
16249	T	C
16254	A	C
16255	G	A
16256	C	T
16258	A	G
16260	C	T
16261	C	T
16264	C	T
16265	A	C
16266	C	T
16270	C	T
16274	G	A
16278	C	T
16284	A	G
16290	C	T
16291	C	T
16292	C	T
16293	A	G
16294	C	T
16296	C	T
16298	T	C
16304	T	C
16309	A	G
16311	T	C
16317	A	T
16318	A	T
16319	G	A
16320	C	T
16325	T	C

16327	C	T
16355	C	T
16356	T	C
16360	C	T
16362	T	C
16366	C	T
16368	T	C
16390	G	A
16391	G	A
16399	A	G
16519	T	C

（実施例2）

彼らが人間の主要ハプログループ血統の全てを代表するので、実施例1のmtDNAシーケンスは選択された。これらのシーケンスの分析は、再確認された非常に全て人間のmtDNAsが単一の母性木に帰属する効果があって、アフリカにおいて根づかせた（R.L.キャンほか（ネイチャー325）：31-36 (1987);M. J.ジョンソンほか（MoleczelarEvolution 19（1983の）ジャーナル）：255-271;D.C.ウォレスその他、グローバルなMitochondrial DNA Variation、そして、Native Americans"inのOriginHumankind、M. Aloisi、B.バッタリア、E. Carafoli、G. A.ダニエーリ、Eds.のOrigin（ベニス（IOSプレス（2000）））;M.イングマンほか（（2000の）ネイチャー408）：708-13;そして、D.C.Wallaceほか（（1999の）Gene 238）：211-230).これらのmtDNAシーケンスの分岐図は、図1に示される。ハプログループは、木の枝に示される。シーケンス進化の較正は、mtDNAの符号化領域のために、評価して、650万年前（MYA）の人間のチンパンジ相違時間に、基礎をおいた（M.グッドマンその他、（1998の）MolPhylogenet。Evol.9:585-98）、人間のmtDNA系統発生at-200の最も最近の普通の原型（MRCA）の時間の評価をできるようにする、000は何年も前に、現れる（YBP）、そして、各々の主要ハプログループ（表5）のためのMRCAの時間の評価。

【００９８】

【表５】

Table 5 Coalescence dates for haplogroups*

Haplogroup	Sample sizes	Time to MRCA ± s.e. ($\times 10^{-4}$ mutations per np) [a]	Time to MRCA ± s.e. ($\times 10^{3}$ years) [b]
chimp+human	1 + 104	818.05 ± 0.75	6,500
humans	104	24.88 ± 0.90	198 ± 19
L0	8	17.92 ± 1.87	142 ± 17
L1	9	17.81 ± 1.77	142 ± 17
L2	7	11.57 ± 1.30	91.9 ±11.8
N	50	8.09 ± 0.53	64.3 ± 5.8
A	4	4.06 ± 0.92	32.3 ± 7.6
R	37	7.66 ± 0.51	60.9 ± 5.5
HV	15	3.61 ± 0.73	28.7 ± 6.1
H	11	2.40 ± 0.40	19.1 ± 3.4
V	3	1.71 ± 0.60	13.6 ± 4.8
JT	7	6.29 ± 0.74	50.0 ± 6.7
J	4	4.33 ± 0.87	34.4 ± 7.2
T	3	1.40 ± 0.55	11.1 ± 4.4
U	4	6.51 ± 0.66	51.7 ± 6.2
M	22	8.15 ± 0.74	64.8 ± 7.1
CZ	10	5.91 ± 0.87	47.0 ± 7.6
C	9	3.56 ± 0.65	28.3 ± 5.5
D	6	4.19 ± 0.67	33.3 ± 5.7
G	3	4.75 ± 0.93	37.7 ± 7.8

* The high probability of reverse mutations in the control region led us to calculate the times to the MRCAs using the entire mtDNA, excluding the control region (np 577-16023).
[a] Based on this value we estimated the average sequence evolution rate as $(1.26 \pm 0.08) \times 10^{-8}$ per nucleotide per year, using the HKY85 model (M. Hasegawa et al., (1985) *J Mol. Evol.* 22:160-74 (1985)).
[b] Standard errors calculated from the inverse hessian at the maximum of the likelihood do not include any uncertainty in the calibration point, and were calculated using the delta method. The coalescence times of the various haplogroups may well be underestimated because of their small sample size.

*我々を導かれる制御領域の復帰突然変異の高い確率は、制御領域（np577-16023）を除いて、全てのmtDNAを使用しているMRCAsに、時、算出する。に基づいて我々が平均シーケンス進化率を推定したこの値（1.26 0。08）HKY85モデルを使用して年につきヌクレオチドにつきx10-（M.ハセガワほか（（1985の）Mol envol 22）：160-74（1985））。可能性の最大で逆の麻布から算出されるbスタンダード・エラーは、いかなる不確実性も較正位置に含まなくて、デルタ方法を使用して算出された。さまざまなハプログループの合併時間は、それらの少ないサンプル・サイズを原因として生じるので過小評価されるだろう。

【００９９】

（実施例3：Inter-Cotttinentalな創設者イベント）

mtDNA木の最も著しい特徴は、1つの大陸からもう一方への移行と関係しているmtDNA血統の数の著しい減少である。例えば、人間がアフリカからユーラシアまで引っ越すときに、ミトコンドリア血統の数は多数から2つの血統まで減少した。北東部アフリカがアフリカのハプログループL0-L2だけから、ヨーロッパでアジアのmtDNA血統の原種まで、アフリカのmtDNAバリエーションの全ての範囲を含むと共に、2つのアフリカのmtDNA血統、マクロハプログループMおよびN（それは65,000のYBPについて起こった）だけはアフリカにユーラシアにコロニーを作るのを任せた。さらに、マクロハプログループ MおよびNのMRCAsの時代は副マクロハプログループRと同様に類似している。そして、ユーラシアの植民地化と関連する迅速な人口展開を提案する。

【０１００】

同様に、人間が後で中央アジアから米州まで移動するときに、血統の数は多数乃至約5から再び減少した。アジアの多様性、それでも、この多様性が実質的に減少する大きなmtDNAがシベリアおよび5つのmtDNA ハプログループ（A、B、C、DおよびX）（それは28,000-34についてアジアにおいて起こった）だけであってそこで、000YBPは、うまくアメリカを占領するためにベーリング陸橋を渡った。人間のmtDNA ハプログループ移動は、図2において表される。

【０１０１】

（実施例4）

更なる分析は、どの対立遺伝子が主要ハプログループを記述していて、副ハプログループを選んで、マクロハプログループを選んだかについて示した。mtDNAヌクレオチド位置および関連した対立遺伝子は、図3に示される。データは分岐図として配置される。そうすると、左上のグループはその権利に対立遺伝子の全てを含む。垂直棒は、棒の右側に対する対立遺伝子が棒の左にグループの全てに存在することを示す。図3のハプログループデータは、表6および7にまとめられる。下位のハプログループデータは、表8および9にまとめられる。各々のグループは、それの下で一覧を示す対立遺伝子を含む。

【０１０２】

【表６】

Table 6

L0	L1	L2	L3	C	D	E	G	Z
1048T	2352C	325T	2352C	3552C	4883T	16227G	4833G	11078G
3516A	3796C	680C	8618C	4715G	5178A		8200C	16185T
4312T	5951G	2416C	10086C	7196A	8414T		16017C	16224C
4586C	5984G	2758G	10398A	8584A	14668T		16129A	16260T
5442C	6071C	4158G	10819G	9545G	15487T			
6185C	9072G	8206A	14212C	13263G	16362C	16362C		
8113A	10586A	9221G	16124C	14318C				
8251A	12810G	11944C	16278T	16298C				16298C
9347G	13485G	13803G	16362C	16327T				
9402C	3666A	13958C		489C	489C	489C	489C	489C
9818T	7055G	16278T		10400T	10400T	10400T	10400T	10400T
10589A	7389C	16390G		14783C	14783C	14783C	14783C	14783C
10664T	13789C			15043A	15043A	15043A	15043A	15043A
10915C	14178C	15301A	15301A	15301A	15301A	15301A	15301A	15301A
12007A								
13276G								
13506T								
825A	825A							
2758A	2758A							
2885C	2885C							
7146G	7146G							
8468T	8468T							
8655T	8655T							
10688A	10688A							
10810C	10810C							
13105G	13105G							
769A	769A							
1018A	1018A							
3594T	3594T							
4104G	4104G							
7256T	7256T							
7521A	7521A							
13650T	13650T							

【０１０３】

【表７】

Table 7

A	I	W	X	B	F	Y	V	H	U	J	T
663G	4529T	204C	1719A		12406A	7933G	72C	2706A	3197C	295T	11812G
16290T	10034C	207A	3516G		16304C	8392A	4580A	7028C	4646C	489C	12633T
16319A	16129A	1243C	6221C			16126C	15904T		7768G	12612G	14233G
	16391A	5046A	14470C			16231C	16298C		9055A	13708A	16163C
		5460A	16189C	16189C		16266T	73A	73A	11332T	16069T	16186T
		8251A	16278T				11719G	11719G	11467G		16189C
		8994A					14766C	14766C	12308G		1888A
		11947G							12372A		4917G
		15884C							13104G		8697A
		16292T							14070G		10463C
									15907G		13368A
									16051G		14905A
									16129C		15607G
									16172C		15928A
									16189C		16294T
									16219G		
									16224C	4216C	4216C
									16249C	11251G	11251G
									16270T	15452A	15452A
									16311T	16126C	16126C
									16318T		
									16343G		
									16356C		
				12705C	12705C	12705C	12705C	12705C	12705C	12705C	12705C
				16223C	16223C	16223C	16223C	16223C	16223C	16223C	16223C
8701A	8701A	8701A	8701A	8701A	8701A	8701A	8701A	8701A	8701A	8701A	8701A
9540T	9540T	9540T	9540T	9540T	9540T	9540T	9540T	9540T	9540T	9540T	9540T
10873T	10873T	10873T	10873T	10873T	10873T	10873T	10873T	10873T	10873T	10873T	10873T

【０１０４】

【表８】

Table 8

L1a1	L1a2	L1b1	L1b2	L2a	L2b	L2c	L3a	L3b	L3c	L3d
4586C	8113A	2352C	3796C	13803G	4158G	325T	23252C	8618C	10086C	10398
9818T	8251A		5951G			680C	10819G	16124C	16124C	
1048T	1048T		5984G			13958C	14212C		16278T	
3516A	3516A		6071C	2416C	2416C	2416C			16362C	
4312T	4312T		9072G	2758G	2758G	2758G				
5442C	5442C		10586A	8206A	8206A	8206A				
6185C	6185C		12810G	9221G	9221G	9221G				
9402C	9402C		13485G	11944C	11944C	11944C				
9347G	9347G	3666A	3666A	16278T	16278T	16278T				
10589A	10589A	7055G	7055G	16390G	16390G	16390G				
10664T	10664T	7389C	7389C	15301A	15301A	15301A	15301A	15301A	15301A	15301A
10915C	10915C	13789C	13789C							
12007A	12007A	14178C	14178C							
13276G	13276G									
13506T	13506T									
825A	825A	825A	825A							
2758A	2758A	2758A	2758A							
2885C	2885C	2885C	2885C							
7146G	7146G	7146G	7146G							
8468T	8468T	8468T	8468T							
8655T	8655T	8655T	8655T							
10688A	10688A	10688A	10688A							
10810C	10810C	10810C	10810C							
13105G	13105G	13105G	13105G							
769A	769A	769A	769A	769A	769A	769A				
1018A	1018A	1018A	1018A	1018A	1018A	1018A				
3594T	3594T	3594T	3594T	3594T	3594T	3594T				
4104G	4104G	4104G	4104G	4104G	4104G	4104G				
7256T	7256T	7256T	7256T	7256T	7256T	7256T				
7521A	7521A	7521A	7521A	7521A	7521A	7521A				
13650T	13650T	13650T	13650T	13650T	13650T	13650T				

【０１０５】

【表９】

Table 9

UK	U7	U6	U5	U4	U3	U2	U1	T*	T1
9055A	16318T	16172C	3197C	4646C	16343G	15907G	13104G	11812G	12633T
16224C		16219G	7768G	11332T		16051G	14070G	14233G	16163C
16311T			16270T	16356C		16129C	16189C		16186T
							16249C		16189C
11467G	11467G	11467G	11467G	11467G	11467G	11467G	11467G	1888A	1888A
12308G	12308G	12308G	12308G	12308G	12308G	12308G	12308G	4917G	4917G
12372A	12372A	12372A	12372A	12372A	12372A	12372A	12372A	8697A	8697A
								10463C	10463C
								13368A	13368A
								14905A	14905A
								15607G	15607G
								15928A	15928A
								16294T	16294T
								4216C	4216C
								11251G	11251G
								15452A	15452A
								16126C	16126C
12705C	12705C	12705C	12705C	12705C	12705C	12705C	12705C	12705C	12705C
16223C	16223C	16223C	16223C	16223C	16223C	16223C	16223C	16223C	16223C
8701A	8701A	8701A	8701A	8701A	8701A	8701A	8701A	8701A	8701A
9540T	9540T	9540T	9540T	9540T	9540T	9540T	9540T	9540T	9540T
10873T	10873T	10873T	10873T	10873T	10873T	10873T	10873T	10873T	10873T

（実施例5）

図3のデータの更なる分析は、ハプログループを診断することに役立つヌクレオチド対立遺伝子の集合を示した。ハプログループの全ておよび図3の副ハプログループを診断することに役立つヌクレオチド対立遺伝子の一組は、表10にリストされる。ハプログループを診断する多くの等価な方法が、ある。テストだけまたは2、3の場所だけを必要としている方法の実施例は、あとに続く。対立遺伝子は、mtDNAを含んでいる人間のサンプルにおいて識別される。ハプログループL0は、4586C、9818Tまたは8113Aを識別することによって診断されることができる。ハプログループ L1は、825A、2758A、2885C、7146G、8468T、8655T、10688A、10810Cまたは13105Gを識別することによって診断されることができる。ハプログループL2は、2416C、2758G、8206A、9221G、11944Cまたは16390Gを識別することによって診断されることができる。ハプログループ L3は、10819G、14212C、8618C、10086C、16362C、10398Aまたは16124Cを識別することによって診断されることができる。ハプログループCは、3552C、4715G、7196A、8584A、9545G、13263G、14318Cまたは16327Tを識別することによって診断されることができる。ハプログループDは、4883T、5178A、8414T、14668Tまたは15487Tを識別することによって診断されることができる。ハプログループ Eは、16227Gを識別することによって診断されることができる。ハプログループGは、4833G、8200Cまたは16017Cを識別することによって診断されることができる。ハプログループ Zは、11078G、16185Tまたは16260Tを識別することによって診断されることができる

。ハプログループAは、663G、16290Tまたは16319Aを識別することによって診断されることができる。私が4529T、10034Cまたは16391Aを識別して診断されることができるハプログループ。ハプログループWは、204C、207A、1243C、5046A、5460A、8994A、11947G、15884Cまたは16292Tを識別することによって診断されることができる。ハプログループXは、1719A、3516G、6221のCまたは14470Cを識別することによって診断されることができる。ハプログループ Fは、12406Aまたは16304Cを識別することによって診断されることができる。ハプログループYは、7933G、8392A、16231Cまたは16266Tを識別することによって診断されることができる。ハプログループ Uは、3197C、4646C、7768G、9055A、11332T、13104G、14070G、15907G、16051G、16129C、16172C、16219G、16249C、16270T、16311T、16318T、16343Gまたは16356Cを識別することによって診断されることができる。ハプログループJは、295T、12612G、13708Aまたは16069Tを識別することによって診断されることができる。ハプログループ Tは、11812G、12633T、14233G、16163C、16186T、1888A、4917G、8697A、10463C、13368A、14905A、15607G、15928Aまたは16294Tを識別することによって診断されることができる。ハプログループVは、72C、4580Aまたは15904Tを識別することによって診断されることができる。ハプログループ Hは、2706Aまたは7028Cを識別することによって診断されることができる。ハプログループBの診断はより複雑である。そして、3つのステップを必要とする。ハプログループ Bは、16189Cを識別することによって診断されることができる;そして、1719A、3516G、6221C、14470Cまたは16278Tの欠如をそばに識別すること;そして、1888A、4216C、4917G、8697A、10463C、11251G、11467G、12308G、12372A、12633T、13104G、13368A、14070G、14905A、15452A、15607G、15928A、16126C、16163C、16186T、16249Cまたは16294Tの欠如を識別することによって。

【０１０６】

【表１０】

Table 10
Nucleotide Alleles Useful for Diagnosing Human Haplogroups

72	C
204	C
207	A
295	T
663	G
825	A
1243	C
1719	A
1888	A
2416	C
2706	A
2758	A
2758	G
2885	C
3197	C
3516	G
3552	C
4216	C
4529	T
4580	A
4586	C
4646	C
4715	G
4833	G
4883	T
4917	G
5046	A
5178	A
5460	A
6221	C
7028	C
7146	G
7196	A
7768	G
7933	G
8113	A
8200	C
8206	A
8392	A
8414	T
8468	T
8584	A
8618	C
8655	T
8697	A
8994	A
9055	A
9221	G
9545	G
9818	T
10034	C
10086	C
10398	A
10463	C
10688	A
10810	C
10819	G
11078	G
11251	G
11332	T
11467	G
11812	G
11944	C
11947	G
12308	G
12372	A
12406	A
12612	G
12633	T
13104	G
13105	G
13263	G
13368	A
13708	A
14070	G
14212	C
14233	G
14318	C
14470	C
14668	T
14905	A
15301	A
15452	A
15487	T
15607	G
15884	C
15904	T
15907	G
15928	A
16017	C
16051	G
16069	T
16124	C
16126	C
16129	C
16163	C
16172	C
16185	T
16186	T
16219	G
16227	G
16231	C
16249	C
16260	T
16266	T
16270	T
16278	T
16290	T
16292	T
16294	T
16304	C
16311	T
16318	T
16319	A
16327	T
16343	G
16356	C
16362	C
16390	G
16391	A

追加的な対立遺伝子は、表11に含まれる。これらの対立遺伝子は、ハプログループを診断するために上記したそれらにとって、等価な方法を設計することに役立つ。表11の対立

遺伝子は、マクロハプログループを診断する効率的な方法を設計することに役立つ。表10および11および図3のデータは、また、副ハプログループを識別することに役立つ。本発明は方法を人間のサンプルにおいて同一化することによって副ハプログループL1a1を診断することに提供する。そして、ヌクレオチド対立遺伝子のうちの1つが4586Cおよび9818Tからなる群から選択される。本発明は方法を人間のサンプルにおいて同一化することによって副ハプログループL1a2を診断することに提供する。そして、ヌクレオチド対立遺伝子のうちの1つが8113Aおよび8251Aからなる群から選択される。本発明は、方法を2352C人間のサンプル、ヌクレオチド対立遺伝子および3666A、7055G、7389C、13789Cおよび14178Cからなる群から選択されるヌクレオチド対立遺伝子のうちの1つにおいて同一化することによって副ハプログループL1b1を診断することに提供する。本発明は、方法を人間のサンプル、3796Cからなる群から選択されるヌクレオチド対立遺伝子のうちの1つ、5951G、5984G、6071C、9072G、10586A、12810Gおよび13485Gにおいて同一化することによって副ハプログループL1b2を診断することに提供する。本発明は、13803G、人間のサンプルにおいて識別することによって副ハプログループ L2aを診断する方法にヌクレオチド対立遺伝子を提供する。本発明は、4158G、人間のサンプルにおいて識別することによって副ハプログループL2bを診断する方法にヌクレオチド対立遺伝子を提供する。本発明は、方法を人間のサンプル、325Tからなる群から選択されるヌクレオチド対立遺伝子のうちの1つ、680Cおよび13958Cにおいて同一化することによって副ハプログループL2cを診断することに提供する。本発明は、方法を人間のサンプル、2325Cからなる群から選択されるヌクレオチド対立遺伝子のうちの1つ、10819Gおよび14212Cにおいて同一化することによって副ハプログループL3aを診断することに提供する。本発明は、8618C、人間のサンプルにおいて識別することによって副ハプログループ L3bを診断する方法にヌクレオチド対立遺伝子を提供する。本発明は、10086C、人間のサンプルにおいて識別することによって副ハプログループL3cを診断する方法にヌクレオチド対立遺伝子を提供する。本発明は、10398A、人間のサンプルにおいて識別することによって副ハプログループ L3dを診断する方法にヌクレオチド対立遺伝子を提供する。本発明は方法を人間のサンプルにおいて同一化することによって副ハプログループUkを診断することに提供する。そして、ヌクレオチド対立遺伝子のうちの1つがグループ成っているof9055Aおよび1631のITから選ばれる。本発明は、16318T、人間のサンプルにおいて識別することによって副ハプログループU7を診断する方法にヌクレオチド対立遺伝子を提供する。本発明は方法を人間のサンプルにおいて同一化することによって副ハプログループ U6を診断することに提供する。そして、ヌクレオチド対立遺伝子のうちの1つが16172Cおよび16219Gからなる群から選択される。本発明は、方法を人間のサンプル、3197Cからなる群から選択されるヌクレオチド対立遺伝子のうちの1つ、7768Gおよび16270Tにおいて同一化することによってsub-ハプログループU5を診断することに提供する。本発明は、方法を、16356C、人間のサンプル、4646Cからなる群から選択されるヌクレオチド対立遺伝子のうちの1つ、11332Tにおいて同一化することによって副ハプログループU4を診断することに提供する。本発明は、16343G、人間のサンプルにおいて識別することによって副ハプログループ U3を診断する方法にヌクレオチド対立遺伝子を提供する。本発明は、方法を人間のサンプル、15907Gからなる群から選択されるヌクレオチド対立遺伝子のうちの1つ、16051Gおよび16129Cにおいて同一化することによって副ハプログループU2を診断することに提供する。本発明は、方法を人間のサンプル、13104Gからなる群から選択されるヌクレオチド対立遺伝子のうちの1つ、14070G、16189Cおよび16249Cにおいて同一化することによって副ハプログループU1を診断することに提供する。本発明は方法を人間のサンプルにおいて同一化することによって副ハプログループ T*を診断することに提供する。そして、ヌクレオチド対立遺伝子のうちの1つが11812Gおよび14233Gからなる群から選択される。本発明は、方法を人間のサンプル、12633Tからなる群から選択されるヌクレオチド対立遺伝子のうちの1つ、16163Cおよび16186Tにおいて同一化することによって副ハプログループT1を診断することに提供する。

【０１０７】

【表１１】

Table 11
Nucleotide Alleles Useful for Diagnosing Human Haplogroups and Macro-Haplogroups

72	C
73	A
204	C
207	A
295	T
325	T
489	C
663	G
680	C
769	A
825	A
1018	A
1048	T
1243	C
1719	A
1888	A
2352	C
2416	C
2706	A
2758	A
2758	G
2885	C
3197	C
3516	A
3516	G
3552	C
3594	T
3666	A
3796	C
4104	G
4158	G
4216	C
4312	T
4529	T
4580	A
4586	C

4646	C
4715	G
4833	G
4883	T
4917	G
5046	A
5178	A
5442	C
5460	A
5951	G
5984	G
6071	C
6185	C
6221	C
7028	C
7055	G
7146	G
7196	A
7256	T
7389	C
7521	A
7768	G
7933	G
8113	A
8200	C
8206	A
8251	A
8392	A
8414	T
8468	T
8584	A
8618	C
8655	T
8697	A
8701	A
8994	A
9055	A
9072	G
9221	G
9347	G

9402	C
9540	T
9545	G
9818	T
10034	C
10086	C
10398	A
10400	T
10463	C
10586	A
10589	A
10664	T
10688	A
10810	C
10819	G
10873	T
10915	C
11078	G
11251	G
11332	T
11467	G
11719	G
11812	G
11944	C
11947	G
12007	A
12308	G
12372	A
12406	A
12612	G
12633	T
12705	C
12810	G
13104	G
13105	G
13263	G
13276	G
13368	A
13485	G
13506	T

13650	T
13708	A
13789	C
13803	G
13958	C
14070	G
14178	C
14212	C
14233	G
14318	C
14470	C
14668	T
14766	C
14783	C
14905	A
15043	A
15301	A
15452	A
15487	T
15607	G
15884	C
15904	T
15907	G
15928	A
16017	C
16051	G
16069	T
16124	C
16126	C
16129	A
16129	C
16163	C
16172	C
16185	T
16186	T
16189	C
16219	G
16223	C
16224	C
16227	G
16231	C
16249	C
16260	T
16266	T
16270	T
16278	T
16290	T
16292	T
16294	T
16298	C
16304	C
16311	T
16318	T
16319	A
16327	T
16343	G
16356	C
16362	C
16390	G
16391	A

ハプログループを診断する等価な方法は、825A、2758A、2885C、7146G、8468T、8655T、10688A、10810Cまたは13105Gのうちの1つの存在を識別することによってハプログループL0を診断している;そして、3666A、7055G、7389C、13789Cまたは14178Cのうちの1つの欠如を識別すること。他の等価な方法は、図3のデータに由来することができて、本発明

の範囲内である。

【０１０８】

（実施例6）

LebersHereditary Optic Neuropathy（LHON）は、ミトコンドリアＤＮＡ（mtDNA）突然変異によって生じる盲目の形式である。4つの突然変異、3460A、11778A、14484Cおよび14459Aは、世界的に90%以上のLHONを占めて、designated"primary"mutationsである。主たる突然変異は強くキャリアにLHONの素因を作る、対照で見つからなくてあって、Complexの全てある私遺伝子、そして、各々によって共起しない。11778Aおよび14484C突然変異がヨーロッパのmtDNAハプログループ J（9%のヨーロッパの由来のmtDNAsで見つかって）に関連して予想されるより、しばしば発生したことを証明された。そして、mtDNAの中の相乗作用の相互作用に突然変異が疾患表現の確率を増やしたことを示唆した。表3にリストされるヌクレオチド対立遺伝子の除去を含んで、主たるLHON突然変異のない2つのロシアのLHON系統のシーケンス分析は、2つの新規な複合体を示した私突然変異、3635A、そして、4640C。静脈血液サンプルは、家族から得られた。ゲノムDNAは、Chelex100（鯨座、Emberyville、CA、USA）を使用している軟膜血液分数から分離された。mtDNAは拡大されたPCR 2-3kb断片（100桁Centriconに純化される）の、そして、サイクル-配列された使用しているBigDyeTerminators（ABI/パーキン・エルマー鯨座）およびABI Prism 377はDNAシーケンサを自動化した。突然変異は、組合せを誤られた下塗り後の突然変異に特有の規制酵素消化を使用して確認されたPCR白血球mtDNAの拡大（茶色のM.D.その他、（1995の）Human Mutat。6:311-325).

（実施例7）

10663C、新規な主たるLHONmtDNA突然変異Complexを影響を及ぼす私、遺伝子は3つのコーカサスのLHON系統、いずれが帰属したか全部、ハプログループ J. These 3つの系統においてhomoplasmicだった。そして、唯一のハプログループは周知の、主たるLHON突然変異を港にしなかったJ関連のLHON系統（17の中の）であった。完全なmtDNAシーケンスを使用しているハプログループJの広範囲の系統発生の分析は、10663C異型にこの背景上の起こられた3つの独立時間があることを証明した。この突然変異は、200以上のnon-ハプログループJヨーロッパの対照、74のハプログループ J患者および制御部mtDNAsまたは主たる突然変異のない36人の推定のLHON患者に存在しなかった。部分的なComplex、私は離脱する10663C-含んでいるリンパ芽球の見つけた、そして、サイブリッド・ミトコンドリア。このように、10663C突然変異に、起こられた3つの独立時間（ハプログループJ上の、そして、周知のLHON突然変異のないLHON患者だけの各時間）がある。それが表現のためのハプログループ Jによって提供される遺伝子の背景を必要とするように見えるという点で、これは全ての病原性mtDNA突然変異の中でユニークな10663C突然変異をする。これらの結果は、更なる証拠をハプログループJの、そして、付加的な方法で相互に作用している方法of"mild"mtDNA突然変異のための役割に疾患表現を引き起こすように素因を作ることに提供する。軽度のND6np 14484およびND3 np 10663レーバーのHereditary Optic Neuropathy（LHON）を有するヨーロッパ人ミスセンス突然変異それらもmtDNAハプログループ J.を所有する場合、より多くが盲目の傾向がある

（実施例8）

大陸間のmtDNAシーケンス放射線、中立不偏のモデルがタジマのDを使用しているmtDNAシーケンス異型およびFuの配布において、見つけるため検査された標準からの偏差およびLiD*試験における人口統計学的要因の重要性を評価する（Y. X.Fu、W. H. Li、（1993の）Genetics 133：693-709。そして、F.タジマ、（1989の）Genetics123,585-95）。独自に起こっている全ての突然変異および選択的にニュートラルについては、集団遺伝学の標準の中立不偏のモデルは、恒常的なサイズのランダムな嵌合人口を装う。ペアをなすmtDNAシーケンス差の大陸度数分布は、A.R.ロジャーズ、HarpendingしているH.、（1992の）Mol.Biolの方法を使用している迅速な人口展開に対する検査に算出された。Evol.9:552-569。

【０１０９】

アフリカのmtDNAシーケンス（n= 32）のために、結果は有意に標準の中立不偏のモデルからそれなかった、そして、違いが計数するペアをなすシーケンスの度数分布は広くて耳障りだった。これらの結果の両方とも、アフリカの人口が長い間比較的安定していたモデルと整合している。対照的に、non-アフリカのmtDNAs（n= 72）は、中性からの非常に有意な偏差を示した（タジマは、D =-2である。43（P < 0.01）;FuおよびLi D* =-5。09、P <0.02）、ベル形の度数分布ofpairwiseと同程度よくシーケンス差。このように、これらの結果は、度数分布をゆがめていた人口展開と整合している（L.Excoffier（J. Mol.）Evol.30：125-39（1990）およびD. A. Merriwetherほか（1991の）J. Mol.Evol33：543-555）。

【０１１０】

よりよくこれらの人口統計学的影響の局所の配布を定めるために、ユーラシア・サンプルは、ヨーロッパでアジアのプラスのアメリカインディアンに分けられた。全てのヨーロッパのmtDNAsの分析も、標準の中立不偏のモデルからの重要な逸脱を明らかにした（タジマは、D =-2である。19、P < 0。01;FuandLiD*=-3。31、P< 0。02).ヨーロッパのmtDNAsのためのペアをなすシーケンス差の配布は2つの鋭いピークを明かした。そして、2つの主要拡張位相をほのめかした。ハプログループHおよびV mtDNAsがサンプルから削除されるときに、これらのピークで最も最近のものは失われた。それゆえに、ハプログループ H（それは40%の現代のヨーロッパのmtDNAs（A.Torroniほか（Human Genetics 62,1137-1152（1998）アメリカのジャーナル））を表して、19,000のYBPのMRCAを有する）は、比較的最近ヨーロッパにおいて優れているようになった。

【０１１１】

集計されたアジアでアメリカインディアンのmtDNAs（n = 41）の分析も、標準の中立不偏のモデルからの重要な逸脱を明らかにした（タジマのD =-2.28、P <0.01、FuおよびLi D* =-4。31 ;迅速な人口展開と整合したペアをなす違いの幅広い、ベル形の配布を同じく明らかにしているP < 0.02）。

【０１１２】

アジアのアメリカインディアンハプログループ A、B、C、DおよびX mtDNAs（n = 26）が別に分析されるときに、それらもFuおよびLi D*試験のための中性からの有意な偏差を示して、（D*=-2。65（P < 0.05））タジマのD試験（D =-1.60（ns））のためでない。ペアをなすシーケンス差のそれらの配布はまた、強く単峰形だった。そして、人々がシベリアおよびBeringiaで、そして、米州に移動したので、人口が拡大したことを示した。

【０１１３】

（実施例9：人間のmtDNA遺伝子の可変の置換突然変異率）

選択が大陸間のmtDNAシーケンス変化の突然の変動を引き起こすことにおける重要な要因であったかどうか決定するために、非同義の、同義のベース置換の数は、各々の主な大陸スペースの植民地化に貢献したそれらのハプログループの全ての13のmtDNAタンパク質遺伝子のために分析された：アフリカ人、ヨーロッパ人およびアメリカインディアン。例えば、アジアのアメリカインディアンハプログループAからのthe'Native Americans"the mtDNAsのために、B、C、DおよびXは、結合された。ランダムな突然変異が創設者人口にたまるので、ハプログループからのアジアのアメリカインディアンmtDNAsは結合された、そして、新しい環境において有利であると判明するそれらのmtDNAsは豊かにされる。それゆえに、ハプログループの突然変異を創立することは、血統の大陸成功において重要である。我々は、それから比率を比較することによって各々の大陸の植民地化の間、非各々のmtDNA遺伝子のための同義のヌクレオチド置換対可能な選択的な効果の陽性反応が出た。同義の突然変異比率にsynonynousなnon-の増加は、選択が機能的に変えられたタンパク質の普及を支持したことを示唆する。

【０１１４】

一度だけ各々の変化を計数して、同義の突然変異に対するnonsynonymousの比率の比較は、いくつかの遺伝子（表12）のための大陸間の大きな変化を明らかにした。非同義の突然変異の蓄積の著しい増加は、アフリカ人、Cytbおよびアメリカインディアンのヨーロッ

パ人およびATP6のCOHIのND3について参照された。各々の遺伝子のための非同義で同義の突然変異の数は、また、Two尾のあるFisherExact Testを使用しているP値を計算することによって異なる大陸の間で比較された。そして、全てのmtDNA遺伝子（表12）の合計のためのアフリカ人およびヨーロッパ人との間に、ATP6のためのアフリカ人およびアメリカインディアンとの間に、これはアフリカ人および両方のヨーロッパ人間の有意な差およびCOIIIのためのアメリカインディアンを現した。それゆえに、この分析は、選択が大陸mtDNAタンパク質バリエーションを成形することの役をつとめたという仮説を支える。

【０１１５】

【表１２】

Table 12*

Gene	Number of Polymorphic Sites									Two-Tail FET P-value		
	African			European			Native American			Afr vs Eur	Afr vs Am	Eur vs Am
	N-syn	Syn	Ratio	Syn	N-syn	Ratio	Syn	N-syn	Ratio			
ND1	10	17	0.59	5	5	***1.00***	4	4	***1.00***	0.71	0.69	1.00
ND2	9	22	0.41	4	9	0.44	3	7	0.43	1.00	1.00	1.00
ND3	6	2	***3.00***	1	3	0.33	1	4	0.25	0.22	0.10	1.00
ND4L	0	7	0.00	0	1	0.00	1	4	0.25	1.00	0.42	1.00
ND4	4	35	0.11	2	13	0.15	3	12	0.25	1.00	0.38	1.00
ND5	15	31	0.48	8	20	0.40	2	14	0.14	0.80	0.19	0.28
ND6	2	14	0.14	1	6	0.17	3	5	0.60	1.00	0.29	0.57
Cytb	11	19	0.58	14	9	***1.56***	5	12	0.42	0.10	0.75	0.60
COI	7	30	0.23	0	9	0.00	0	13	0.00	0.32	0.17	1.00
COII	3	19	0.16	0	4	0.00	2	6	0.33	1.00	0.59	0.52
COIII	1	13	0.08	6	5	***1.20***	7	10	0.70	***0.02***	***0.05***	0.70
ATP6	3	15	0.20	5	6	0.83	7	5	***1.40***	0.20	***0.05***	0.68
ATP8	2	3	0.67	2	0		1	3	0.33	0.43	1.00	0.40
Total	73	227	0.32	48	90	0.53	39	99	0.39	***0.03***	0.41	0.30

* mtDNA遺伝子の同義の突然変異数対置換。Rplmt=置換突然変異（比率= rplmt/無声映画）。FET = Fisher Exact Test。Afr =アフリカ（Eur =ヨーロッパ）Am =アメリカインディアン。太いイタリックの多形サイトの比率は、観察されるいくつかのより高い値を強調する。Two-TailedされたFETの下の太いイタリックのそれらは、0.05のレベルで重要である比較を示す。

【０１１６】

（実施例10）

各々の突然変異は、上記の分析から、ハプログループの範囲内のその頻度にかかわりなく、それ重要なだけである‖節点の突然変異およびover-の重要性が重要性を強調するunder-emphasizes末期の個人的な多形性。この方法の変形例として、我々が算出したこと修正非同義の（kn）、そして、同義の（Ks）突然変異頻度、そうすると、算出することによってその遺伝子に作用している相対的な選択的な制約を決定するkc値｛k=-ln（Ifa/KS）｝。高さkc、値は高いタンパク質シーケンスの保護および低いアミノ酸バリエーションを表す安い価格が低いタンパク質の保護および高いアミノ酸バリエーションを表す、（N.Neckelmannその他、（1987の）会報Natl.Acad。科学USA 84：7580-7584)。

【０１１７】

kc、各々の人間のmtDNA遺伝子のための値は、人間のmtDNAシーケンス（図4）の完全な大域的コレクション全体に比較された。ATP6遺伝子は、以前、それが非常に異種間比較（N.Neckelmannほか（（1987の）会報Natl. Acad. Sci.USA 84：7580-7584））において、比較的節約されることを示したにもかかわらず、人間のmtDNAの最も少なく節約された遺伝子であった。

【０１１８】
（実施例11）
ATP6のより高い異種間保護は、匹敵することによって確認されたkcチンパンジ（パン穴居人）およびコビトチンパンジ（パンpaniscus）対人間の値;8つの霊長類種（ヒヒ、ボルネオおよびスマトラ島オランウータン、テナガザル、ゴリラ、低地ゴリラ、コビトチンパンジおよびチンパンジ）対人間;そして、13の多様な哺乳類の種（ウシ、マウス、ネコ、イヌ、ブタ、ネズミ、サイ、ウマ、テナガザル、ゴリラ、オランウータン、コビトチンパンジ、チンパンジ）（図3）対人間。このように、ATP6が種の間で非常に節約されると共に、それは人間の範囲内であまり十分に節約されない。これらの結果は、他の遺伝子のために観察される異種間保護対、減少する内部特性と整合している（C.A. Wiseほか（（1998の）Genetics 148）：409-21、そして、ミトコンドリア・タンパク質バリエーションがシトクロムｃオキシダーゼ遺伝子に示すように人間および他の霊長類において加速されるという仮説については）（L.I.Grossmanその他、（2001の）Mol. Phylogerlet.Evol.18:26-36)。
【０１１９】
（実施例12）
個々のmtDNAタンパク質遺伝子が異なる人間の大陸人口のそれらの選択的な制約において異なるという可能性を更に調査するkc、大陸ハプログループの各々の一組からの全ての13のmtDNAタンパク質遺伝子のための値は、算出された：アフリカ人、ヨーロッパ人およびアメリカインディアン。それがペアをなす比較によって大陸の一組のmtDNAsを切り離した累積的な選択的な圧力kc、値は各々のmtDNA（表13）の遺伝子のために算出された。mtDNAタンパク質の比較kc、アフリカ人対ヨーロッパ人の価値は、3つの遺伝子（ND1、cytbおよびGOUT）がヨーロッパ人のかなりより低いシーケンスの保護を有することを明らかにした。比較のkc、アフリカのmtDNA遺伝子対アメリカインディアンの値は、アメリカインディアンのかなりより低いシーケンスの保護を有した6つの遺伝子（ND4、ND6、COII、COIII、ATP6およびATP8）を現した。最後に、ヨーロッパ人またはアメリカインディアン対アフリカ人のkC値の比較で、4つのmtDNA遺伝子（ND3、ND5、cytbおよびCOI）がアフリカ人のかなりより低いシーケンスの保護を有することが分かった。最も大きな中で違いkc、値はCOIIIおよびATP6のアフリカ人およびアメリカインディアンの間の比較のために、そして、アフリカ人およびヨーロッパ人（表13）間のCOIIIのために参照された。
【０１２０】
【表１３】

Table 13*

GENES	African sequences (n=32)	European sequences (n=31)	T-test P value	Native American sequences {A,B,C,D,X} (n=26)	T-test P value
ND1	2.08 ± 1.18	0.27 ± 1.90	P<0.0001	2.07 ± 1.92	NS
ND2	1.72 ± 1.07	1.57 ± 1.85	NS	1.81 ± 1.11	NS
ND3	0.51 ± 1.87	0.91 ± 2.32	NS	1.70 ± 1.32	P<0.01
ND4L	*	*	NS	2.41 ± 3.83	*
ND4	3.49 ± 1.34	3.39 ± 2.23	NS	2.20 ± 1.19	P<0.001
ND5	1.78 ± 0.71	2.20 ± 1.20	NS	3.63 ± 3.56	P<0.01
ND6	2.51 ± 1.19	3.13 ± 3.99	NS	1.15 ± 1.52	P<0.001
Cytb	1.89 ± 0.96	0.34 ± 1.51	P<0.0001	2.46 ± 1.15	P<0.05
COI	2.37 ± 0.95	3.85 ± 3.93	P<0.05	*	*
COII	2.73 ±1.32	*	*	1.74 ± 2.12	P<0.05
COIII	4.65 ± 3.94	0.94 ± 2.08	P<0.0001	2.11 ± 1.26	P<0.01
ATP6	2.31 ± 1.28	1.48 ± 2.28	NS	-0.14 ± 1.34	P<0.0001
ATP8	2.62 ± 1.89	*	*	1.25 ± 1.94	P<0.01

* 選択的な制約の係数の評価（kc）遺伝子および領域のそばの層になる。kc、値および標準偏差は、アフリカ人、ヨーロッパでアジア系アメリカ人のハプログループA,B、C、DおよびX mtDNAタンパク質-符号化遺伝子のために算出した。*はkC値が算出されることができなかったことを示す。これは、次のことの故である。いずれのKもまたはKaは0であった。ハプログループXはNative-アメリカのシーケンスだけによって表される。そして、ヨーロッパのXシーケンスが除外される。

【０１２１】

一緒に必要とされて、これらのデータは、人間が異なる大陸にコロニーを作ったので、異なる選択的な力が個々のmtDNA遺伝子に作用したことを示す。さらに、mtDNAタンパク質シーケンスの観察された違いは、彼らが熱帯で亜熱帯アフリカから、そして、温暖なユーラシアおよび北極シベリアおよびBeringiaに移住したので、人間が経験があるようにする気象的な移行と相関する。大陸間の最も高いアミノ酸シーケンス・バリエーションを示したmtDNA遺伝子は、COinおよびATP6であった。

【０１２２】

（実施例13）

実施例9-12において識別される進化的にな重要な遺伝子にあっている表3のヌクレオチド対立遺伝子は、進化の重要性のために分析された。進化的にな重要な対立遺伝子は、進化的にな重要な遺伝子にあって、アミノ酸変化を引き起こす。ND1の進化的にな重要なヌクレオチド対立遺伝子のリスト、ND2、ND3、ND4、ND5、ND6、Cytb、COI、COII、COin、ATP6およびATP8は、表14に現れる。表14のケンブリッジ・ヌクレオチド対立遺伝子は、進化的にに重要である。ケンブリッジ対立遺伝子を含んで、これらのアミノ酸対立遺伝子は、進化的にに重要である。アミノ酸対立遺伝子の位置は、表3にリストされるヌクレオチド対立遺伝子の位置によって識別される。表14にリストされない他の進化的にな重要なヌクレオチド対立遺伝子は、同じコドンの範囲内である隣接したヌクレオチド場所および表14にリストされる同じアミノ酸のためのコードで対立遺伝子を含む。

【０１２３】

【表１４】

Table 14

Evolutionarily Significant Human Mitochondrial Nucleotide and Amino Acid Alleles

Gene	Genome Location	Cambridge Nucleotide	Non-Cambridge Nucl. Allele	Cambridge Amino Acid	Non-Cambridge AA Allele
ND1	3308	T	C	M	T
ND1	3316	G	A	A	T
ND1	3394	T	C	Y	H
ND1	3505	A	G	T	A
ND1	3547	A	G	I	V
ND1	3565	A	G	T	A
ND1	3644	T	C	V	A
ND1	3796	A	T	T	A
ND1	3796	A	G	T	S
ND1	3796	A	C	T	P
ND1	3808	A	G	T	A
ND1	3866	T	C	I	T
ND1	4025	C	T	T	M
ND1	4040	C	T	T	M

ND1	4048	G	A	D	N
ND1	4123	A	G	I	V
ND1	4216	T	C	Y	H
ND1	4225	A	G	M	V
ND1	4232	T	C	I	T
ND2	4491	G	A	V	I
ND2	4506	A	G	I	V
ND2	4512	G	A	A	T
ND2	4596	G	A	V	I
ND2	4695	T	C	V	I
ND2	4767	A	G	M	V
ND2	4824	A	G	T	A
ND2	4833	A	G	T	A
ND2	4917	A	G	N	D
ND2	4960	C	T	A	G
ND2	5043	G	T	A	S
ND2	5046	G	A	V	I
ND2	5178	C	A	L	M
ND2	5262	G	A	A	T
ND2	5263	C	T	A	V
ND2	5331	C	A	L	I
ND2	5442	T	C	F	L
ND2	5460	G	A	A	T
COI	6150	G	A	V	I
COI	6253	T	C	M	T
COI	6324	G	A	A	T
COI	6366	G	A	V	I
COI	6607	T	C	F	S
COI	7146	A	G	T	A
COI	7257	A	G	I	V
COI	7347	G	A	V	I
COI	7389	T	C	Y	H
COI	7444	G	A	TER	K
COII	7664	G	A	A	T
COII	7673	A	G	I	V
COII	7697	G	A	V	I
COII	8027	G	A	A	T
COII	8142	C	T	A	V
ATP8	8387	G	A	V	M
ATP8	8414	C	T	L	F
ATP8	8448	T	C	M	T
ATP8	8460	A	G	N	S
ATP8	8472	C	T	P	L
ATP8	8553	C	T	S	L
ATP6	8545	G	A	A	T
ATP6	8563	A	G	T	A
ATP6	8566	A	G	I	V
ATP6	8584	G	A	A	T
ATP6	8618	T	C	I	T

ATP6	8701	A	G	T	A
ATP6	8705	T	C	M	T
ATP6	8764	G	A	A	T
ATP6	8794	C	T	H	Y
ATP6	8836	A	G	M	V
ATP6	8860	A	G	T	A
ATP6	8875	T	C	F	L
ATP6	8962	A	G	T	A
ATP6	9053	G	A	S	N
ATP6	9055	G	A	A	T
ATP6	9077	T	C	I	T
ATP6	9103	T	C	F	L
ATP6	9136	A	G	I	V
ATP6	9151	A	G	I	V
COIII	9237	G	A	V	M
COIII	9325	T	C	M	T
COIII	9355	A	G	N	S
COIII	9456	A	G	I	V
COIII	9402	A	C	T	P
COIII	9477	G	A	V	I
COIII	9559	C	G	P	R
COIII	9591	G	A	V	I
COIII	9667	A	G	N	S
COIII	9682	T	C	M	T
COIII	9822	C	A	L	I
COIII	9957	T	C	F	L
COIII	9966	G	A	V	I
ND3	10086	A	G	N	D
ND3	10086	A	C	N	H
ND3	10152	G	C	E	Q
ND3	10182	G	C	D	H
ND3	10197	G	A	A	T
ND3	10321	T	C	V	A
ND3	10398	A	G	T	A
ND4L	10609	T	C	C	R
ND4	10816	A	T	K	N
ND4	10920	C	T	P	L
ND4	11016	G	A	S	N
ND4	11078	A	G	I	V
ND4	11150	G	A	A	T
ND4	11172	A	G	N	S
ND4	11177	C	T	P	S
ND4	11654	A	G	T	A
ND4	11909	A	G	T	A
ND4	11963	G	A	V	I
ND4	11969	G	A	A	T
ND4	12083	T	G	S	A
ND4	12134	T	C	S	P
ND5	12346	C	T	H	Y

ND5	12358	A	G	T	A
ND5	12361	A	G	T	A
ND5	12373	A	G	T	A
ND5	12397	A	G	T	A
ND5	12406	T	A	V	I
ND5	12635	T	C	I	T
ND5	12850	A	G	I	V
ND5	12940	G	A	A	T
ND5	12967	A	C	T	P
ND5	13104	A	G	I	V
ND5	13105	A	G	I	V
ND5	13135	G	A	A	T
ND5	13145	G	A	S	N
ND5	13276	A	G	M	V
ND5	13477	G	A	A	T
ND5	13651	A	G	T	A
ND5	13660	A	G	N	D
ND5	13708	G	A	A	T
ND5	13759	G	A	A	T
ND5	13780	A	G	I	V
ND5	13789	T	C	Y	H
ND5	13819	T	C	F	L
ND5	13880	C	A	S	Y
ND5	13886	T	C	L	P
ND5	13924	C	T	P	S
ND5	13927	A	T	S	C
ND5	13928	G	C	S	T
ND5	13958	G	C	G	A
ND5	13966	A	G	T	A
ND5	14000	T	A	L	Q
ND5	14059	A	G	I	V
ND5	14128	A	G	T	A
ND6	14178	T	C	I	V
ND6	14272	C	G	L	F
ND6	14318	T	C	N	S
ND6	14319	T	C	N	D
ND6	14384	G	C	A	G
ND6	14459	G	A	A	V
ND6	14484	T	C	M	V
ND6	14502	T	C	I	V
ND6	14571	T	A	S	C
CytB	14766	C	T	T	I
CytB	14769	A	G	N	S
CytB	14793	A	G	H	R
CytB	14798	T	C	F	L
CytB	14861	G	A	A	T
CytB	14862	C	T	A	V
CytB	14979	T	C	I	T
CytB	15110	G	A	A	T

CytB	15113	A	G	T	A
CytB	15204	T	C	I	T
CytB	15218	A	C	T	P
CytB	15218	A	G	T	A
CytB	15238	C	G	I	M
CytB	15257	G	A	D	N
CytB	15261	G	A	S	N
CytB	15317	G	A	A	T
CytB	15318	C	T	A	V
CytB	15323	G	A	A	T
CytB	15326	A	G	T	A
CytB	15431	G	A	A	T
CytB	15452	C	A	L	I
CytB	15497	G	A	G	S
CytB	15519	T	C	L	P
CytB	15663	T	C	I	T
CytB	15731	G	A	A	T
CytB	15746	A	G	I	V
CytB	15803	G	A	V	M
CytB	15806	G	A	A	T
CytB	15812	G	A	V	M
CytB	15824	A	G	T	A
CytB	15849	C	T	T	I
CytB	15884	G	C	A	P

本発明の方法を使用している生理的状況に対する素因を伴う表14の対立遺伝子のサブセットは、表15にリストされる。

【０１２４】

【表１５】

Table 15

Amino Acid Alleles Associated with Physiological Conditions in this Invention

Gene	Genome Location	Nucleotide Alleles Useful for Diagnosing Haplogroups	Amino Acid Alleles Useful for Diagnosing Haplogroups	Haplogroups Diagnosable by Alleles
ND1	3796	C	P	(L1b2)
ND2	4833	G	A	G
ND2	4917	G	D	T
ND2	5046	A	I	W
ND2	5178	A	M	D
ND2	5442	C	L	L0
ND2	5460	A	T	W
COI	7146	G	A	L0, L1
COI	7389	C	H	L1
ATP8	8414	T	F	D
ATP6	8584	A	T	C
ATP6	8618	C	T	(L3b)
ATP6	8701	A	T	A, I, W, X, B, F, Y, U, J, T, V, H

ATP6	9055	A	T	(Uk)
COIII	9402	C	P	L0
ND3	10086	C	H	(L3c)
ND3	10398	A	T	(L3d)
ND4	11078	G	V	Z
ND5	12406	A	I	F
ND5	13104	G	V	(U1)
ND5	13105	G	V	L0, L1
ND5	13276	G	V	L0
ND5	13708	A	T	J
ND5	13789	C	H	L1
ND5	13958	C	A	(L2c)
ND6	14178	C	V	L1
ND6	14318	C	S	C
CytB	14766	C	T	V, H
CytB	15452	A	I	J, T
CytB	15884	C	P	W

（実施例14：ATP6の大陸に特有のアミノ酸置換）
更に人間の大陸に特有のATP6アミノ酸置換の生物学的重要性を調査するために、39の動物の種mtDNAs（12匹の霊長類、22匹の他の哺乳類、4匹の非哺乳類の脊椎動物およびショウジョウバエ）を使用している各々の可変の人間の位置のためのアミノ酸の保護は、分析された。これは、特定のmtDNAハプログループを伴うATP6置換の節約されて進化的にでそれゆえに、重要な多数が潜在的に機能的に変わることを明らかにした、アミノ酸。
【０１２５】
ATP6のコドン59（T59A（ヌクレオチド場所8701-8703））のアラニン置換に対するトレオニンは、Worldの残りからマクロハプログループNのmtDNAsを切り離す。位置59の極地のトレオニンは、全ての大きな類人猿および旧世界の若干の猿において節約される。
【０１２６】
マクロハプログループ Mのハプログループの中で、関連したシベリアのアメリカインディアンハプログループ CおよびZは、A20T（ヌクレオチド位置づけ8584-8586）異型によって詳細に描写される。この位置で見つかる無極性アミノ酸は、全ての動物の種で起こるMacaca属（Papio）Balaenopteraおよびショウジョウバエ。
【０１２７】
マクロハプログループ Nのハプログループの中で、非R血統Nlbは、2つの特徴的なアミノ酸置換を収容している：位置104のメチオニンが全ての哺乳類およびトレオニンにおいて節約されるM104V（ヌクレオチド位置づけ8836-8838）およびT146A.（ヌクレオチド位置づけ8962-8964）は、146を配置する全ての動物のmtDNAs.の全体にわたって節約するさらに、T146A置換は、ATPシンターゼのカップリング効率を変えて、NARPおよびリー症候群（i.Trounce、S.ニール、D. C.ウォレス、アメリカ合衆国91,8334-8338（1994）の米国科学アカデミーのProceedings）を引き起こす病原性突然変異L156Rと同じ膜内外a-helixの範囲内である。
【０１２８】
また、マクロハプログループ Nの、ハプログループ A mtDNAsは、H90Y（ヌクレオチド位置づけ8794-8796）アミノ酸置換を収容している。この位置のヒスチジンは全部で節約される胎盤であるPongo、CebusおよびLoxodontaを除いて哺乳類、そして、非常に保全された領域の中で起こる。さらに、異種のグループのtRNALYs-COII9bp削除をもたらしているmtDNAs一つの、そして、任意にハプログループ Bに割り当てられて、1つのmtDNAは、F193L（ヌクレオチド位置づけ9103-9105）置換を収容していた。この位置は、Pongo、Papio、CebusおよびErinaceusを除いて全ての哺乳類において保全される。
【０１２９】

シーケンスが異なる種のこの比較において使用した各々のmtDNAが１人か２人の個人だけに由来するので、まれな偏位ケースが人間のそれらに対応するそれらの種の環境的に適応突然変異の蓄積によることはあり得る。このように、上記のATP6アミノ酸多形性は、進化的にな適応突然変異のために予想される特徴を有する。

【０１３０】

【表１６】

Table 16

Nucleotide Locus	Nucleotide Alleles	WIPO code
64	CT	y
72	TC	y
73	AG	r
89	TC	y
93	AG	r
95	AC	m
114	CT	y
143	GA	r
146	TC	y
150	CT	y
151	CT	y
152	TC	y
153	AG	r

171	GA	r
180	TC	y
182	CT	y
183	AG	r
185	GAT	d
185	GAT	d
186	CA	m
189	ACG	v
189	ACG	v
194	CT	y
195	TAC	h
195	TAC	h
198	CT	y
199	TC	y
200	AG	r
204	TC	y
207	GA	r
208	TC	y
210	AG	r
212	TC	y
215	AG	r
217	TC	y
225	GA	r
227	AG	r
228	GA	r
235	AG	r
236	TC	y
247	GA	r
250	TC	y
252	TC	y
263	AG	r
291	AG	r
295	CT	y
297	AG	r
316	GA	r
317	CAG	v
317	CAG	v
320	CT	y
325	CT	y
340	CT	y
357	AG	r
373	AG	r
400	TG	k
408	TA	w
418	CT	y
456	CT	y
462	CT	y
465	CT	y
467	CT	y

471	TC	y
480	TC	y
482	TC	y
489	TC	y
493	AG	r
499	GA	r
508	AG	r
593	TC	y
597	CT	y
663	AG	r
678	TC	y
680	TC	y
709	GA	r
710	TC	y
721	TC	y
750	AG	r
769	GA	r
825	TA	w
827	AG	r
850	TC	y
921	TC	y
930	GA	r
961	TCG	b
961	TCG	b
1018	GA	r
1041	AG	r
1048	CT	y
1119	TC	y
1189	TC	y
1243	TC	y
1290	CT	y
1382	AC	m
1406	TC	y
1415	GA	r
1420	TC	y
1438	AG	r
1442	GA	r
1503	GA	r
1598	GA	r
1700	TC	y
1703	CT	y
1706	CT	y
1709	GA	r
1715	CT	y
1719	GA	r
1736	AG	r
1738	TC	y
1780	TC	y
1811	AG	r

1888	GA	r
1927	GA	r
2000	CT	y
2060	AG	r
2092	CT	y
2245	ACG	v
2245	ACG	v
2263	CA	m
2308	AG	r
2332	CT	y
2352	TC	y
2358	AG	r
2380	CT	y
2416	TC	y
2483	TC	y
2581	AG	r
2639	CT	y
2650	CT	y
2706	AG	r
2755	AG	r
2758	GA	r
2768	AG	r
2789	CT	y
2792	AG	r
2834	CT	y
2836	CA	m
2857	TC	y
2863	TC	y
2885	TC	y
3010	GA	r
3083	TC	y
3197	TC	y
3200	TA	w
3202	TC	y
3204	CT	y
3206	CT	y
3221	AG	r
3290	TC	y
3308	TC	y
3316	GA	r
3372	TC	y
3394	TC	y
3438	GA	r
3450	CT	y
3480	AG	r
3505	AG	r
3513	CT	y
3516	CGA	v
3516	CGA	v

3547	AG	r
3549	CT	y
3552	TCA	h
3552	TCA	h
3565	AG	r
3594	CT	y
3644	TC	y
3666	GA	r
3693	GA	r
3699	CT	y
3720	AG	r
3756	AG	r
3796	AGTC	n
3796	ACGT	n
3808	AG	r
3816	AG	r
3834	GA	r
3843	AG	r
3847	TC	y
3866	TC	y
3918	GA	r
3921	CA	m
3927	AG	r
3970	CT	y
3981	AG	r
4025	CT	y
4040	CT	y
4044	AG	r
4048	GA	r
4086	CT	y
4104	AG	r
4117	TC	y
4122	AG	r
4123	AG	r
4158	AG	r
4203	AG	r
4216	TC	y
4221	CT	y
4225	AG	r
4232	TC	y
4248	TC	y
4312	CT	y
4336	TC	y
4370	TC	y
4388	AG	r
4454	TA	w
4491	GA	r
4506	AG	r
4508	CT	y

4512	GA	r
4529	ATC	h
4529	ATC	h
4541	GA	r
4580	GA	r
4586	TC	y
4596	GA	r
4646	TC	y
4655	GA	r
4688	TC	y
4695	TC	y
4715	AG	r
4742	TC	y
4767	AG	r
4769	AG	r
4820	GA	r
4824	AG	r
4833	AG	r
4841	GA	r
4883	CT	y
4907	TC	y
4917	AG	r
4960	CT	y
4977	TC	y
4994	AG	r
5004	TC	y
5027	CT	y
5036	AG	r
5043	GT	k
5046	GA	r
5063	TC	y
5096	TC	y
5108	TC	y
5147	GA	r
5153	AG	r
5178	CA	m
5231	GA	r
5237	GA	r
5255	CT	y
5262	GA	r
5263	CT	y
5285	AG	r
5300	CT	y
5330	CA	m
5331	CA	m
5390	AG	r
5393	TC	y
5417	GA	r
5426	TC	y

5442	TC	y
5460	GA	r
5465	TC	y
5471	GA	r
5492	TC	y
5495	TC	y
5580	TC	y
5581	AG	r
5601	CT	y
5603	CT	y
5606	CT	y
5633	CT	y
5655	TC	y
5711	AG	r
5773	GA	r
5811	AG	r
5814	TC	y
5821	GA	r
5826	TC	y
5843	AG	r
5951	AG	r
5984	AG	r
5987	CT	y
6026	GA	r
6029	CT	y
6045	CT	y
6071	TC	y
6077	CT	y
6104	CT	y
6150	GA	r
6152	TC	y
6164	CT	y
6167	TC	y
6182	GA	r
6185	TC	y
6221	TC	y
6227	TC	y
6253	TC	y
6257	GA	r
6324	GA	r
6366	GA	r
6371	CT	y
6392	TC	y
6473	CT	y
6491	CA	m
6524	TC	y
6548	CT	y
6587	CT	y
6607	TC	y

6680	TC	y
6713	CT	y
6719	TC	y
6734	GA	r
6752	AG	r
6770	AG	r
6776	TC	y
6815	TC	y
6827	TC	y
6875	CA	m
6938	CT	y
6962	GA	r
6989	AG	r
7028	CT	y
7052	AG	r
7055	AG	r
7058	TA	w
7076	AG	r
7146	AG	r
7154	AG	r
7175	TC	y
7196	CA	m
7202	AG	r
7226	GA	r
7256	CT	y
7257	AG	r
7271	AG	r
7274	CT	y
7319	TC	y
7337	GA	r
7347	GA	r
7389	TC	y
7403	AG	r
7424	AG	r
7444	GA	r
7476	CT	y
7493	CT	y
7521	GA	r
7561	TC	y
7571	AG	r
7600	GA	r
7624	TA	w
7645	TC	y
7648	CT	y
7660	TC	y
7664	GA	r
7673	AG	r
7675	CT	y
7693	CT	y

7694	CT	y
7697	GA	r
7744	TC	y
7765	AG	r
7768	AG	r
7771	AG	r
7858	CT	y
7861	TC	y
7864	CT	y
7867	CT	y
7933	AG	r
7948	CT	y
7999	TC	y
8014	AG	r
8020	GA	r
8027	GA	r
8080	CT	y
8087	TC	y
8113	CA	m
8142	CT	y
8149	AG	r
8152	GA	r
8155	GA	r
8185	TC	y
8200	TC	y
8206	GA	r
8248	AG	r
8251	GA	r
8260	TC	y
8269	GA	r
8271-8279	accccctct/-	
8286	TC	y
8292	GA	r
8298	TC	y
8344	AG	r
8387	GA	r
8389	AG	r
8392	GA	r
8404	TC	y
8414	CT	y
8428	CT	y
8448	TC	y
8460	AG	r
8468	CT	y
8472	CT	y
8473	TC	y
8485	GA	r
8545	GA	r
8553	CT	y

8563	AG	r
8566	AG	r
8577	AG	r
8584	GA	r
8618	TC	y
8655	CT	y
8697	GA	r
8701	AG	r
8703	CT	y
8705	TC	y
8709	CT	y
8721	AG	r
8733	TC	y
8764	GA	r
8781	CA	m
8784	AG	r
8790	GA	r
8793	TC	y
8794	CT	y
8805	AG	r
8836	AG	r
8838	GA	r
8856	GA	r
8860	AG	r
8875	TC	y
8877	TC	y
8911	TC	y
8913	AG	r
8928	TC	y
8943	CT	y
8962	AG	r
8994	GA	r
9042	CT	y
9053	GA	r
9055	GA	r
9072	AG	r
9077	TC	y
9090	TC	y
9093	AC	m
9103	TC	y
9114	AG	r
9120	AG	r
9123	GA	r
9136	AG	r
9151	AG	r
9156	AG	r
9174	TC	y
9221	AG	r
9237	GA	r

9242	AG	r
9248	CT	y
9263	AG	r
9272	CT	y
9296	CT	y
9311	TC	y
9325	TC	y
9335	CT	y
9347	AG	r
9355	AG	r
9356	CT	y
9377	AG	r
9402	AC	m
9449	CT	y
9456	AG	r
9477	GA	r
9509	TC	y
9536	CT	y
9540	TC	y
9545	AG	r
9548	GA	r
9554	GA	r
9559	CG	s
9575	GA	r
9591	GA	r
9599	CT	y
9632	AG	r
9647	TC	y
9667	AG	r
9682	TC	y
9698	TC	y
9755	GA	r
9818	CT	y
9822	CA	m
9824	TA	w
9911	CT	y
9932	GA	r
9950	TC	y
9957	TC	y
9966	GA	r
9977	TC	y
10034	TC	y
10086	ACG	v
10086	ACG	v
10115	TC	y
10118	TC	y
10142	CT	y
10151	AG	r
10152	GC	s

10172	GA	r
10182	GC	s
10197	GA	r
10238	TC	y
10253	TC	y
10256	TC	y
10310	GA	r
10313	AG	r
10321	TC	y
10325	GA	r
10358	AG	r
10370	TC	y
10398	AG	r
10400	CT	y
10410	TC	y
10414	GT	k
10427	GA	r
10463	TC	y
10499	AG	r
10505	TC	y
10550	AG	r
10586	GA	r
10589	GA	r
10609	TC	y
10637	CT	y
10640	TC	y
10646	GA	r
10659	CT	y
10664	CT	y
10667	TC	y
10688	GA	r
10736	CT	y
10790	TC	y
10792	AG	r
10793	CT	y
10804	AG	r
10810	TC	y
10819	AG	r
10828	TC	y
10873	TC	y
10876	AG	r
10894	CT	y
10915	TC	y
10920	CT	y
10939	CT	y
10966	TC	y
10984	CG	s
11002	AG	r
11016	GA	r

11017	TC	y
11023	AG	r
11078	AG	r
11092	AG	r
11147	TC	y
11150	GA	r
11167	AG	r
11172	AG	r
11176	GA	r
11177	CT	y
11215	CT	y
11251	AG	r
11257	CT	y
11296	CT	y
11299	TC	y
11332	CT	y
11362	AG	r
11365	TC	y
11377	GA	r
11467	AG	r
11476	CT	y
11536	CT	y
11590	AG	r
11611	GA	r
11641	AG	r
11653	AG	r
11654	AG	r
11674	CT	y
11701	TC	y
11719	GA	r
11722	TC	y
11767	CT	y
11812	AG	r
11854	TC	y
11884	AG	r
11887	GA	r
11893	AG	r
11899	TC	y
11909	AG	r
11914	GA	r
11944	TC	y
11947	AG	r
11959	AG	r
11963	GA	r
11969	GA	r
12007	GA	r
12049	CT	y
12070	GA	r
12083	TG	k

12121	TC	y
12134	TC	y
12153	CT	y
12172	AG	r
12175	TC	y
12234	AG	r
12236	GA	r
12239	CT	y
12248	AG	r
12308	AG	r
12346	CT	y
12358	AG	r
12361	AG	r
12372	GA	r
12373	AG	r
12397	AG	r
12406	GA	r
12414	TC	y
12477	TC	y
12501	GA	r
12507	AG	r
12519	TC	y
12528	GA	r
12540	AG	r
12612	AG	r
12630	GA	r
12633	CT	y
12635	TC	y
12669	CT	y
12672	AG	r
12693	AG	r
12705	CT	y
12720	AG	r
12738	TC	y
12768	AG	r
12771	GA	r
12810	AG	r
12822	AG	r
12850	AG	r
12879	TC	y
12882	CT	y
12930	AT	w
12940	GA	r
12948	AG	r
12967	AC	m
12972	AG	r
12999	AG	r
13020	TC	y
13059	CT	y

13068	AG	r
13101	AC	m
13104	AG	r
13105	AG	r
13135	GA	r
13143	TC	y
13145	GA	r
13149	AG	r
13194	GA	r
13197	CT	y
13212	CT	y
13221	AG	r
13263	AG	r
13276	AG	r
13281	TC	y
13368	.GA	r
13440	CG	s
13477	GA	r
13485	AG	r
13494	CT	y
13500	TCG	b
13500	TCG	b
13506	CT	y
13512	AG	r
13563	AG	r
13590	GA	r
13594	AG	r
13602	TC	y
13611	AG	r
13617	TC	y
13641	TC	y
13650	CT	y
13651	AG	r
13660	AG	r
13708	GA	r
13722	AG	r
13734	TC	y
13759	GA	r
13780	AG	r
13789	TC	y
13803	AG	r
13812	TC	y
13818	TC	y
13819	TC	y
13827	AG	r
13880	CA	m
13886	TC	y
13914	CA	m
13924	CT	y

13927	AT	w
13928	GC	s
13958	GC	s
13965	TC	y
13966	AG	r
13980	GA	r
14000	TA	w
14016	GA	r
14020	TC	y
14022	AG	r
14025	TC	y
14034	TC	y
14059	AG	r
14070	AGT	d
14070	AGT	d
14088	TC	y
14094	TC	y
14097	CT	y
14118	AG	r
14128	AG	r
14148	AG	r
14152	AG	r
14167	CT	y
14178	TC	y
14182	TC	y
14200	TC	y
14203	AG	r
14209	AG	r
14212	TC	y
14215	TC	y
14221	TC	y
14233	AG	r
14272	CG	s
14284	CT	y
14308	TC	y
14311	TC	y
14318	TC	y
14319	TC	y
14371	TC	y
14374	TC	y
14384	GC	s
14455	CT	y
14459	GA	r
14470	TC	y
14484	TC	y
14488	TC	y
14502	TC	y
14560	GA	r
14566	AG	r

14569	GA	r
14571	TA	w
14580	AG	r
14587	AG	r
14605	AG	r
14668	CT	y
14693	AG	r
14766	CT	y
14769	AG	r
14783	TC	y
14793	AG	r
14798	TC	y
14812	CT	y
14836	AG	r
14861	GA	r
14862	CT	y
14905	GA	r
14911	CT	y
14971	TC	y
14974	CG	s
14979	TC	y
15016	CT	y
15034	AG	r
15043	GA	r
15110	GA	r
15113	AG	r
15115	TC	y
15136	CT	y
15172	GA	r
15204	TC	y
15217	GA	r
15218	AC	m
15229	TC	y
15238	CG	s
15244	AG	r
15257	GA	r
15261	GA	r
15301	GA	r
15317	GA	r
15318	CT	y
15323	GA	r
15326	AG	r
15346	GA	r
15358	AG	r
15431	GA	r
15442	AG	r
15452	CA	m
15466	GA	r
15470	TC	y

15487	AT	w
15497	GA	r
15514	TC	y
15519	TC	y
15535	CT	y
15607	AG	r
15626	CT	y
15629	TC	y
15646	CT	y
15661	CT	y
15663	TC	y
15670	TC	y
15724	AG	r
15731	GA	r
15746	AG	r
15766	AG	r
15784	TC	y
15793	CT	y
15803	GA	r
15806	GA	r
15812	GA	r
15824	AG	r
15833	CT	y
15849	CT	y
15884	GC	s
15900	TC	y
15904	CT	y
15907	AG	r
15924	AG	r
15927	GA	r
15928	GA	r
15930	GA	r
15932	TC	y
15939	CT	y
15941	TC	y
15942	TC	y
15968	TC	y
16017	TC	y
16038	AG	r
16051	AG	r
16069	CT	y
16071	CT	y
16075	TC	y
16086	TC	y
16093	TC	y
16108	CT	y
16111	CT	y
16114	CA	m
16124	TC	y

16126	TC	y
16129	GCA	v
16129	GCA	v
16140	TC	y
16144	TC	y
16145	GA	r
16147	CT	y
16148	CT	y
16153	GA	r
16162	AG	r
16163	AC	m
16166	AC	m
16167	CT	y
16168	CT	y
16169	CT	y
16171	AG	r
16172	TC	y
16175	AG	r
16176	CT	y
16182	AC	m
16183	AC	m
16184	CT	y
16185	CT	y
16186	CT	y
16187	CT	y
16188	CAG	v
16188	CAG	v
16189	TC	y
16192	CT	y
16193	CT	y
16207	AG	r
16209	TC	y
16212	AG	r
16213	GA	r
16214	CT	y
16217	TC	y
16219	AG	r
16223	CT	y
16224	TC	y
16227	AG	r
16229	TC	y
16230	AG	r
16231	TC	y
16232	CT	y
16234	CT	y
16235	AG	r
16239	CT	y
16241	AG	r
16242	CT	y

16243	TC	y
16245	CT	y
16247	AG	r
16249	TC	y
16254	AC	m
16255	GA	r
16256	CT	y
16257	CT	y
16258	AG	r
16260	CT	y
16261	CT	y
16264	CT	y
16265	AC	m
16266	CT	y
16268	CT	y
16270	CT	y
16271	TC	y
16274	GA	r
16278	CT	y
16284	AG	r
16286	CG	s
16287	CT	y
16288	TC	y
16290	CT	y
16291	CT	y
16292	CT	y
16293	AG	r
16294	CT	y
16296	CT	y
16298	TC	y
16304	TC	y

16309	AG	r
16311	TC	y
16316	AG	r
16317	AT	w
16318	AT	w
16319	GA	r
16320	CT	y
16324	TC	y
16325	TC	y
16326	AG	r
16327	CT	y
16343	AG	r
16344	CT	y
16354	CT	y
16355	CT	y
16356	TC	y
16357	TC	y
16360	CT	y
16362	TC	y
16366	CT	y
16368	TC	y
16390	GA	r
16391	GA	r
16399	AG	r
16438	GA	r
16439	CA	m
16483	GA	r
16519	TC	y
16527	CT	y

シーケンス・リストの参照

配列番号１は、表3において一覧を示すにつれて、1は本発明のヌクレオチド対立遺伝子を含んでいる理論上の人間のmtDNAゲノム・シーケンスである。

【０１３１】

配列番号2は、ケンブリッジSequence（GenBankAccession数J01415）と呼ばれている人間のmtDNA参照シーケンスである。

【図面の簡単な説明】

【０１３２】

【図１】図1は、mtDNAがシーケンスおよび2つの霊長類シーケンスを完了することを104人の人間のコンセンサス隣-接続している木に明らかにする。数は、つまみ革値（500の全体のつまみ革の%は、複製する）と一致する（フェルゼンシュタイン（J.（1993の）PHYLIP（系統発生InferencePackage））3.53c.著者、Genetics部、ワシントン大学、シアトル、WAのそばの配布する）。最大Likelihood（mL）およびUPGMAは、大陸に特有のmtDNAハプログループに関して一貫した分岐命令を与えた。シーケンス：11-53 :Genbank AF346963-AF347015 (4);E21U：GenbankX93334（AlLla）：Genbank D38112（カム修正）：一致することを修正される（R. M.アンドルーズほか（ネイチャーGenetics23,147（1999）））Genbank NC001807;残りは、Application Binary Interface 377を使用している本発明において発生する48のシーケンスである。9bp削除を除くすきまおよび削除と同様に、疾患に関係した患者サンプルの特定の突然変異は、この分析から締め出された（ヌクレオチドが、8281まで8272を配置する（np））。ハプログループA、B、C、DおよびXは、ユーラシアおよび米州から引き出された。ハプログループ名は、大文字によって示される。P. paniscusおよびP.穴居人mtDNAシーケンスが、outgroupsとして使われた。L0およびL1がすでに含む

ハプログループは、L1aおよびmtDNAsに、それぞれL1b（Y.S.チェンほか、Human Genetics 66アメリカのジャーナル、1362-1383（2000））を割り当てた。
【図２Ａ】図2Aは、世界中で人間のハプログループの移動を示す。+/-、+/+、or-/-同等Dde110394およびAlu 110397。*は、Rsa 116329に等しい。突然変異率は、2.2-2である。9%の／100万年。時間評価は、YBP（何年も前にある）である。
【図２Ｂ】図2Bは、世界中で人間のハプログループの移動を示す。+/-、+/+、or-/-同等Dde110394およびAlu 110397。*は、Rsa 116329に等しい。突然変異率は、2.2-2である。9%の／100万年。時間評価は、YBP（何年も前にある）である。
【図３】図3は、21の主要人間のハプログループ、21の副ハプログループおよびいくつかのマクロハプログループを記載しているヌクレオチド対立遺伝子の一覧を示している分岐図を示す。左上のグループは、それらの権利に対立遺伝子によって記載されている。垂直棒は、棒の左に対する各々のグループが棒の右側に対立遺伝子の全てを有することを示す。
【図４】図4は、選択的な制約を示す（kc値）哺乳類の種の中の比較を有するmtDNAタンパク質遺伝子の。有意性（P< 0.05）は、分散分析、ｔ検定またはTukey-Kramer Multiple Comparisons試験を使用して決定された。使用する大部分のプログラムは、DNAspからある（J.RozasおよびR. Rozas（（1999の）Bioinformatics 15）：174-5).DNAシーケンス相違は、DIVERGEプログラム（ウィスコンシンPackageVersion 10.0、Genetics Computer Group（GCG）、マディソン、WI）を使用して分析された。全ての13のmtDNA遺伝子のために、データは人間、P.穴居人と比較した人間、P.paniscusと比較した人間および霊長類の9つの種のために示される。ATP6およびATP8だけのために、データはまた、哺乳類の14の種のために示される。
【配列表】

SEQ ID NO:1

```
   1 gatcacaggt ctatcaccct attaaccact cacgggagct ctccatgcat ttggtatttt
  61 cgtytggggg gyrtgcacgc gatagcatyg cgrgmcgctg gagccggagc accytatgtc
 121 gcagtatctg tctttgattc ctrccycaty yyrttattta tcgcacctac rttcaatatty
 181 ayrgdmgavc atayhtayyr aagygtrytr aytartyaat gcttrtrrga cataryaata
 241 acaattraay gyctgcacag ccrctttcca cacagacatc ataacaaaaa rtttycrcca
 301 aacccccccct cccccrvtty tggcyacagc acttaaacay atctctgcca aaccccraaa
 361 acaaagaacc ctracaccag cctaaccaga tttcaaattk tatctttwgg cggtatgyac
 421 ttttaacagt caccccccaa ctaacacatt attttyccct cycaytycca yactactaay
 481 cycatcaaya carcccccrc ccatcctrcc cagcacacac acaccgctgc taaccccata
 541 ccccgaacca accaaacccc aaagacaccc cccacagttt atgtagctta ccycctyaaa
 601 gcaatacact gaaaatgttt agacgggctc acatcacccc ataaacaaat aggtttggtc
 661 ctrgcctttc tattagcycy tagtaagatt acacatgcaa gcatccccry tccagtgagt
 721 ycaccctcta aatcaccacg atcaaaaggr acaagcatca agcacgcarc aatgcagctc
 781 aaaacgctta gcctagccac acccccacgg gaaacagcag tgatwarcct ttagcaataa
 841 acgaaagtty aactaagcta tactaacccc agggttggtc aatttcgtgc cagccaccgc
 901 ggtcacacga ttaacccaag ycaatagaar ccggcgtaaa gagtgtttta gatcaccccc
 961 bccccaataa agctaaaact cacctgagtt gtaaaaaact ccagttgaca caaaatarac
1021 tacgaaagtg gctttaacat rtctgaayac acaatagcta agacccaaac tgggattaga
1081 taccccacta tgcttagccc taaacctcaa cagttaaayc aacaaaactg ctcgccagaa
1141 cactacgagc cacagcttaa aactcaaagg acctggcggt gcttcatayc cctctagagg
1201 agcctgttct gtaatcgata aaccccgatc aacctcacca ccycttgctc agcctatata
1261 ccgccatctt cagcaaaccc tgatgaaggy tacaaagtaa gcgcaagtac ccacgtaaag
1321 acgttaggtc aaggtgtagc ccatgaggtg gcaagaaatg ggctacattt tctaccccag
1381 amaaactacga tagcccttat gaaacytaag ggtcraaggy ggatttagca gtaaactrag
1441 artagagtgc ttagttgaac agggccctga agcgcgtaca caccgcccgt caccctcctc
1501 aartatactt caaaggacat ttaactaaaa cccctacgca tttatataga ggagacaagt
1561 cgtaacatgg taagtgtact ggaaagtgca cttggacraa ccagagtgta gcttaacaca
1621 aagcacccaa cttacactta ggagatttca acttaacttg accgctctga gctaaaccta
1681 gccccaaacc cactccaccy taytaycara caacyttarc caaaccattt acccarayaa
1741 agtataggcg atagaaattg aaacctggcg caatagatay agtaccgcaa gggaaagatg
1801 aaaaattata rccaagcata atatagcaag gactaacccc tataccttct gcataatgaa
1861 ttaactagaa ataactttgc aaggagarcc aaagctaaga cccccgaaac cagacgagct
1921 acctaaraac agctaaaaga gcacacccgt ctatgtagca aaatagtggg aagatttata
1981 ggtagaggcg acaaacctay cgagcctggt gatagctggt tgtccaagat agaatcttag
2041 ttcaacttta aatttgcccr cagaaccctc taaatcccct tgtaaattta aytgttagtc
2101 caaagaggaa cagctctttg gacactagga aaaaaccttg tagagagagt aaaaaattta
2161 acacccatag taggcctaaa agcagccacc aattaagaaa gcgttcaagc tcaacaccca
2221 ctacctaaaa aatcccaaac atatvactga actcctcaca ccmaattgga ccaatctatc
2281 accctataga agaactaatg ttagtatrag taacatgaaa acattctcct cygcataagc
2341 ctgcgtcaga tyaaaacrct gaactgacaa ttaacagccy aatatctaca atcaaccaac
2401 aagtcattat tacccycact gtcaacccaa cacaggcatg ctcataagga aaggttaaaa
2461 aaagtaaaag gaactcggca aaycttaccc cgcctgttta ccaaaaacat cacctctagc
2521 atcaccagta ttagaggcac cgcctgccca gtgacacatg tttaacggcc gcggtaccct
2581 raccgtgcaa aggtagcata atcacttgtt ccttaaatag ggacctgtat gaatggctyc
2641 acgagggtty agctgtctct tacttttaac cagtgaaatt gacctgcccg tgaagaggcg
2701 ggcatracac agcaagacga gaagacccta tggagcttta atttattaat gcaarcarta
2761 cctaacarac ccacaggtcc taaactacya arcctgcatt aaaaatttcg gttggggcga
2821 cctcggagca gaaycmaacc tccgagcagt acatgcyaag acytcaccag tcaaagcgaa
2881 ctacyatact caattgatcc aataacttga ccaacggaac aagttaccct agggataaca
2941 gcgcaatcct attctagagt ccatatcaac aatagggttt acgacctcga tgttggatca
3001 ggacatcccr atggtgcagc cgctattaaa ggttcgtttg ttcaacgatt aaagtcctac
3061 gtgatctgag ttcagaccgg agyaatccag gtcggtttct atcta-cttc aaattcctcc
3121 ctgtacgaaa ggacaagaga aataaggcct acttcacaaa gcgccttccc ccgtaaatga
3181 tatcatctca acttagyatw ayaycyacac ccacccaaga rcagggtttg ttaagatggc
3241 agagcccggt aatcgcataa aacttaaaac tttacagtca gaggttcaay tcctcttctt
3301 aacaacayac ccatgrccaa cctcctactc ctcattgtac ccattctaat cgcaatggca
3361 ttcctaatgc tyaccgaacg aaaaattcta ggcyatatac aactacgcaa aggccccaac
```

```
3421 gttgtaggcc cctacggrct actacaaccy ttcgctgacg ccataaaact cttcaccaar
3481 gagcccctaa aacccgccac atctrccatc acyctvtaca tcaccgcccc gaccttrgct
3541 ctcaccatyg chcttctact atgarccccc ctccccatac ccaaccccct ggtyaacctc
3601 aacctaggcc tcctatttat tctagccacc tctagcctag ccgyttactc aatcctctga
3661 tcaggrtgag catcaaactc aaactacgcc ctratcggyg cactgcgagc agtagcccar
3721 acaatctcat atgaagtcac cctagccatc attctrctat caacattact aataagtggc
3781 tcctttaacc tctccncccct tatcacarca caagarcacc tctgattact cctrccatca
3841 tgrcccytgg ccataatatg atttayctcc acactagcag agaccaaccg aaccccccttc
3901 gaccttgccg aaggggartc mgaactrgtc tcaggcttca acatcgaata cgccgcaggc
3961 cccttcgccy tattcttcat rgccgaatac acaaacatta ttataataaa caccctcacc
4021 actayaatct tcctaggaay aacrtatrac gcactctccc ctgaactcta cacaacatat
4081 tttgtyacca agaccctact tctracctcc ctgttcytat grrttcgaac agcatacccc
4141 cgattccgct acgaccarct catacacctc ctatgaaaaa acttcctacc actcacccta
4201 gcrttactta tatgayatgt ytccrtaccc aytacaatct ccagcatycc ccctcaaacc
4261 taagaaatat gtctgataaa agagttactt tgatagagta aataatagga gyttaaaccc
4321 ccttatttct aggacyatga gaatcgaacc catccctgag aatccaaaay tctccgtgcc
4381 acctatcrca ccccatccta aagtaaggtc agctaaataa gctatcgggc ccataccccg
4441 aaaatgttgg ttawaccctt cccgtactaa ttaatcccct ggcccaaccc rtcatctact
4501 ctaccrtytt trcaggcaca ctcatcachg cgctaagctc rcactgattt tttacctgag
4561 taggcctaga aataaacatr ctagcyttta ttccarttct aaccaaaaaa ataaaccctc
4621 gttccacaga agctgccatc aagtayttcc tcacrcaagc aaccgcatcc ataatccttc
4681 taatagcyat cctcytcaac aatatactct ccggrcaatg aaccataacc aatactacca
4741 aycaatactc atcattaata atcatartrg ctatagcaat aaaactagga atagccccct
4801 ttcacttctg agtcccagar gttrcccaag gcrcccctct racatccggc ctgcttcttc
4861 tcacatgaca aaaactagcc ccyatctcaa tcatatacca aatctcyccc tcactaracg
4921 taagccttct cctcactctc tcaatcttat ccatcatagy aggcagttga ggtggaytaa
4981 accaAaccca gctrcgcaaa atcytagcat actcctcaat tacccayata ggatgrataa
5041 takcarttct accgtacaac ccyaacataa ccattcttaa tttaactatt tatatyatcc
5101 taactacyac cgcattccta ctactcaact taaactccag caccacracc ctrctactat
5161 ctcgcacctg aaacaagmta acatgactaa caccсttaat tccatccacc ctcctctccc
5221 taggaggcct rcccccrcta accggctttt tgccyaaatg grycattatc gaagaattca
5281 caaaraacaa tagcctcaty atccccacca tcatagccac catcaccctm mttaacctct
5341 acttctacct acgcctaatc tactccacct caatcacact actccccatr tcyaacaacg
5401 taaaaataaa atgacarttt gaacayacaa aacccacccc aytcctcccc acactcatcr
5461 ccctyaccac rctactccta cctatctccc cyttyatact aataatctta tagaaattta
5521 ggttaaatac agaccaagag ccttcaaagc cctcagtaag ttgcaatact taatttctgy
5581 racagctaag gactgcaaaa ycycaytctg catcaactga acgcaaatca gcyactttaa
5641 ttaagctaag ccctyactag accaatggga cttaaaccca caaacactta gttaacagct
5701 aagcacccta rtcaactggc ttcaatctac ttctcccgcc gccgggaaaa aaggcgggag
5761 aagccccggc agrtttgaag ctgcttcttc gaatttgcaa ttcaatatga raaycacctc
5821 rgagcyggta aaaagaggcc tarcccctgt ctttagattt acagtccaat gcttcactca
5881 gccattttac ctcaccccca ctgatgttcg ccgaccgttg actattctct acaaaccaca
5941 aagacattgg racactatac ctattattcg gcgcatgagc tggrgtycta ggcacagctc
6001 taagcctcct tattcgagcc gagctrggyc agccaggcaa ccttytaggt aacgaccaca
6061 tctacaacgt yatcgtyaca gcccatgcat ttgtaataat cttyttcata gtaataccca
6121 tcataatcgg aggctttggc aactgactar tyccccctaat aatyggygcc cccgatatgg
6181 crttycccccg cataaacaac ataagcttct gactcttacc yccctcyctc ctactcctgc
6241 tcgcatctgc tayagtrgag gccggagcag gaacaggttg aacagtctac cctcccttag
6301 cagggaacta ctcccaccct ggarcctccg tagacctaac catcttctcc ttacacctag
6361 caggtrtctc ytctatctta ggggccatca ayttcatcac aacaattatc aatataaaac
6421 cccctgccat aacccaatac caaacgcccc tcttcgtctg atccgtccta atyacagcag
6481 tcctacttct mctatctctc ccagtcctag ctgctggcat cacyatacta ctaacagacc
6541 gcaacctyaa caccaccttc ttcgaccccg ccggaggagg agacccyatt ctataccaac
6601 acctatyctg atttttcggt caccctgaag tttatattct tatcctacca ggcttcggaa
6661 taatctccca tattgtaacy tactactccg gaaaaaaaga accatttgga tayataggya
6721 tggtctgagc tatratatca attggcttcc trgggtttat cgtgtgagcr caccayatat
6781 ttacagtagg aatagacgta gacacacgag catayttcac ctccgcyacc ataatcatcg
6841 ctatccccac cggcgtcaaa gtatttagct gactmgccac actccacgga agcaatatga
6901 aatgatctgc tgcagtgctc tgagccctag gattcatytt tcttttcacc gtaggtggcc
```

6961 tractggcat tgtattagca aactcatcrc tagacatcgt actacacgac acgtactacg
7021 ttgtagcyca cttccactat gtcctatcaa trggrgcwgt atttgccatc ataggrggct
7081 tcattcactg atttccccta ttctcaggct acaccctaga ccaaacctac gccaaaatcc
7141 atttcrctat catrttcatc ggcgtaaatc taacyttctt cccacaacac tttctmggcc
7201 trtccggaat gccccgacgt tactcrgact accccgatgc atacaccaca tgaaayrtcc
7261 tatcatctgt rggytcattc atttctctaa cagcagtaat attaataatt ttcatgatyt
7321 gagaagcctt cgcttcraag cgaaaartcc taatagtaga agaaccctcc ataaacctgg
7381 agtgactaya tggatgcccc ccrccctacc acacattcga agarcccgta tacataaaat
7441 ctaracaaaa aaggaaggaa tcgaaccccc caaagytggt ttcaagccaa ccycatggcc
7501 tccatgactt tttcaaaaag rtattagaaa aaccatttca taactttgtc aaagttaaat
7561 yataggctaa rtcctatata tcttaatggc acatgcagcr caagtaggtc tacaagacgc
7621 tacwtcccct atcatagaag agctyatyac ctttcatgay cacrccctca tartyatttt
7681 ccttatctgc ttyytartcc tgtatgccct tttcctaaca ctcacaacaa aactaactaa
7741 tacyaacatc tcagacgctc aggaratrga raccgtctga actatcctgc ccgccatcat
7801 cctagtcctc atcgccctcc catccctacg catcctttac ataacagacg aggtcaayga
7861 yccytcyctt accatcaaat caattggcca ccaatggtac tgaacctacg agtacaccga
7921 ctacggcgga ctratcttca actcctayat acttccccca ttattcctag aaccaggcga
7981 cctgcgactc cttgacgtyg acaatcgagt agtrctcccr attgaarccc ccattcgtat
8041 aataattaca tcacaagacg tcttgcactc atgagctgty cccacaytag gcttaaaaac
8101 agatgcaatt ccmggacgtc taaaccaaac cactttcacc gytacacgrc crggrgtata
8161 ctacggtcaa tgctctgaaa tctgyggagc aaaccacagy ttcatrccca tcgtcctaga
8221 attaattccc ctaaaaatct ttgaaatrgg rcccgtatty accctatarc accccctcta
8281 cccccyctag arcccacygt aaagctaact tagcattaac cttttaagtt aaagattaag
8341 agarccaaca cctctttaca gtgaaatgcc ccaactaaat actaccrtrt grcccaccat
8401 aatyaccccc ataytcctta cactattyct catcacccaa ctaaaaayat taaacacaar
8461 ctaccacyta cyyccctcac caaarcccat aaaaataaaa aattataaca aaccctgaga
8521 accaaaatga acgaaaatct gttcrcttca ttyattgccc ccrcartcct aggcctrccc
8581 gccrcagtac tgatcattct atttccccct ctattgaycc ccacctccaa atatctcatc
8641 aacaaccgac taatyaccac ccaacaatga ctaatcaaac taacctcaaa acaaatrata
8701 rcyayacaya acactaaagg rcgaacctga tcycttatac tagtatcctt aatcattttt
8761 attrccacaa ctaacctcct mggrctcctr ccyyactcat ttacrccaac cacccaacta
8821 tctataaacc tagccrtrgc catcccctta tgagcrggcr cagtgattat aggcytycgc
8881 tctaagatta aaaatgccct agcccacttc ytrccacaag gcacaccyac acccсttatc
8941 ccyatactag ttattatcga arccatcagc ctactcattc aaccaatagc cctrgccgta
9001 cgcctaaccg ctaacattac tgcaggccac ctactcatgc ayctaattgg aarcrccacc
9061 ctagcaatat craccaytaa ccttccctcy acmcttatca tcytcacaat tctrattctr
9121 ctractatcc tagaartcgc tgtcgcctta rtccargcct acgttttcac actyctagta
9181 agcctctacc tgcacgacaa cacataatga cccaccaatc rcatgcctat catatartaa
9241 arcccagycc atgaccccta acrggggccc tytcagccct cctaatgacc tccggyctag
9301 ccatgtgatt ycacttccac tccayaacgc tcctyatact aggcctrcta accaryacac
9361 taaccatata ccaatgrtgg cgcgatgtaa cacgagaaag cmcataccaa ggccaccaca
9421 caccacctgt ccaaaaaggc cttcgatayg ggatartcct atttattacc tcagaarttt
9481 ttttcttcgc aggatttttc tgagccttyt accactccag cctagcccct accccycaay
9541 taggrggrca ctgrccccsa acaggcatca ccccrctaaa tcccctagaa rtcccactyc
9601 taaacacatc cgtattactc gcatcaggag trtcaatcac ctgagcycac catagtctaa
9661 tagaaarcaa ccgaaaccaa ayaattcaag cactgctyat tacaatttta ctgggtctct
9721 attttaccct cctacaagcc tcagagtact tcgartctcc cttcaccatt tccgacggca
9781 tctacggctc aacatttttt gtagccacag gcttccaygg amtwcacgtc attattggct
9841 caactttcct cactatctgc ttcatccgcc aactaatatt tcactttaca tccaaacatc
9901 actttggctt ygaagccgcc gcctgatact grcattttgt agatgtggty tgactayttc
9961 tgtatrtctc catctaytga tgagggtctt actcttttag tataaatagt accgttaact
10021 tccaattaac tagytttgac aacattcaaa aaagagtaat aaacttcgcc ttaattttaa
10081 taatcvacac cctcctagcc ttactactaa taatyatyac attttgacta ccacaactca
10141 ayggctacat rsaaaaatcc acccсttacg artgcggctt csaccctata tccccсrccc
10201 gcgtcccttt ctccataaaa ttcttcttag tagctatyac cttcttatta ttygayctag
10261 aaattgccct ccttttaccc ctaccatgag ccctacaaac aactaacctr ccrctaatag
10321 ytatrtcatc cctcttatta atcatcatcc tagccctrag tctggcctay gagtgactac
10381 aaaaaggatt agactgarcy gaattggtay atakttaaa caaaacraat gatttcgact
10441 cattaaatta tgataatcat atytaccaaa tgcccctcat ttacataaat attatactrg

10501 cattyaccat ctcacttcta ggaatactag tatatcgctc acacctcatr tcctccctac
10561 tatgcctaga aggaataata ctatcrctrt tcattatagc tactctcaya accctcaaca
10621 cccactccct cttagcyaay attgtrccta ttgccatayt agtyttygcc gcctgcgaag
10681 cagcggtrgg cctagcccta ctagtctcaa tctccaacac atatggccta gactaygtac
10741 ataacctaaa cctactccaa tgctaaaact aatcgtccca acaattatay trytaccact
10801 gacrtgacty tccaaaaarc acataatytg aatcaacaca accacccaca gcctaattat
10861 tagcatcatc ccyctrctat tttttaacca aatyaacaac aacctattta gctgytcccy
10921 aacctttcc tccgacccyc taacaacccc cctcctaata ctaacyacct gactcctacc
10981 cctsacaatc atggcaagcc arcgccactt atccarygaa ccrctatcac gaaaaaaact
11041 ctacctctct atactaatct ccctacaaat ctccttartt ataacattca crgccacaga
11101 actaatcata ttttatatct tcttcgaaac cacacttatc cccaccytgr ctatcatcac
11161 ccgatgrggc arccarycag aacgcctgaa cgcaggcaca tacttcctat tctayaccct
11221 agtaggctcc cttcccctac tcatcgcact ratttayact cacaacaccc taggctcact
11281 aaacattcta ctactyacyc tcactgccca agaactatca aactcctgag cyaaCaactt
11341 aatatgacta gcttacacaa trgcytttat agtaaarata cctctttacg gactccactt
11401 atgactccct aaagcccatg tcgaagcccc catcgctggg tcaatagtac ttgccgcagt
11461 actcttraaa ctaggyggct atggtataat acgcctcaca ctcattctca accccctgac
11521 aaaacacata gcctayccct tccttgtact atccctatga ggcataatta taacaagctc
11581 catctgcctr cgacaaacag acctaaaatc rctcattgca tactcttcaa tcagccacat
11641 rgccctcgta gtrrcagcca ttctcatcca aacyccctga agcttcaccg gcgcagtcat
11701 yctcataatc gcccacggrc tyacatcctc attactattc tgcctagcaa actcaaacta
11761 cgaacgyact cacagtcgca tcataatcct ctctcaagga cttcaaactc trctcccact
11821 aatagctttt tgatgacttc tagcaagcct cgcyaacctc gccttacccc ccactattaa
11881 cctrctrgga garctctcyg tgctagtarc cacrttctcc tgatcaaata tcactctcct
11941 actyacrgga ctcaacatrc tartcacarc cctatactcc ctctacatat ttaccacaac
12001 acaatgrggc tcactcaccc accacattaa caacataaaa ccctcattya cacgagaaaa
12061 caccctcatr ttcatacacc takcccccat tctcctccta tccctcaacc ccgacatcat
12121 yaccgggttt tccycttgta aatatagttt aaycaaaaca tcagattgtg artcygacaa
12181 cagaggctta cgacccctta tttaccgaga aagctcacaa gaactgctaa ctcrtrccyc
12241 catgtctrac aacatggctt tctcaacttt taaaggataa cagctatcca ttggtcttag
12301 gccccaaraa ttttggtgca actccaaata aaagtaataa ccatgyacac tactatarcc
12361 rccctaaccc trrcttccct aattcccccc atccttrcca ccctcrttaa cccyaacaaa
12421 aaaaactcat acccccatta tgtaaaatcc attgtcgcat ccacctttat tatcagyctc
12481 ttccccacaa caatattcat rtgcctrgac caagaagtya ttatctcraa ctgacactgr
12541 gccacaaccc aaacaaccca gctctcccta agcttcaaac tagactactt ctccataata
12601 ttcatccctg trgcattgtt cgttacatgr tcyaycatag aattctcact gtgatatata
12661 aactcagayc craacattaa tcagttcttc aartatctac tcatyttcct aattaccatr
12721 ctaatcttag ttaccgcyaa caacctattc caactgttca tcggctgrga rggcgtagga
12781 attatatcct tcttgctcat cagttgatgr tacgcccgag crgatgccaa cacagcagcc
12841 attcaagcar tcctatacaa ccgtatcggc gatatcggyt tyatcctcgc cttagcatga
12901 tttatcctac actccaactc atgagacccw caacaaatar cccttctraa cgctaatcca
12961 agcctcmccc crctactagg cctcctccta gcagcagcrg gcaaatcagc ccaattaggy
13021 ctccacccct gactcccctc agccatagaa ggccccacyc cagtctcrgc cctactccac
13081 tcaagcacta tagttgtagc mggrrtcttc ttactcatcc gcttccaccc cctarcagaa
13141 aayarcccrc taatccaaac tctaacacta tgcttaggcg ctatcaccac tctrttygca
13201 gcagtctgcg cycttacaca raatgacatc aaaaaaatcg tagccttctc cacttcaagt
13261 carctaggac tcatartagt yacaatcggc atcaaccaac cacacctagc attcctgcac
13321 atctgtaccc acgccttctt caaagccata ctatttatgt gctccggrtc catcatccac
13381 aaccttaaca atgaacaaga tattcgaaaa ataggaggac tactcaaaac catacctcts
13441 acttcaacct ccctcaccat tggcagccta gcattarcag gaatrccttt cctyacaggb
13501 ttctaytcca argaccacat catcgaaacc gcaaacatat catacacaaa cgcctgagcc
13561 ctrtctatta ctctcatcgc tacctccctr acargcgcct ayagcactcg rataatyctt
13621 ctcaccctaa caggtcaacc ycgcttcccy rcccttactr acattaacga aaataacccc
13681 accctactaa accccattaa acgcctgrca gccggaagcc trttcgcagg attyctcatt
13741 actaacaaca tttccccrc atcccccttc caaacaacar tccccctcya cctaaaactc
13801 acrgccctcg cygtcacyyt cctaggrctt ctaacagccc tagacctcaa ctacctaacc
13861 aacaaactta aaataaaatm cccacyatgc acattttatt tctccaacat actmggattc
13921 tacyctwsca tcacacaccg cacaatcccc tatctagscc ttctyrcgag ccaaaacctr
13981 cccctactcc tcctagaccw aacctgacta gaaaarctay trccyaaaac aatytcacag

```
14041 caccaaatct ccacctccrt catcacctcd acccaaaaag gcataatyaa actytayttc
14101 ctctctttct tcttcccrct catcctarcc ctactcctaa tcacatarcc trttcccccg
14161 agcaatytca attacaayat ayacaccaac aaacaatgty carccagtra cyacyactaa
14221 ycaacgccca tartcataca aagcccccgc accaatagga tcctcccgaa tsaaccctga
14281 cccytctcct tcataaatta ttcagctycc yacactayya aagtttacca caaccaccac
14341 cccatcatac tctttcaccc acagCacCaa yccyacctcc atcsctaacc ccactaaaac
14401 actcaccaag acctcaaccc ctgaccccca tgcctcagga tactcctcaa tagcyatcrc
14461 tgtagtatay ccaaagacaa ccaycatycc ccctaaataa aytaaaaaaa ctattaaacc
14521 catataacct cccccaaaat tcagaataat aacacacccr accacrccrc waacaatcar
14581 tactaarccc ccataaatag gagarggctt agaagaaaac cccacaaacc ccattactaa
14641 acccacactc aacagaaaca aagcatayat cattattctc gcacggacta carccacgac
14701 caatgatatg aaaaaccatc gttgtatttc aactacaaga acaccaatga ccccaatacg
14761 caaaaytarc cccctaataa aaytaattaa ccrctcaytc atcgacctcc cyaccccatc
14821 caacatctcc gcatgrtgaa acttcggctc actccttggc ryctgcctga tcctccaaat
14881 caccacagga ctattcctag ccatrcacta ytcaccagac gcctcaaccg ccttttcatc
14941 aatcgcccac atcactcgag acgtaaatta yggstgaayc atccgctacc ttcacgccaa
15001 tggcgcctca atattytttta tctgcctctt cctrcacatc ggrcgaggcc tatattacgg
15061 atcatttctc tactcagaaa cctgaaacat cggcattatc ctcctgcttr carcyatagc
15121 aacagccttc ataggytatg tcctcccgtg aggccaaata tcattctgag grgccacagt
15181 aattacaaac ttactatccg ccayccata cattggrmca gacctagtyc aatgaatstg
15241 aggrggctac tcagtaraca rtcccaccct cacacgattc tttacctttc acttcatctt
15301 rcccttcatt attgcarycc tarcarcact ccacctccta ttcttrcacg aaacgggrtc
15361 aaacaacccc ctaggaatca cctcccattc cgataaaatc accttccacc cttactacac
15421 aatcaaagac rccctcggct trcttctctt cmttctctcc ttaatracay taacactatt
15481 ctcaccwgac ctcctargcg acccagacaa ttayacccya gccaacccct taaayacccc
15541 tccccacatc aagcccgaat gatatttcct attcgcctac acaattctcc gatccgtccc
15601 taacaarcta ggaggcgtcc ttgccytayt actatccatc ctcatyctag caataatccc
15661 yaycctccay atatccaaac aacaaagcat aatatttcgc ccactaagcc aatcacttta
15721 ttgrctccta rccgcagacc tcctcrttct aacctgaatc ggaggrcaac cagtaagcta
15781 cccytttacc atyattggac aartarcatc crtactatac ttcrcaacaa tcytaatcct
15841 aataccaayt atctccctaa ttgaaaacaa aatactcaaa tggscctgtc cttgtagtay
15901 aaaytartac accagtcttg taarccrrar aygaaaacyt yyttccaagg acaaatcaga
15961 gaaaaagyct ttaactccac cattagcacc caaagctaag attctaattt aaactaytct
16021 ctgttctttc atggggargc agatttgggt rccacccaag tattgactya yccaycaaca
16081 accgcyatgt atytcgtaca ttactgcyag ycamcatgaa tatygyacvg taccataaay
16141 actyrayyac ctrtagtaca trmaamyyya ryccryatca ammyyyyvyc cyyatgctta
16201 caagcargya crryaaycra ccyycarcyr yyayrcatya ryygyarcyc caamryyrcy
16261 yctymycyay yagratayca acarasyyay yyrycytyaa cagyacatrg yacatrwwry
16321 catyyrycgt acatagcaca ttryagtcaa atcyyyycty gycccyaygg atgacccccc
16381 tcagataggr rtcccttgrc caccatcctc cgtgaaatca atatcccgca caagagtrmt
16441 actctcctcg ctccgggccc ataacacttg gggtagcta aartgaactg tatccgacat
16501 ctggttccta cttcagggyc ataaagycta aatagcccac acgttcccct taaataagac
16561 atcacgatg
```

SEQ ID NO:2

Human Mitochondrial DNA Revised Cambridge Reference Sequence

Note: the sequence below is a modified version of the 2001 Revised Cambridge Reference Sequence (GenBank #NC_001807, a derivation of #J01415). Corrections have been made and annotated per the reanalysis by Andrews et al (1999). Rare polymorphisms are shown with flanking green nucleotides. Corrected sequencing errors are shown with flanking red nucleotides. Clicking on the corrected nucleotides will return to the annotation. A summary table of the Andrews et al revisions is available online.

This sequence differs from GenBank #NC_001807 in that:

HISTORICAL NUCLEOTIDE NUMBERS ARE MAINTAINED. 3106del is maintained in the sequence as a gap.
Rare polymorphism 750A is retained, as in both the original Anderson et al (1981) paper & in the sequence reanalysis by Andrews et al.
This Revised Cambridge Reference Sequence (RCRS) has eighteen annotated nucleotides.

Seven are rare polymorphisms: Nucleotides 263A, 311C-315C, 750A, 1438A, 4769A, 8860A, and 15326A are considered to be rare polymorphisms. See summary table. Rare polymorphisms are shown with flanking green nucleotides.
Eleven are error corrections: Nucleotides 3106del, 3423T, 4985A, 9559C, 11335C, 13702C, 14199T, 14272C, 14365C, 14368C, 14766C are corrections of the original Cambridge sequence. The errors in the original Cambridge sequence have been attributed to sequencing errors (8 instances) and to the inclusion of bovine (2 instances) or HeLa (1 instance) DNA. See summary table.
Corrected sequencing errors are shown with flanking red nucleotides.

```
   1 gatcacaggt ctatcaccct attaaccact cacgggagct ctccatgcat ttggtatttt
  61 cgtctggggg gtatgcacgc gatagcattg cgagacgctg gagccggagc accctatgtc
 121 gcagtatctg tctttgattc ctgcctcatc ctattattta tcgcacctac gttcaatatt
 181 acaggcgaac atacttacta aagtgtgtta attaattaat gcttgtagga cataataata
 241 acaattgaat gtctgcacag ccactttcca cacagacatc ataacaaaaa atttccacca
 301 aaccccccct cccccgcttc tggccacagc acttaaacac atctctgcca aaccccaaaa
 361 acaaagaacc ctaacaccag cctaaccaga tttcaaattt tatcttttgg cggtatgcac
 421 ttttaacagt caccccccaa ctaacacatt attttcccct cccactccca tactactaat
 481 ctcatcaata caacccccgc ccatcctacc cagcacacac acaccgctgc taaccccata
 541 ccccgaacca accaaacccc aaagacaccc cccacagttt atgtagctta cctcctcaaa
 601 gcaatacact gaaaatgttt agacgggctc acatcacccc ataaacaaat aggtttggtc
 661 ctagcctttc tattagctct tagtaagatt acacatgcaa gcatccccgt tccagtgagt
 721 tcaccctcta aatcaccacg atcaaaagga acaagcatca agcacgcagc aatgcagctc
 781 aaaacgctta gcctagccac acccccacgg gaaacagcag tgattaacct ttagcaataa
 841 acgaaagttt aactaagcta tactaacccc agggttggtc aatttcgtgc cagccaccgc
 901 ggtcacacga ttaacccaag tcaatagaag ccggcgtaaa gagtgtttta gatcaccccc
 961 tccccaataa agctaaaact cacctgagtt gtaaaaaact ccagttgaca caaaatagac
1021 tacgaaagtg gctttaacat atctgaacac acaatagcta agacccaaac tgggattaga
1081 taccccacta tgcttagccc taaacctcaa cagttaaatc aacaaaactg ctcgccagaa
1141 cactacgagc cacagcttaa aactcaaagg acctggcggt gcttcatatc cctctagagg
1201 agcctgttct gtaatcgata aaccccgatc aacctcacca cctcttgctc agcctatata
1261 ccgccatctt cagcaaaccc tgatgaaggc tacaaagtaa gcgcaagtac ccacgtaaag
1321 acgttaggtc aaggtgtagc ccatgaggtg gcaagaaatg ggctacattt tctaccccag
1381 aaaactacga tagcccttat gaaacttaag ggtcgaaggt ggatttagca gtaaactaag
1441 agtagagtgc ttagttgaac agggccctga agcgcgtaca caccgcccgt caccctcctc
1501 aagtatactt caaaggacat ttaactaaaa cccctacgca tttatataga ggagacaagt
1561 cgtaacatgg taagtgtact ggaaagtgca cttggacgaa ccagagtgta gcttaacaca
1621 aagcacccaa cttacactta ggagatttca acttaacttg accgctctga gctaaaccta
1681 gccccaaacc cactccacct tactaccaga caaccttagc caaaccattt acccaaataa
1741 agtataggcg atagaaattg aaacctggcg caatagatat agtaccgcaa gggaaagatg
1801 aaaaattata accaagcata atatagcaag gactaacccc tataccttct gcataatgaa
1861 ttaactagaa ataactttgc aaggagagcc aaagctaaga cccccgaaac cagacgagct
1921 acctaagaac agctaaaaga gcacacccgt ctatgtagca aaatagtggg aagatttata
1981 ggtagaggcg acaaacctac cgagcctggt gatagctggt tgtccaagat agaatcttag
2041 ttcaacttta aatttgccca cagaaccctc taaatcccct tgtaaattta actgttagtc
2101 caaagaggaa cagctctttg gacactagga aaaaaccttg tagagagagt aaaaaattta
2161 acacccatag taggcctaaa agcagccacc aattaagaaa gcgttcaagc tcaacaccca
2221 ctacctaaaa aatcccaaac atataactga actcctcaca cccaattgga ccaatctatc
2281 accctataga agaactaatg ttagtataag taacatgaaa acattctcct ccgcataagc
2341 ctgcgtcaga ttaaaacact gaactgacaa ttaacagccc aatatctaca atcaaccaac
```

2401 aagtcattat taccctcact gtcaacccaa cacaggcatg ctcataagga aaggttaaaa
2461 aaagtaaaag gaactcggca aatcttaccc cgcctgttta ccaaaaacat cacctctagc
2521 atcaccagta ttagaggcac cgcctgccca gtgacacatg tttaacggcc gcggtaccct
2581 aaccgtgcaa aggtagcata atcacttgtt ccttaaatag ggacctgtat gaatggctcc
2641 acgagggttc agctgtctct tacttttaac cagtgaaatt gacctgcccg tgaagaggcg
2701 ggcataacac agcaagacga gaagacccta tggagcttta atttattaat gcaaacagta
2761 cctaacaaac ccacaggtcc taaactacca aacctgcatt aaaaatttcg gttggggcga
2821 cctcggagca gaacccaacc tccgagcagt acatgctaag acttcaccag tcaaagcgaa
2881 ctactatact caattgatcc aataacttga ccaacggaac aagttaccct agggataaca
2941 gcgcaatcct attctagagt ccatatcaac aatagggttt acgacctcga tgttggatca
3001 ggacatcccg atggtgcagc cgctattaaa ggttcgtttg ttcaacgatt aaagtcctac
3061 gtgatctgag ttcagaccgg agtaatccag gtcggtttct atcta-cttc aaattcctcc
3121 ctgtacgaaa ggacaagaga aataaggcct acttcacaaa gcgccttccc ccgtaaatga
3181 tatcatctca acttagtatt atacccacac ccacccaaga acagggtttg ttaagatggc
3241 agagcccggt aatcgcataa aacttaaaac tttacagtca gaggttcaat tcctcttctt
3301 aacaacatac ccatggccaa cctcctactc ctcattgtac ccattctaat cgcaatggca
3361 ttcctaatgc ttaccgaacg aaaaattcta ggctatatac aactacgcaa aggccccaac
3421 gttgtaggcc cctacgggct actacaaccc ttcgctgacg ccataaaact cttcaccaaa
3481 gagcccctaa aacccgccac atctaccatc accctctaca tcaccgcccc gaccttagct
3541 ctcaccatcg ctcttctact atgaaccccc ctccccatac ccaacccct ggtcaacctc
3601 aacctaggcc tcctatttat tctagccacc tctagcctag ccgtttactc aatcctctga
3661 tcagggtgag catcaaactc aaactacgcc ctgatcggcg cactgcgagc agtagcccaa
3721 acaatctcat atgaagtcac cctagccatc attctactat caacattact aataagtggc
3781 tcctttaacc tctccaccct tatcacaaca caagaacacc tctgattact cctgccatca
3841 tgacccttgg ccataatatg atttatctcc acactagcag agaccaaccg aaccccttc
3901 gaccttgccg aaggggagtc cgaactagtc tcaggcttca acatcgaata cgccgcaggc
3961 cccttcgccc tattcttcat agccgaatac acaaacatta ttataataaa caccctcacc
4021 actacaatct tcctaggaac aacatatgac gcactctccc ctgaactcta cacaacatat
4081 tttgtcacca agaccctact tctaacctcc ctgttcttat gaattcgaac agcatacccc
4141 cgattccgct acgaccaact catacacctc ctatgaaaaa acttcctacc actcacccta
4201 gcattactta tatgatatgt ctccataccc attacaatct ccagcattcc ccctcaaacc
4261 taagaaatat gtctgataaa agagttactt tgatagagta aataatagga gcttaaaccc
4321 ccttatttct aggactatga gaatcgaacc catccctgag aatccaaaat tctccgtgcc
4381 acctatcaca ccccatccta aagtaaggtc agctaaataa gctatcgggc ccataccccg
4441 aaaatgttgg ttataccctt cccgtactaa ttaatcccct ggcccaaccc gtcatctact
4501 ctaccatctt tgcaggcaca ctcatcacag cgctaagctc gcactgattt tttacctgag
4561 taggcctaga aataaacatg ctagctttta ttccagttct aaccaaaaaa ataaaccctc
4621 gttccacaga agctgccatc aagtatttcc tcacgcaagc aaccgcatcc ataatccttc
4681 taatagctat cctcttcaac aatatactct ccggacaatg aaccataacc aatactacca
4741 atcaatactc atcattaata atcataatag ctatagcaat aaaactagga atagcccct
4801 ttcacttctg agtcccagag gttacccaag gcaccctct gacatccggc ctgcttcttc
4861 tcacatgaca aaaactagcc cccatctcaa tcatatacca aatctctccc tcactaaacg
4921 taagccttct cctcactctc tcaatcttat ccatcatagc aggcagttga ggtggattaa
4981 accaAaccca gctacgcaaa atcttagcat actcctcaat tacccacata ggatgaataa
5041 tagcagttct accgtacaac cctaacataa ccattcttaa tttaactatt tatattatcc
5101 taactactac cgcattccta ctactcaact taaactccag caccacgacc ctactactat
5161 ctcgcacctg aaacaagcta acatgactaa caccettaat tccatccacc ctcctctccc
5221 taggaggcct gcccccgcta accggctttt tgcccaaatg ggccattatc gaagaattca
5281 caaaaaacaa tagcctcatc atccccacca tcatagccac catcaccctc cttaacctct
5341 acttctacct acgcctaatc tactccacct caatcacact actccccata tctaacaacg
5401 taaaaataaa atgacagttt gaacatacaa aacccacccc attcctcccc acactcatcg
5461 cccttaccac gctactccta cctatctccc cttttatact aataatctta tagaaattta
5521 ggttaaatac agaccaagag ccttcaaagc cctcagtaag ttgcaatact taatttctgt
5581 aacagctaag gactgcaaaa ccccactctg catcaactga acgcaaatca gccactttaa
5641 ttaagctaag cccttactag accaatggga cttaaaccca caaacactta gttaacagct
5701 aagcacccta atcaactggc ttcaatctac ttctcccgcc gccgggaaaa aaggcgggag
5761 aagccccggc aggtttgaag ctgcttcttc gaatttgcaa ttcaatatga aaatcacctc
5821 ggagctggta aaaagaggcc taacccctgt ctttagattt acagtccaat gcttcactca
5881 gccattttac ctcaccccca ctgatgttcg ccgaccgttg actattctct acaaaccaca

5941 aagacattgg aacactatac ctattattcg gcgcatgagc tggagtccta ggcacagctc
6001 taagcctcct tattcgagcc gagctgggcc agccaggcaa ccttctaggt aacgaccaca
6061 tctacaacgt tatcgtcaca gcccatgcat ttgtaataat cttcttcata gtaataccca
6121 tcataatcgg aggctttggc aactgactag ttcccctaat aatcggtgcc cccgatatgg
6181 cgtttccccg cataaacaac ataagcttct gactcttacc tccctctctc ctactcctgc
6241 tcgcatctgc tatagtggag gccggagcag gaacaggttg aacagtctac cctcccttag
6301 cagggaacta ctcccaccct ggagcctccg tagacctaac catcttctcc ttacacctag
6361 caggtgtctc ctctatctta ggggccatca atttcatcac aacaattatc aatataaaac
6421 cccctgccat aacccaatac caaacgcccc tcttcgtctg atccgtccta atcacagcag
6481 tcctacttct cctatctctc ccagtcctag ctgctggcat cactatacta ctaacagacc
6541 gcaacctcaa caccaccttc ttcgaccccg ccggaggagg agaccccatt ctataccaac
6601 acctattctg atttttcggt caccctgaag tttatattct tatcctacca ggcttcggaa
6661 taatctccca tattgtaact tactactccg gaaaaaaaga accatttgga tacataggta
6721 tggtctgagc tatgatatca attggcttcc tagggtttat cgtgtgagca caccatatat
6781 ttacagtagg aatagacgta gacacacgag catatttcac ctccgctacc ataatcatcg
6841 ctatccccac cggcgtcaaa gtatttagct gactcgccac actccacgga agcaatatga
6901 aatgatctgc tgcagtgctc tgagccctag gattcatctt tcttttcacc gtaggtggcc
6961 tgactggcat tgtattagca aactcatcac tagacatcgt actacacgac acgtactacg
7021 ttgtagccca cttccactat gtcctatcaa taggagctgt atttgccatc ataggaggct
7081 tcattcactg atttccccta ttctcaggct acaccctaga ccaaacctac gccaaaatcc
7141 atttcactat catattcatc ggcgtaaatc taactttctt cccacaacac tttctcggcc
7201 tatccggaat gccccgacgt tactcggact accccgatgc atacaccaca tgaaacatcc
7261 tatcatctgt aggctcattc atttctctaa cagcagtaat attaataatt ttcatgattt
7321 gagaagcctt cgcttcgaag cgaaaagtcc taatagtaga agaaccctcc ataaacctgg
7381 agtgactata tggatgcccc ccaccctacc acacattcga agaacccgta tacataaaat
7441 ctagacaaaa aaggaaggaa tcgaaccccc caaagctggt ttcaagccaa ccccatggcc
7501 tccatgactt tttcaaaaag gtattagaaa aaccatttca taactttgtc aaagttaaat
7561 tataggctaa atcctatata tcttaatggc acatgcagcg caagtaggtc tacaagacgc
7621 tacttcccct atcatagaag agcttatcac ctttcatgat cacgccctca taatcatttt
7681 ccttatctgc ttcctagtcc tgtatgccct tttcctaaca ctcacaacaa aactaactaa
7741 tactaacatc tcagacgctc aggaaataga aaccgtctga actatcctgc ccgccatcat
7801 cctagtcctc atcgccctcc catccctacg catcctttac ataacagacg aggtcaacga
7861 tccctccctt accatcaaat caattggcca ccaatggtac tgaacctacg agtacaccga
7921 ctacggcgga ctaatcttca actcctacat acttccccca ttattcctag aaccaggcga
7981 cctgcgactc cttgacgttg acaatcgagt agtactcccg attgaagccc ccattcgtat
8041 aataattaca tcacaagacg tcttgcactc atgagctgtc cccacattag gcttaaaaac
8101 agatgcaatt cccggacgtc taaaccaaac cactttcacc gctacacgac cggggtata
8161 ctacggtcaa tgctctgaaa tctgtggagc aaaccacagt ttcatgccca tcgtcctaga
8221 attaattccc ctaaaatct ttgaaatagg gcccgtattt accctatagc accccctcta
8281 ccccctctag agcccactgt aaagctaact tagcattaac cttttaagtt aaagattaag
8341 agaaccaaca cctctttaca gtgaaatgcc ccaactaaat actaccgtat ggcccaccat
8401 aattaccccc atactcctta cactattcct catcacccaa ctaaaaatat taaacacaaa
8461 ctaccaccta cctccctcac caaagcccat aaaaataaaa aattataaca aaccctgaga
8521 accaaaatga acgaaaatct gttcgcttca ttcattgccc ccacaatcct aggcctaccc
8581 gccgcagtac tgatcattct atttccccct ctattgatcc ccacctccaa atatctcatc
8641 aacaaccgac taatcaccac ccaacaatga ctaatcaaac taacctcaaa acaaatgata
8701 accatacaca acactaaagg acgaacctga tctcttatac tagtatcctt aatcattttt
8761 attgccacaa ctaacctcct cggactcctg cctcactcat ttacaccaac cacccaacta
8821 tctataaacc tagccatggc catcccctta tgagcgggcA cagtgattat aggctttcgc
8881 tctaagatta aaaatgccct agcccacttc ttaccacaag gcacacctac accccttatc
8941 cccatactag ttattatcga aaccatcagc ctactcattc aaccaatagc cctggccgta
9001 cgcctaaccg ctaacattac tgcaggccac ctactcatgc acctaattgg aagcgccacc
9061 ctagcaatat caaccattaa ccttccctct acacttatca tcttcacaat tctaattcta
9121 ctgactatcc tagaaatcgc tgtcgcctta atccaagcct acgttttcac acttctagta
9181 agcctctacc tgcacgacaa cacataatga cccaccaatc acatgcctat catatagtaa
9241 aacccagccc atgaccccta acaggggccc tctcagccct cctaatgacc tccggcctag
9301 ccatgtgatt tcacttccac tccataacgc tcctcatact aggcctacta accaacacac
9361 taaccatata ccaatgatgg cgcgatgtaa cacgagaaag cacataccaa ggccaccaca
9421 caccacctgt ccaaaaaggc cttcgatacg ggataatcct atttattacc tcagaagttt

9481 ttttcttcgc aggatttttc tgagcctttt accactccag cctagcccct acccccccaat
9541 taggagggca ctggccccCa acaggcatca ccccgctaaa tcccctagaa gtcccactcc
9601 taaacacatc cgtattactc gcatcaggag tatcaatcac ctgagctcac catagtctaa
9661 tagaaaacaa ccgaaaccaa ataattcaag cactgcttat tacaatttta ctgggtctct
9721 attttaccct cctacaagcc tcagagtact tcgagtctcc cttcaccatt tccgacggca
9781 tctacggctc aacatttttt gtagccacag gcttccacgg acttcacgtc attattggct
9841 caactttcct cactatctgc ttcatccgcc aactaatatt tcactttaca tccaaacatc
9901 actttggctt cgaagccgcc gcctgatact ggcattttgt agatgtggtt tgactatttc
9961 tgtatgtctc catctattga tgagggtctt actcttttag tataaatagt accgttaact
10021 tccaattaac tagttttgac aacattcaaa aaagagtaat aaacttcgcc ttaattttaa
10081 taatcaacac cctcctagcc ttactactaa taattattac attttgacta ccacaactca
10141 acggctacat agaaaaatcc acccccttacg agtgcggctt cgaccctata tcccccgccc
10201 gcgtcccttt ctccataaaa ttcttcttag tagctattac cttcttatta tttgatctag
10261 aaattgccct ccttttaccc ctaccatgag ccctacaaac aactaacctg ccactaatag
10321 ttatgtcatc cctcttatta atcatcatcc tagccctaag tctggcctat gagtgactac
10381 aaaaaggatt agactgaacc gaattggtat atagtttaaa caaaacgaat gatttcgact
10441 cattaaatta tgataatcat atttaccaaa tgcccctcat ttacataaat attatactag
10501 catttaccat ctcacttcta ggaatactag tatatcgctc acacctcata tcctccctac
10561 tatgcctaga aggaataata ctatcgctgt tcattatagc tactctcata accctcaaca
10621 cccactccct cttagccaat attgtgccta ttgccatact agtctttgcc gcctgcgaag
10681 cagcggtggg cctagcccta ctagtctcaa tctccaacac atatggccta gactacgtac
10741 ataacctaaa cctactccaa tgctaaaact aatcgtccca acaattatat tactaccact
10801 gacatgactt tccaaaaaac acataatttg aatcaacaca accacccaca gcctaattat
10861 tagcatcatc cctctactat tttttaacca aatcaacaac aacctattta gctgttcccc
10921 aacctttttcc tccgaccccc taacaacccc cctcctaata ctaactacct gactcctacc
10981 cctcacaatc atggcaagcc aacgccactt atccagtgaa ccactatcac gaaaaaaact
11041 ctacctctct atactaatct ccctacaaat ctccttaatt ataacattca cagccacaga
11101 actaatcata ttttatatct tcttcgaaac cacacttatc cccaccttgg ctatcatcac
11161 ccgatgaggc aaccagccag aacgcctgaa cgcaggcaca tacttcctat tctacaccct
11221 agtaggctcc cttcccctac tcatcgcact aatttacact cacaacaccc taggctcact
11281 aaacattcta ctactcactc tcactgccca agaactatca aactcctgag ccaacaactt
11341 aatatgacta gcttacacaa tagcttttat agtaaagata cctctttacg gactccactt
11401 atgactccct aaagcccatg tcgaagcccc catcgctggg tcaatagtac ttgccgcagt
11461 actcttaaaa ctaggcggct atggtataat acgcctcaca ctcattctca acccccctgac
11521 aaaacacata gcctacccct tccttgtact atccctatga ggcataatta taacaagctc
11581 catctgccta cgacaaacag acctaaaatc gctcattgca tactcttcaa tcagccacat
11641 agccctcgta gtaacagcca ttctcatcca aacccccctga agcttcaccg gcgcagtcat
11701 tctcataatc gcccacgggc ttacatcctc attactattc tgcctagcaa actcaaacta
11761 cgaacgcact cacagtcgca tcataatcct ctctcaagga cttcaaactc tactcccact
11821 aatagctttt tgatgacttc tagcaagcct cgctaacctc gccttacccc ccactattaa
11881 cctactggga gaactctctg tgctagtaac cacgttctcc tgatcaaata tcactctcct
11941 acttacagga ctcaacatac tagtcacagc cctatactcc ctctacatat ttaccacaac
12001 acaatggggc tcactcaccc accacattaa caacataaaa ccctcattca cacgagaaaa
12061 caccctcatg ttcatacacc tatcccccat tctcctccta tccctcaacc ccgacatcat
12121 taccgggttt tcctcttgta aatatagttt aaccaaaaca tcagattgtg aatctgacaa
12181 cagaggctta cgacccctta tttaccgaga aagctcacaa gaactgctaa ctcatgcccc
12241 catgtctaac aacatggctt tctcaacttt taaaggataa cagctatcca ttggtcttag
12301 gccccaaaaa ttttggtgca actccaaata aaagtaataa ccatgcacac tactataacc
12361 accctaaccc tgacttccct aattcccccc atccttacca ccctcgttaa ccctaacaaa
12421 aaaaactcat acccccatta tgtaaaatcc attgtcgcat ccacctttat tatcagtctc
12481 ttccccacaa caatattcat gtgcctagac caagaagtta ttatctcgaa ctgacactga
12541 gccacaaccc aaacaaccca gctctcccta agcttcaaac tagactactt ctccataata
12601 ttcatccctg tagcattgtt cgttacatgg tccatcatag aattctcact gtgatatata
12661 aactcagacc caaacattaa tcagttcttc aaatatctac tcatcttcct aattaccata
12721 ctaatcttag ttaccgctaa caacctattc caactgttca tcggctgaga gggcgtagga
12781 attatatcct tcttgctcat cagttgatga tacgcccgag cagatgccaa cacagcagcc
12841 attcaagcaa tcctatacaa ccgtatcggc gatatcggtt tcatcctcgc cttagcatga
12901 tttatcctac actccaactc atgagaccca caacaaatag cccttctaaa cgctaatcca
12961 agcctcaccc cactactagg cctcctccta gcagcagcag gcaaatcagc ccaattaggt

13021 ctccacccct gactcccctc agccatagaa ggccccaccc cagtctcagc cctactccac
13081 tcaagcacta tagttgtagc aggaatcttc ttactcatcc gcttccaccc cctagcagaa
13141 aatagcccac taatccaaac tctaacacta tgcttaggcg ctatcaccac tctgttcgca
13201 gcagtctgcg cccttacaca aaatgacatc aaaaaaatcg tagccttctc cacttcaagt
13261 caactaggac tcataatagt tacaatcggc atcaaccaac cacacctagc attcctgcac
13321 atctgtaccc acgccttctt caaagccata ctatttatgt gctccgggtc catcatccac
13381 aaccttaaca atgaacaaga tattcgaaaa ataggaggac tactcaaaac catacctctc
13441 acttcaacct ccctcaccat tggcagccta gcattagcag gaatcccttt cctcacaggt
13501 ttctactcca aagaccacat catcgaaacc gcaaacatat catacacaaa cgcctgagcc
13561 ctatctatta ctctcatcgc tacctccctg acaagcgcct atagcactcg aataattctt
13621 ctcaccctaa caggtcaacc tcgcttcccc acccttacta acattaacga aaataacccc
13681 accctactaa accccattaa acgcctggca gccggaagcc tattcgcagg atttctcatt
13741 actaacaaca tttcccccgc atcccccttc caaacaacaa tccccctcta cctaaaactc
13801 acagccctcg ctgtcacttt cctaggactt ctaacagccc tagacctcaa ctacctaacc
13861 aacaaactta aaataaaatc cccactatgc acattttatt tctccaacat actcggattc
13921 taccctagca tcacacaccg cacaatcccc tatctaggcc ttcttacgag ccaaaacctg
13981 cccctactcc tcctagacct aacctgacta gaaaagctat tacctaaaac aatttcacag
14041 caccaaatct ccacctccat catcacctca acccaaaaag gcataattaa actttacttc
14101 ctctctttct tcttcccact catcctaacc ctactcctaa tcacataacc tattcccccg
14161 agcaatctca attacaatat atacaccaac aaacaatgtt caaccagtaa ctactactaa
14221 tcaacgccca taatcataca aagcccccgc accaatagga tcctcccgaa tCaaccctga
14281 cccctctcct tcataaatta ttcagcttcc tacactatta aagtttacca caaccaccac
14341 cccatcatac tctttcaccc acagcaccaa tcctacctcc atcgctaacc ccactaaaac
14401 actcaccaag acctcaaccc ctgaccccca tgcctcagga tactcctcaa tagccatcgc
14461 tgtagtatat ccaaagacaa ccatcattcc ccctaaataa attaaaaaaa ctattaaacc
14521 catataacct cccccaaaat tcagaataat aacacacccg accacaccgc taacaatcaa
14581 tactaaaccc ccataaatag gagaaggctt agaagaaaac cccacaaacc ccattactaa
14641 acccacactc aacagaaaca aagcatacat cattattctc gcacggacta caaccacgac
14701 caatgatatg aaaaaccatc gttgtatttc aactacaaga acaccaatga ccccaatacg
14761 caaaactaac cccctaataa aattaattaa ccactcattc atcgacctcc ccaccccatc
14821 caacatctcc gcatgatgaa acttcggctc actccttggc gcctgcctga tcctccaaat
14881 caccacagga ctattcctag ccatgcacta ctcaccagac gcctcaaccg ccttttcatc
14941 aatcgcccac atcactcgag acgtaaatta tggctgaatc atccgctacc ttcacgccaa
15001 tggcgcctca atattcttta tctgcctctt cctacacatc gggcgaggcc tatattacgg
15061 atcatttctc tactcagaaa cctgaaacat cggcattatc ctcctgcttg caactatagc
15121 aacagccttc ataggctatg tcctcccgtg aggccaaata tcattctgag gggccacagt
15181 aattacaaac ttactatccg ccatcccata cattgggaca gacctagttc aatgaatctg
15241 aggaggctac tcagtagaca gtcccaccct cacacgattc tttacctttc acttcatctt
15301 gcccttcatt attgcagccc tagcaacact ccacctccta ttcttgcacg aaacgggatc
15361 aaacaacccc ctaggaatca cctcccattc cgataaaatc accttccacc cttactacac
15421 aatcaaagac gccctcggct tacttctctt ccttctctcc ttaatgacat taacactatt
15481 ctcaccagac ctcctaggcg acccagacaa ttataccctа gccaacccct taaacacccc
15541 tccccacatc aagcccgaat gatatttcct attcgcctac acaattctcc gatccgtccc
15601 taacaaacta ggaggcgtcc ttgccctatt actatccatc ctcatcctag caataatccc
15661 catcctccat atatccaaac aacaaagcat aatatttcgc ccactaagcc aatcacttta
15721 ttgactccta gccgcagacc tcctcattct aacctgaatc ggaggacaac cagtaagcta
15781 cccttttacc atcattggac aagtagcatc cgtactatac ttcacaacaa tcctaatcct
15841 aataccaact atctccctaa ttgaaaacaa aatactcaaa tgggcctgtc cttgtagtat
15901 aaactaatac accagtcttg taaaccggag atgaaaacct tttccaagg acaaatcaga
15961 gaaaaagtct ttaactccac cattagcacc caaagctaag attctaattt aaactattct
16021 ctgttctttc atggggaagc agatttgggt accacccaag tattgactca cccatcaaca
16081 accgctatgt atttcgtaca ttactgccag ccaccatgaa tattgtacgg taccataaat
16141 acttgaccac ctgtagtaca taaaaaccca atccacatca aaaccccctc cccatgctta
16201 caagcaagta cagcaatcaa ccctcaacta tcacacatca actgcaactc caaagccacc
16261 cctcacccac taggatacca acaaacctac ccacccttaa cagtacatag tacataaagc
16321 catttaccgt acatagcaca ttacagtcaa atcccttctc gtccccatgg atgacccccc
16381 tcagataggg gtcccttgac caccatcctc cgtgaaatca atatcccgca caagagtgct
16441 actctcctcg ctccgggccc ataacacttg gggtagcta aagtgaactg tatccgacat
16501 ctggttccta cttcagggtc ataaagccta aatagcccac acgttcccct taaataagac
16561 atcacgatg

【図１】

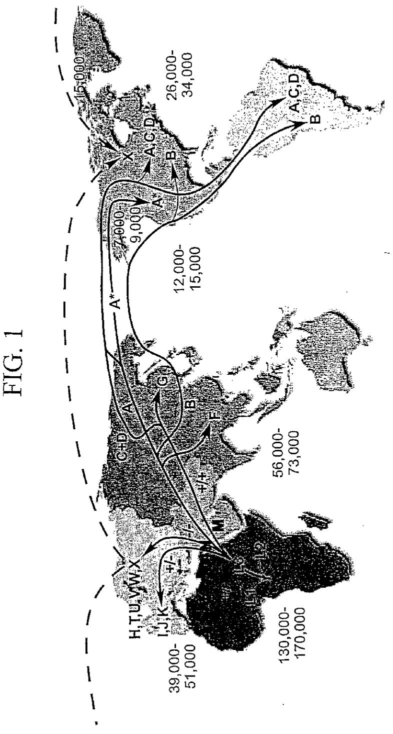


FIG. 1

【図２Ａ】

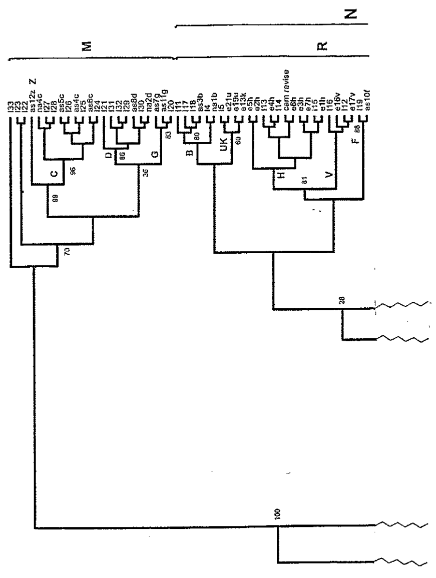


FIG. 2A

【図２Ｂ】

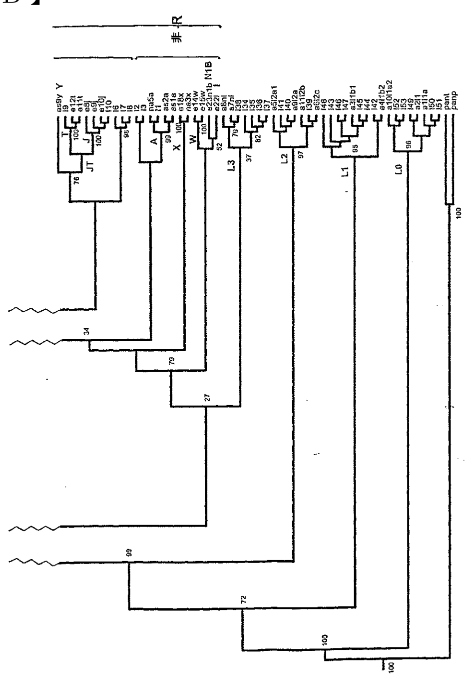


FIG. 2B

【図３】

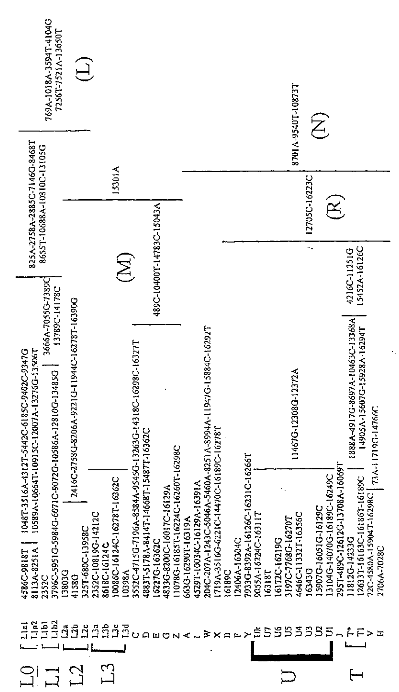
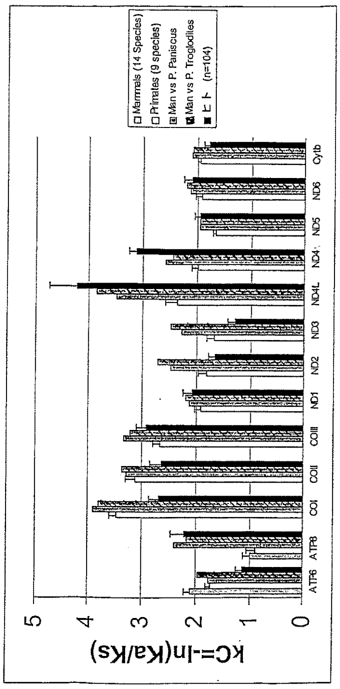


FIG. 3

【図４】

FIG. 4


kC=-ln(Ka/Ks)
0
1
2
3
4
5
ATP6
ATP8
COI
COII
COIII
ND1
ND2
ND3
ND4L
ND4
ND5
ND6
Cytb
Mammals (14 Species)
Primates (9 species)
Man vs P. Paniscus
Man vs P. Troglodites
ヒト (n=104)

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US02/28471

A. CLASSIFICATION OF SUBJECT MATTER
IPC(7) : C12Q 1/68; C07H 21/02; 21/04
US CL : 435/6; 536/23.1, 23.5, 24.1, 24.3, 24.31, 24.33
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 435/6; 536/23.1, 23.5, 24.1, 24.3, 24.31, 24.33

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Please See Continuation Sheet

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	ANDREWS et al. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nature Genetics. October 1999, Vol. 23, page 147, see complete document.	1, 2, 23, and 63-81
Y	NCBI Database, National Center for Biotechnology Information, National Library of Medicine, NIH (Bethesda, MD, USA) Accession No. J01415, 18 April 2000.	1, 2, 23 and 63-81
Y	US 5,837,832 A (CHEE et al) 17 November 1998 (17.11.1998), especially columns 5 and 6.	63-81

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
06 March 2003 (06.03.2003)	18 APR 2003

Name and mailing address of the ISA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	Valerie Bell-Harris for Carla Myers
Facsimile No. (703)305-3230	Telephone No. 703-308-0196

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US02/28471

Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

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

1. ☐ Claim Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claim Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claim Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:
Please See Continuation Sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☒ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.: Please See Continuation Sheet

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest ☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

PCT/US02/28471

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Prior to setting forth the inventions, it is noted that the groupings have been modified in view of the comments submitted in the response of February 12, 2003. The term "allele" has been deleted and replaced with the phrase "nucleotide locus" in order to clarify any misunderstandings regarding the invention to be searched. Both allelic variants at each nucleotide locus will be searched for each elected invention, i.e., 64C and 64T. Furthermore, groups 1-958 and 983-1940 include each of the haplotypes set forth in the present disclosure. As stated on page 97 of the disclosure, "SEQ ID NO: 1 is a theoretical human mtDNA genome sequence containing the nucleotide alleles of this invention as listed in table 3." Accordingly, in the following groupings, table 3 has been referenced as listing each of the nucleotide alleles and nucleotide loci. The disclosure makes clear that a single nucleotide allele may be detected as indicative of a haplotype and thereby the claims are not limited to methods in which a specific set of allelic variants are to be detected. Rather, the claims include detecting a single allele set forth in Table 3. In view of Applicants statements in the response of February 12, 2003, the invention to be searched as "the first invention" is considered to be the nucleotide locus of 825 and in particular, the alleles 825T and 825A.

Groups 1-958, claims 1-23, drawn to methods for haplotyping comprising determining whether a sample from a mitochondrial nucleic acid of a human contains one of the alleles as set forth in Table 3. Upon election of one of the groups, please specify the "nucleotide locus" to be searched (i.e., one of the nucleotide positions listed in Table 3, for example, nucleotide position 64 with respect to SEQ ID NO: 1).

Group 959, claims 24, 25, and 53-58, drawn to a machine readable storage device comprising one or more alleles.

Groups 960-971, claims 26-38, drawn to a method for identifying a gene. It is noted that Groups 960-971 correspond to the genes ND1, ND2, ND3, ND4, ND5, ND6, cytb, COI, COII, COIII, ATP6, and ATP8, respectively. For example, if Group 960 is elected, then the claims will be searched to the extent that they are limited to methods of for identifying the ND1 gene. Upon election of one these groups, please specific the gene to be identified.

Group 972, claims 39-52, drawn to a method for diagnosing a physiological predisposition by comparing haplotypes from different geographic locations.

Groups 973-979, claims 59-61, drawn to a method for diagnosing LHON by detecting a nucleotide allele. It is noted that Groups 973-979 correspond to the alleles at positions 458, 295, 12612, 13708, 16069, 3635 and 4640, respectively. For example, if Group 973 is elected, then the claims will be searched to the extent that they apply are limited to method of for diagnosing LHON by detecting the 458 allele. Upon election of one these groups, please specific the gene to be identified.

Groups 980-982, claim 62, drawn to a method for diagnosing the likelihood of blindness. It is noted that Groups 980-982, correspond to alleles 11778, 14484 and 10663, respectively. For example, if Group 980 is elected, then the claims will be searched to the extent that they apply are limited to method of for identifying the 1778 allele. Upon election of one these groups, please specific the allele to be detected.

Groups 983-1940, claims 63-81, drawn to libraries, arrays, and methods of using arrays comprising a nucleic acid of 7 to 30 nucleotides comprising one of the alleles set forth in Table 3. Upon election of one of the groups, please specify the number of the "nucleotide locus" to be searched (i.e., one of the nucleotide positions listed in Table 3, for example, nucleotide position 64 with respect to SEQ ID NO: 1).

The inventions listed as Groups 1-1940 do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

The first claimed invention, claim 1, lacks unity because the invention represents methods which do not provide a special technical feature over the art. Groups 1-958 of claims 1-23 are broadly drawn to methods for haplotyping by detecting any allele of the

INTERNATIONAL SEARCH REPORT

PCT/US02/28471

human mtDNA. As set forth in the disclosure, methods for haplotying human mtDNA are well known in the art and specifically are taught by Wallace et al, Ingman et al., Maca-Meyer et al., Schurr et al, and Macaulay et al (see page 1of the disclosure). Since the mtDNA of humans is known in the art, as were methods of haplotyping mtDNA, as admitted in the specification, there is no special technical feature linking the recited groups, as would be necessary to fulfill the requirement for unity of invention.

It is also noted that each of the claims of groups 1-958, 960-971, 973-979, 980-982 and 983-1940 has been presented in improper Markush format, as distinct products and distinct methods are improperly joined in the claims. Each allele site and each molecule containing said allele site is structurally and functionally distinct from each other and has a different special technical feature than each other. The chemical structure of each polymorphism and of each molecule containing the same differ from each other. For example, a polynucleotide comprising the allele at position 64 is chemically, structurally, and functionally different from a molecule comprising the allele at position 72. As the products and methods encompassed by the claims do not share a special technical feature, the distinct products and methods may not properly be presented in the alternative. Accordingly, the claims have been separated into a number of groups corresponding to the number of different inventions encompassed by the claims, and the claims will be searched only as they read upon the invention of the elected group. For the same reasons, the remainder of the claims have been separated in a number of groups corresponding to the number of different inventions encompassed thereby.

Further, the groups comprising the storage device and arrays are drawn to distinct products lacking the same or corresponding special technical features. The nucleic acids of the arrays are composed of nucleotides and function in, e.g., methods of nucleic acid hybridization. The storage devices are composed of, e.g., a CPU, a display device, an input device, etc. and function in, e.g., methods of electronic sequence comparison. As products of different sets of the groups differ from each other in structure, function, and effect, they do not belong to a recognized class of chemical compound, or have both a "common property or activity" and a common structure as would be required to show that the inventions are "of a similar nature".

Additionally, each of the claimed methods have different objectives and require different process steps. The methods of Groups 1-958 require steps of determining a nucleotide sequence and identifying haplotypes. The claims of Groups 960-971 are drawn to methods for identifying a gene. The claims of Group 972 are drawn to a method to diagnose a predisposition to a physiological disease by comparing haplotypes from different geographical locations. The claims of groups 973-979 are drawn to methods for diagnosing LHON by detecting an allele. The claims of groups 980-982 are drawn to methods for diagnosing the likelihood of blindness by detecting the presence of an allele. In addition to differences in objectives, effects, and method steps, it is again noted that the claims of the present Groups are not directed to the detection or identification of molecules having the same or common special technical feature, for the reasons discussed above.

Continuation of Box II Item 3:
claims 1, 2 and 23 with respect to the allele at nucleotide locus 825; and claims 63-81 with respect to allele 12373G

Continuation of B. FIELDS SEARCHED Item 3:
DIALOG: MEDLINE, BIOSIS, EMBASE, SCISEARCH; STN: CA, GENBANK; WEST: US, EP, JP, WO Patents
search terms: mitochondrial, MTND5, ND5, NADH dehydrogenase 5, array, hybridization, polymorphism, mutation

フロントページの続き

(51)Int.Cl.[7]　　　　　　　　　　ＦＩ　　　　　　　　　　　　　テーマコード（参考）
　　　　　　　　　　　　　　　　　Ｃ１２Ｎ　15/00　　　　Ｆ

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

(72)発明者　ホッセイニ，　セイド
　　　　　アメリカ合衆国　ジョージア　３００９７，　ドゥルス，　アンバーレイ　ウェイ　７３２０
(72)発明者　ミシュマー，　ダン
　　　　　アメリカ合衆国　カリフォルニア　９２６１２，　アーバイン，　スタンフォード　コート　５６３
(72)発明者　ルイーズーペシーニ，　エドゥアルド
　　　　　アメリカ合衆国　カリフォルニア　９２６１２，　アーバイン，　パークビュー　レーン　３７００，　アパートメント　２０シー
(72)発明者　ロット，　マリー
　　　　　アメリカ合衆国　ジョージア　３０３４５，　アトランタ，　レンジウッド　テラス　２８３６

Ｆターム(参考)　4B024 AA11　CA01　HA11
　　　　　　　4B063 QA11　QA17　QQ42　QQ52　QR08　QR55　QR82　QS25　QS34

patsnap

专利名称(译)	人线粒体DNA多态性，单倍群，与生理状态的关联，以及基因分型阵列		
公开(公告)号	JP2005525082A	公开(公告)日	2005-08-25
申请号	JP2003523626	申请日	2002-08-30
[标]申请(专利权)人(译)	埃默里大学		
申请(专利权)人(译)	埃默里大学		
[标]发明人	ウォラスダグラスシー ホッセイニセイド ミシュマーダン ルイーズペシーニエドゥアルド ロットマリー		
发明人	ウォラス, ダグラス シー. ホッセイニ, セイド ミシュマー, ダン ルイーズ-ペシーニ, エドゥアルド ロット, マリー		
IPC分类号	G01N33/53 C12N15/09 C12Q1/68 G01N33/566 G01N37/00		
CPC分类号	C12Q1/6883 C12Q2600/156		
FI分类号	C12N15/00.ZNA.A C12Q1/68.Z G01N33/53.M G01N33/566 G01N37/00.102 C12N15/00.F		
F-TERM分类号	4B024/AA11 4B024/CA01 4B024/HA11 4B063/QA11 4B063/QA17 4B063/QQ42 4B063/QQ52 4B063/QR08 4B063/QR55 4B063/QR82 4B063/QS25 4B063/QS34		
代理人(译)	夏木森下		
优先权	60/316333 2001-08-30 US 2356536 2001-08-31 CA 60/380546 2002-05-13 US		
其他公开文献	JP2005525082A5		
外部链接	Espacenet		

摘要(译)

本发明提供了指示主要人单倍群的人mtDNA多态性和诊断这些单倍群和选择的单倍群的方法。本发明具体提供以下方法：诊断人单倍群的方法，包括步骤：a）提供包含所述人源线粒体核酸的样品;和b）鉴定指示样品中组的至少一个核苷酸等位基因的存在或不存在。

Gene	Physiological Condition	Nucleotide Locus	Cambridge Nucleotide Allele	Physiological Condition Nucleotide Allele	Cambridge Amino Acid Allele	Physiological Condition Amino Acid Allele
MTND1	*MELAS	3308	T	C	M	T
MTND1	*NIDDM; LHON; PEO	3316	G	A	A	T
MTND1	*LHON	3394	T	C	Y	H
MTND1	*NIDDM	3394	T	C	Y	H
MTND1	*ADPD	3397	A	G	M	V
MTND1	*LHON	3460	G	A	A	T
MTND1	*LHON	3496	G	T	A	S
MTND1	*LHON	3497	C	T	A	V
MTND1	*LHON	4136	A	G	Y	C
MTND1	*LHON	4160	T	C	L	P
MTND1	*LHON	4216	T	C	Y	H
MTND2	*LHON	4917	A	G	D	N
MTND2	*LHON	5244	G	A	G	S
MTND2	*AD	5460	G	A	A	T
MTND2	*AD	5460	G	T	A	S
MTCO1	*Myoglobinuria, Exercise Intolerance	5920	G	A	W	Ter
MTCO1	*Multisystem Disorder	6930	G	A	G	Ter
MTCO1	*LHON	7444	G	A	Ter	K
MTCO2	*Mitochondrial Encephalomyopathy	7587	T	C	M	T
MTCO2	*MM	7671	T	A	M	K
MTCO2	*Multisystem Disorder	7896	G	A	W	Ter
MTCO2	*Lactic Acidosis	8042	AT	2 nt del (AT)	M	Ter
MTATP6	*NARP	8993	T	G	L	R
MTATP6	*NARP / Leigh Disease	8993	T	C	L	P
MTATP6	*LHON	9101	T	C	I	T
MTATP6	*FBSN / Leigh Disease	9176	T	C	L	P
MTATP6	*Leigh Disease	9176	T	G	L	R
MTCO3	*LHON	9438	G	A	G	S
MTCO3	*Leigh-like	9537	C	C ins	Q	frameshift
MTCO3	*LHON	9738	G	T	A	S
MTCO3	*LHON	9804	G	A	A	T
MTCO3	*Mitochondrial Encephalopathy	9952	G	A	W	Ter
MTCO3	*PEM; MELAS	9957	T	C	F	L
MTND3	*ESOC	10191	T	C	S	P
MTND4	*MELAS	11084	A	G	T	A
MTND4	*LHON	11778	G	A	R	H
MTND4	*Exercise Intolerance	11832	G	A	W	Ter
MTND4	*DM	12026	A	G	I	V
MTND5	*MELAS	13513	G	A	D	N