

(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2005-502336  
(P2005-502336A)

(43) 公表日 平成17年1月27日(2005.1.27)

(51) Int. Cl. <sup>7</sup>	F I	テーマコード (参考)
C 1 2 N 15/00	C 1 2 N 15/00 Z N A A	4 B O 2 4
A 6 1 K 35/76	A 6 1 K 35/76	4 B O 6 3
A 6 1 K 48/00	A 6 1 K 48/00	4 B O 6 5
A 6 1 P 3/00	A 6 1 P 3/00	4 C O 8 4
A 6 1 P 43/00	A 6 1 P 43/00 1 O 5	4 C O 8 7
	審査請求 未請求 予備審査請求 有 (全 119 頁) 最終頁に続く	

(21) 出願番号 特願2003-519472 (P2003-519472)	(71) 出願人 504051445 ナショナル パブリック ヘルス インス ティテュート フィンランド エフイーエン-00300 ヘルシンキ マンネルヘイミンティエ 166
(86) (22) 出願日 平成14年8月9日 (2002.8.9)	(74) 代理人 110000109 特許業務法人特許事務所サイクス
(85) 翻訳文提出日 平成16年2月9日 (2004.2.9)	(72) 発明者 ベルトーネン レーナ フィンランド 00251 ヘルシンキ ハールトマニンカテュ 8 ナショナル パブリック ヘルス インスティテュート バイオメディカム デパートメント オ ブ モレキュラー メディシン
(86) 国際出願番号 PCT/EP2002/008963	
(87) 国際公開番号 W02003/014342	
(87) 国際公開日 平成15年2月20日 (2003.2.20)	
(31) 優先権主張番号 01119377.8	
(32) 優先日 平成13年8月10日 (2001.8.10)	
(33) 優先権主張国 欧州特許庁 (EP)	
(31) 優先権主張番号 01119528.6	
(32) 優先日 平成13年8月14日 (2001.8.14)	
(33) 優先権主張国 欧州特許庁 (EP)	
(31) 優先権主張番号 60/315,955	
(32) 優先日 平成13年8月31日 (2001.8.31)	
(33) 優先権主張国 米国 (US)	

最終頁に続く

(54) 【発明の名称】 成人型ラクターゼ欠乏症と関連するDNA変異体の同定

(57) 【要約】

本発明は、成人型ラクターゼ欠乏症の一因となる、または成人型ラクターゼ欠乏症を示す、腸内ラクターゼ・フロリジン加水分解酵素 (LPH) 遺伝子の5'部分を含む核酸分子であって、(a) 配列番号1の核酸配列を有する、または含む核酸分子であって、配列番号1のその配列が、図4にも図示され、図8に図示される配列中にも含まれる核酸分子；(b) 配列番号2の核酸配列を有する、または含む核酸分子であって、配列番号2のその配列が、図5にも図示され、図9に図示される配列中にも含まれる核酸分子；(c) その相補鎖が(a)または(b)の核酸分子と厳密な条件下にてハイブリダイズする、少なくとも20ヌクレオチドの核酸分子であって、前記ポリヌクレオチド/核酸分子が、LPH遺伝子から-13910 5'位に相当する位置にシトシン残基を有する核酸分子；(d) その相補鎖が(a)または(b)の核酸分子と厳密な条件下にてハイブリダイズする、少なくとも20ヌクレオチドの核酸分子であって、前記ポリヌクレオチド/核酸分子が、LPH遺伝子から-22018 5'位に相当する位置にグアニン残基を有する核酸分子；からなる群から選択される、核酸分子に関する。本発明はさらに、上述の核酸分子に含まれるSNPの解析に基づく、成人型ラクターゼ欠乏症の存在または素因について試験する方法に関する。さらに、本発明は、成人型ラクターゼ欠乏症の存在または素因の検出に有用な診断組成物およびキットに関する。

## 【特許請求の範囲】

## 【請求項 1】

成人型ラクターゼ欠乏症の一因となる、または成人型ラクターゼ欠乏症を示す、腸内ラクターゼ・フロリジン加水分解酵素 (LPH) 遺伝子の 5' 部分を含む核酸分子であって、  
(a) 配列番号 1 の核酸配列を有する、または含む核酸分子であって、配列番号 1 のその配列が、図 4 にも図示され、図 8 に図示される配列中にも含まれる核酸分子；

(b) 配列番号 2 の核酸配列を有する、または含む核酸分子であって、配列番号 2 のその配列が、図 5 にも図示され、図 9 に図示される配列中にも含まれる核酸分子；

(c) その相補鎖が (a) または (b) の核酸分子と厳密な条件下にてハイブリダイズする、少なくとも 20 ヌクレオチドの核酸分子であって、前記ポリヌクレオチド / 核酸分子が、LPH 遺伝子から - 13910 5' 位に相当する位置にシトシン残基を有する核酸分子；

(d) その相補鎖が (a) または (b) の核酸分子と厳密な条件下にてハイブリダイズする、少なくとも 20 ヌクレオチドの核酸分子であって、前記ポリヌクレオチド / 核酸分子が、LPH 遺伝子から - 22018 5' 位に相当する位置にグアニン残基を有する核酸分子；からなる群から選択される、核酸分子。

10

## 【請求項 2】

腸内ラクターゼ・フロリジン加水分解酵素 (LPH) 遺伝子の 5' 部分を含む核酸分子であって、

(a) 配列番号 3 の核酸配列を有する、または含む核酸分子であって、配列番号 3 のその配列が、図 6 に図示される核酸分子；

(b) 配列番号 4 の核酸配列を有する、または含む核酸分子であって、配列番号 4 のその配列が、図 7 にも図示される核酸分子；

(c) その相補鎖が (a) または (b) の核酸分子と厳密な条件下にてハイブリダイズする核酸分子であって、前記ポリヌクレオチド / 核酸分子が、LPH 遺伝子から - 13910 位に相当する位置にチミジン残基を有する核酸分子；

(d) その相補鎖が (a) または (b) の核酸分子と厳密な条件下にてハイブリダイズする核酸分子であって、前記ポリヌクレオチド / 核酸分子が、LPH 遺伝子から - 22018 位に相当する位置にアデノシン残基を有する核酸分子；からなる群から選択される、核酸分子。

20

30

## 【請求項 3】

ゲノム DNA である、請求項 1 または 2 に記載の核酸分子。

## 【請求項 4】

前記ゲノム DNA が遺伝子の一部である、請求項 3 に記載の核酸分子。

## 【請求項 5】

少なくとも 14 ヌクレオチドを有する、請求項 1 から 4 のいずれか一項に記載の核酸分子の断片であって、LPH 遺伝子のヌクレオチド位 - 13910 またはヌクレオチド位 - 22018 を含む、断片。

## 【請求項 6】

請求項 1 および 3 から 5 のいずれか一項に記載の核酸分子に相補的である、核酸分子。

40

## 【請求項 7】

請求項 2 から 5 のいずれか一項に記載の核酸分子に相補的である、核酸分子。

## 【請求項 8】

請求項 1 および 3 から 5 のいずれか一項に記載の核酸分子を含む、ベクター。

## 【請求項 9】

請求項 2 から 4 のいずれか一項に記載の核酸分子を含む、ベクター。

## 【請求項 10】

LPH 遺伝子のヌクレオチド位 - 13910 またはヌクレオチド位 - 22018 を含む、請求項 1 および 3 から 5 のいずれか一項に記載の核酸分子と、またはその相補鎖と、厳密な条件下でハイブリダイズする、プライマーまたはプライマー対。

50

## 【請求項 1 1】

LPH 遺伝子のヌクレオチド位 - 1 3 9 1 0 またはヌクレオチド位 - 2 2 0 1 8 を含む、請求項 2 から 5 のいずれか一項に記載の核酸分子と、またはその相補鎖と、厳密な条件下でハイブリダイズする、プライマーまたはプライマー対。

## 【請求項 1 2】

請求項 6 に記載のベクターで形質転換された、非ヒト宿主。

## 【請求項 1 3】

請求項 7 に記載のベクターで形質転換された、非ヒト宿主。

## 【請求項 1 4】

細菌、酵母細胞、昆虫細胞、真菌細胞、哺乳動物細胞、植物細胞、遺伝子導入動物または遺伝子導入植物である、請求項 1 2 または 1 3 に記載の非ヒト宿主。 10

## 【請求項 1 5】

請求項 1 および 3 から 6 のいずれか一項に記載の野生型核酸分子とは特異的に結合するが、対応する野生型核酸分子とは特異的に結合しない、抗体またはアプタマーまたはファージ。

## 【請求項 1 6】

請求項 2 から 5 および 7 のいずれか一項に記載の野生型核酸分子とは特異的に結合するが、成人型ラクターゼ欠乏症の一因となる、または成人型ラクターゼ欠乏症を示す、対応する変異配列とは特異的に結合しない、抗体またはアプタマーまたはファージ。

## 【請求項 1 7】

請求項 2、3、4 に記載の野生型核酸分子または請求項 9 に記載のベクターを含む、医薬組成物。 20

## 【請求項 1 8】

請求項 1 から 7 のいずれか一項に記載の核酸分子、請求項 8 または 9 に記載のベクター、請求項 1 1 または 1 2 に記載のプライマーまたはプライマー対、および / または請求項 1 5 または 1 6 に記載の抗体アプタマーおよび / またはファージを含む、診断組成物。

## 【請求項 1 9】

成人型ラクターゼ欠乏症もしくはそれに関連する形質の存在または素因について試験する方法であって、ホモ接合またはヘテロ接合の状態における、請求項 1 および 3 から 6 のいずれか一項に記載の核酸分子の存在に関して、かかる素因を保有する将来の患者またはかかる素因を保有する疑いのあるヒトから採取された試料を試験することを含む方法。 30

## 【請求項 2 0】

成人型ラクターゼ欠乏症もしくはそれに関連する形質の存在または素因について試験する方法であって、ホモ接合またはヘテロ接合の状態における、請求項 2 から 5 および 7 のいずれか一項に記載の核酸分子の存在に関して、かかる素因を保有する将来の患者またはかかる素因を保有する疑いのあるヒトから採取された試料を試験することを含む方法。

## 【請求項 2 1】

前記試験が、成人型ラクターゼ欠乏症の一因となる、または成人型ラクターゼ欠乏症を示す核酸分子に相補的な請求項 6 に記載の相補的核酸分子、またはプローブとしての野生型配列に相補的な請求項 7 に記載の核酸分子を、厳密な条件下にて前記試料に含まれる核酸分子とハイブリダイズさせる段階と、前記ハイブリダイゼーションを検出する段階とを含む、請求項 1 9 または 2 0 に記載の方法。 40

## 【請求項 2 2】

前記ハイブリダイゼーションの産物を制限エンドヌクレアーゼで消化するか、または前記ハイブリダイゼーションの産物を制限エンドヌクレアーゼを用いた消化にかける段階と、前記消化の産物を分析する段階とをさらに含む、請求項 1 9 または 2 1 のいずれか一項に記載の方法。

## 【請求項 2 3】

前記プローブが検出可能に標識される、請求項 2 1 に記載の方法。

## 【請求項 2 4】

前記試験が、請求項 1 から 7 のいずれか一項に記載の核酸分子の少なくとも一部の核酸配列を決定する段階を含み、前記部分が、LPH 遺伝子のヌクレオチド位 - 13910 および / またはヌクレオチド位 - 22018 を含む、請求項 19 または 20 に記載の方法。

【請求項 25】

核酸配列の前記決定が、固相ミニシーケンシング (minisequencing) によって行われる、請求項 24 に記載の方法。

【請求項 26】

前記核酸配列を決定する前に、前記核酸分子の少なくとも前記一部を増幅する段階をさらに含む、請求項 24 に記載の方法。

【請求項 27】

前記試験が、増幅反応を行う段階であって、前記増幅反応に用いられるプライマーの少なくとも一つが、請求項 10 に記載のプライマーであるか、または請求項 10 に記載のプライマー対に属する段階と、増幅産物をアッセイする段階とを含む、請求項 19 または 20 に記載の方法。

【請求項 28】

前記試験が、増幅反応を行う段階であって、前記増幅反応に用いられるプライマーの少なくとも一つが、請求項 11 に記載のプライマーであるか、または請求項 11 に記載のプライマー対に属する段階と、増幅産物をアッセイする段階とを含む、請求項 19 または 20 に記載の方法。

【請求項 29】

前記増幅が、ポリメラーゼ連鎖反応 (PCR) によって行われるか、または前記増幅がポリメラーゼ連鎖反応 (PCR) である、請求項 26 から 28 のいずれか一項に記載の方法。

【請求項 30】

成人型ラクターゼ欠乏症の存在または素因について試験する方法であって、ヒトから採取された試料を、請求項 15 に記載の抗体またはアプタマーまたはファージへの特異的結合についてアッセイする段階を含む方法。

【請求項 31】

成人型ラクターゼ欠乏症の存在または素因について試験する方法であって、ヒトから採取された試料を、請求項 16 に記載の抗体またはアプタマーまたはファージへの特異的結合についてアッセイする段階を含む方法。

【請求項 32】

前記抗体またはアプタマーまたはファージが検出可能に標識される、請求項 30 または 31 に記載の方法。

【請求項 33】

前記試験がイムノアッセイである、請求項 30 から 32 のいずれか一項に記載の方法。

【請求項 34】

前記試料が、血液、血清、血漿、胎児組織、唾液、尿、粘膜組織、粘液、腔組織、腔から採取される胎児組織、皮膚、毛、毛嚢または他のヒト組織である、請求項 19 から 33 のいずれか一項に記載の方法。

【請求項 35】

前記試料からの前記核酸分子が、固体担体に固定化される、請求項 19 から 34 のいずれか一項に記載の方法

【請求項 36】

前記固体担体が、チップ、シリカウエハー、ビーズまたはマイクロタイタープレートである、請求項 35 に記載の方法。

【請求項 37】

成人型ラクターゼ欠乏症の存在または素因を解析するための、請求項 1 から 7 のいずれか一項に記載の核酸分子の使用。

【請求項 38】

10

20

30

40

50

請求項 1 から 7 のいずれか一項に記載の核酸分子、請求項 11 または 12 に記載のプライマーまたはプライマー対、請求項 8 または 9 に記載のベクター、および / または請求項 15 または 16 に記載の抗体アプタマーおよび / またはファージを 1 つまたは複数の容器内に含むキット。

【請求項 39】

遺伝子治療における、請求項 2 から 4 のいずれか一項に記載の核酸分子または請求項 7 に記載のベクターの使用。

【請求項 40】

前記遺伝子治療によって、成人型ラクターゼ欠乏症が治療または予防される、請求項 39 に記載の使用。

10

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、成人型ラクターゼ欠乏症 (adult-type hypolactasia) の一因となる、またはそれを示す、腸内ラクターゼ・フロリジン加水分解酵素 (LPH) 遺伝子の 5' 部分を含む核酸分子に関し、前記核酸分子は、(a) 配列番号 1 の核酸配列を有する、または含む核酸分子であって、配列番号 1 のその配列が、図 4 にも図示され、図 8 に図示される配列中にも含まれる核酸分子；(b) 配列番号 2 の核酸配列を有する、または含む核酸分子であって、配列番号 2 のその配列が、図 5 にも図示され、図 9 に図示される配列中にも含まれる核酸分子；(c) その相補鎖が (a) または (b) の核酸分子と、厳密な条件下にてハイブリダイズする、少なくとも 20 ヌクレオチドの核酸分子であって、前記ポリヌクレオチド / 核酸分子が、LPH 遺伝子から -13910 5' 位に相当する位置にシトシン残基を有する核酸分子；(d) その相補鎖が (a) または (b) の核酸分子と、厳密な条件下にてハイブリダイズする、少なくとも 20 ヌクレオチドの核酸分子であって、前記ポリヌクレオチド / 核酸分子が、LPH 遺伝子から -22018 5' 位に相当する位置にグアニン残基を有する核酸分子；からなる群から選択される。本発明はさらに、成人型ラクターゼ欠乏症の存在、またはその素因について試験する方法であって、上記の核酸分子に含まれる SNP の解析に基づく方法に関する。さらに、本発明は、成人型ラクターゼ欠乏症の存在、またはその素因の検出に有用な診断組成物および診断キットに関する。

20

【0002】

様々な試料が本明細書を通じて記載されている。製造元のマニュアルおよびカタログを含む、これらの試料の開示内容は、参照により本明細書に組み込まれる。

30

【0003】

腸内上皮細胞によって主に発現される、ラクターゼ・フロリジン加水分解酵素 (LPH) は、ラクトース、乳糖をグルコースおよびガラクトース<sup>1</sup>に加水分解する。ラクトースがもはや食事の必須成分ではない場合、LPH 酵素の発現は、哺乳動物において離乳期間に非常に低いレベルに劇的に低下する。ヒトにおいては、成人型ラクターゼ欠乏症またはラクターゼ非存続 (lactase non-persistence) として知られる状態は、ほとんどの人種に発生し、乳糖不耐症のために、成人間での新鮮な乳の使用を厳しく制限するものである。ラクターゼ非存続状態の発症年齢は、タイ人における 1 ~ 2 歳から、フィンランド人中の 10 ~ 20 歳までの範囲にわたり、人種間で異なる<sup>2 ~ 3</sup>。しかしながら、北欧人種および他の数種の人種群においては、LPH 活性は、成人の大部分で一生を通してラクターゼ存続として知られる状態を持続する。その表現型ラクターゼ存続 / 非存続は、遺伝的に決定されることが示されており、その存続状態は、非存続性状態<sup>4 ~ 6</sup>と比較して優性である。

40

【0004】

成人型ラクターゼ欠乏症の現況技術の診断は、乳糖負荷試験 (LTT) に基づく。一晚 (10 時間) 絶食させた後、ラクトース 1 g / kg を 12.5 % 溶液として投与する。最大投与量は 50 g である。ラクトース摂取前およびラクトースを摂取して 20 分および 30 分後に、毛細管血試料を採取する。そのグルコース濃度は、グルコースオキシターゼ法によって決定される (イェルム (Hjelm) およびデ・ベルディエ (de Verdier) 1963)。LT

50

Tの日の腹部症状を記録する。1. 1 mmol/l以上の血中グルコース濃度の最高上昇が、吸収不良の徴候とされる(ガッドマン(Gudman)-ホイヤー(Hoyer)およびハルマン(Harnum) 1968, ジュッシラ(Jussila) 1970, サヒ(Sahi) 1972)。LLTは、偽陽性および偽陰性診断のリスクを10%有し、つまり、LTTの感受性および特異性は約90%である(イソコスキー(Isokoski)ら. 1972, ニューカマー(Newcomer)ら. 1975, Sahi 1983)。肝臓におけるガラクトースの代謝を阻害するエタノール0.3g/kgを投与し(Tygstrup およびLundqvist 1962)、12.5%溶液でラクトース1g/kgを15分後に投与することによって、LLTの精度を高めることができる。

#### 【0005】

最初の、または繰り返し行われたLTTにおける0.2mg/100ml未満の最高上昇を有する小児は、胃内視鏡によって行われる小腸生検に送られる。これは、専門知識が必要とされる観血的手順であり、通常、胃腸病学の専門家によってのみ大学病院で行われている。生検試料は、解剖顕微鏡を用いて組織学的に検査され、粘膜のマルターゼ、スクラーゼおよびラクターゼ活性が決定される(ラウニアラ(Launiala)ら. 1964)。小児におけるラクターゼ欠乏症の診断は、腸内生検の組織構造が正常であり、ラクターゼ活性がタンパク質20U/g未満であり、ラクターゼ/スクラーゼ比0.30未満であれば、あるいは、エタノール投与を用いたLTTにおいては、20mg/100ml未満の血中グルコース濃度の最高上昇および5mg/100ml未満の血中ガラクトース濃度の最高上昇(サヒ(Sahi)ら, 1972)が示されれば、正しいとされる。上述のように、成人型ラクターゼ欠乏症を診断する現在の方法は手間がかかる。LTTは不正確であり、したがって、診断が確定される前に、観血的手順、胃鏡検査が必要である。成人型ラクターゼ欠乏症は一般的であり、非特異的な腹部症状(患者の3分の1が胃痛を訴える)の主な原因であることから、この一般的な健康上の問題の診断法を改善することが明らかに必要である。

#### 【0006】

それにもかかわらず、今までに取り扱い易く、それと同時に、迅速かつ正確な結果が得られる生化学的試験法は開発されていなかった。ゲノムDNA/発現レベルでの疾患原因の解明も同様に成功していない。このように、成人におけるLPH遺伝子のコード領域およびプロモーター領域のシーケンシングによって、ラクターゼ存続性/非存続性と関連するDNA変異は明らかにされておらず、この形質<sup>7-8</sup>と関連するスプライスバリエーションまたはmRNA編集バリエーションの証拠も現れていない。ラクターゼ存続性/非存続性の形質が、ラクターゼ遺伝子内にある、またはラクターゼ遺伝子に隣接するシス作用要素(1つまたは複数)によっておそらく制御されることが、以前の研究によって示されており、強い連鎖不平衡(LD)が、ラクターゼ遺伝子にわたる70kbハプロタイプ全体に認められている<sup>9,10</sup>。いくつかの研究によって、転写調節レベルでLPH遺伝子発現の主要な制御が働くという証明が報告されている<sup>11-13</sup>。しかしながら、LPH遺伝子の発現の転写および転写後制御の両方に影響を及ぼす変異が、成人型ラクターゼ欠乏症の病因に関係することが示唆されている<sup>14-15</sup>。

#### 【0007】

上記の事項を考慮すると、本発明の根底にある技術的課題は、成人型ラクターゼ欠乏症、またはこの疾患に対する素因の正確かつ簡便な診断を可能にする、手段および方法を提供することである。

#### 【0008】

前記技術的課題の解決は、特許請求の範囲において特徴付けられる実施形態によって達成される。

#### 【0009】

このように、本発明は、成人型ラクターゼ欠乏症の一因となる、またはそれを示す、腸内ラクターゼ・フロリジン加水分解酵素(LPH)遺伝子の5'部分を含む核酸分子に関するものであり、前記核酸分子は、(a)配列番号1の核酸配列を有する、または含む核酸分子であって、配列番号1の配列が、図4にも図示され、図8に図示される配列中にも含まれる核酸分子; (b)配列番号2の核酸配列を有する、または含む核酸分子であって、

配列番号2のその配列が、図5にも図示され、図9に図示される配列中にも含まれる核酸分子；(c)その相補鎖が(a)または(b)の核酸分子と、厳密な条件下にてハイブリダイズする、少なくとも20ヌクレオチドの核酸分子であって、前記ポリヌクレオチド/核酸分子が、LPH遺伝子から-139105'位に相当する位置にシトシン残基を有する核酸分子；(d)その相補鎖が(a)または(b)の核酸分子と、厳密な条件下にてハイブリダイズする、少なくとも20ヌクレオチドの核酸分子であって、前記ポリヌクレオチド/核酸分子が、LPH遺伝子から-220185'位に相当する位置にグアニン残基を有する核酸分子；からなる群から選択される。

#### 【0010】

本発明に従って、「腸内ラクターゼ・フロリジン加水分解酵素(LPH)遺伝子」という用語は、ラクトースを、その成分、グルコースおよびガラクトースに加水分解する活性を有する酵素をコードする遺伝子を意味する。その酵素は、E.C.3.2.1.23.62によって特徴付けられる。 10

#### 【0011】

「成人型ラクターゼ欠乏症」という用語は、全世界の人種にかなりの割合で、腸内細胞におけるラクターゼ・フロリジン加水分解酵素(LPH)の酵素活性の「生理学的」減少から起こる常染色体劣性症状であり、乳糖不耐症としても知られる症状を意味する。

#### 【0012】

「成人型ラクターゼ欠乏症の一因となる、またはそれを示す」という用語は、SNP、およびこのように見出された対応する核酸分子がその症状を指し示し、したがって、もしかすると原因にもなり得るという事実を意味する。したがって、この用語には、記載の5'位置が症状を指し示すということが必然的に必要とされる。一方、前記用語には、5'位置が症状の原因であるか、または一因であるということは必ずしも必要ではない。しかし、前記用語から、いずれかのSNPまたは両方のSNPの原因となる、または一因となる役割は除外されない。 20

#### 【0013】

「厳密な条件下でハイブリダイズする」という言葉は、当業者によく知られているが、または従来のプロトコールに従って当業者によって確立することができるハイブリダイゼーション条件を意味する。最も有利なことには、その言葉は、非常に厳密な条件を意味する。各配列に、適した、厳密な条件は、温度、核酸分子の組成、塩の条件等のよく知られているパラメーターに基づいて確立される：例えば、サンプルック(Sambrook)ら、"Molecular Cloning, A Laboratory Manual"; CSH Press, Cold Spring Harbor, 1989またはヒギンズ(Higgins)およびヘイムズ(Hames)(編集者)、"Nucleic acid hybridization, a practical approach", IRL Press, Oxford 1985(参考文献54)、特にブリテン(Britten)&デビッドソン(Davidson)による"Hybridization Strategy"の3~15章を参照のこと。通常の(非常に厳密な)条件は、0.5×SSCおよび0.1%SDS中で65でのハイブリダイゼーションまたは50%ホルムアミド、4×SSCおよび0.1%SDS中で42でのハイブリダイゼーションを含む。通常、ハイブリダイゼーション後には洗浄して、非特異的シグナルを除去する。洗浄条件としては、65、0.2×SSCおよび0.1%SDSまたは2×SSCおよび0.1%SDSまたは0.3×SSCおよび0.1%SDS(25~65にて)などの条件が挙げられる。 30 40

#### 【0014】

本明細書において上述されるように、本発明は、少なくとも20ヌクレオチドのハイブリダイズする(hybridizing)核酸分子に関する；上記の(c)および(d)を参照のこと。さらに、本発明は、少なくとも50、少なくとも100、少なくとも150、または少なくとも200ヌクレオチドの核酸分子にも関する。前記ハイブリダイズ(hybridizing)断片は、少なくとも25、少なくとも50、少なくとも75、または少なくとも100ヌクレオチド、上記の(c)に定義される-13910位の、または上記の(d)に定義される-22018位の5'および3'を含むことが好ましい。

#### 【0015】

「核酸分子」という用語は、天然および非天然核酸分子のどちらも意味する。非天然核酸分子には、cDNAならびにPNAなどの誘導体が含まれる。

【0016】

本明細書全体を通じて、「配列番号\_\_の核酸配列を含む核酸分子・・・」という言葉は、配列番号によって指定される核酸分子よりも少なくとも1ヌクレオチド長い核酸分子を意味する。それと同時に、これらの分子は、例えば配列番号2または1、3または4によって指定される本発明の核酸分子の5'および/または3'末端にわたり、最大で3000ヌクレオチドに及ぶ。

【0017】

驚くべきことに、本発明によれば、2つのラクターゼ欠乏症関連変異体が、MCM6遺伝子の異なるイントロン中に位置付けられる、LPH遺伝子からかなり離れて位置することが見出された。MCM6は、遺伝子ファミリー(MCM2~7)のメンバーであり、確実に細胞周期中に1回だけ起こるようにDNA複製の開始に必要とされる<sup>31</sup>。LPHと異なり、MCM6は、その組織分布において制限されず、MCM6およびLPH転写物のレベルに相関性はない<sup>18</sup>。これらの発見により、これら2種類の遺伝子は、組織特異性または発生制御を提供する、機能的に有意なシス作用要素を全く共有しないことが示唆されるだろう<sup>18</sup>。最もおそらくは、同定される変異体は、LPHおよびMCM6遺伝子の発現に対して異なる機能的な重要性を有する。さらに驚くべきことには、ラクターゼ欠乏症との完全な関連に基づいて、それら(またはそれらのうちの1種類)は、腸内上皮におけるLPH遺伝子の転写物レベルの年齢依存性ダウンレギュレーションに関連するが、MCM6の転写にはほとんど、または全く効果を及ぼさない。

【0018】

実験的に、9つの拡大されたフィンランド人家系において行われる、連鎖解析、対立遺伝子関連解析および拡大されたハプロタイプ解析を用いて、成人型ラクターゼ欠乏症の遺伝子座が、2q21上で47kb区間に限定された。その領域の配列解析によって、単一ヌクレオチド多型(SNP)、すべてのフィンランド人家系において、および異なる4つの集団(人種)からの236個体の試料セットにおいて、成人型ラクターゼ欠乏症で完全に同時分離されるC/T<sub>-13910</sub>が明らかとなった。C/T<sub>-13910</sub>から8kbのテロメア長にある他のSNP G/A<sub>-22018</sub>が7症例以外のすべてにおいて形質と関連した。1047DNAにおけるC/T<sub>-13910</sub> SNPの有病率は、3つの異なる集団における成人型ラクターゼ欠乏症の報告された有病率を反映し、形質のその重要性に更なる証拠を提供した。

【0019】

上記で言及されている、驚くべき発見によって、初めて、LPH遺伝子の上流にある記載の単一ヌクレオチド多型の分子解析に基づく試験システムを確立することが可能となる。どちらのSNPも、成人型ラクターゼ欠乏症の診断、または成人型ラクターゼ欠乏症の素因の診断に堅固な基礎を提供するが、ヌクレオチド位-13910が、単独で、またはヌクレオチド位-22018と組み合わせて解析されることが好ましい。これは、-13910位のSNPが、解析される症例の100%で疾患と関連したのに対して、-22018位のSNPは、すべての症例のうち98%のみで成人型ラクターゼ欠乏症と関連したからである。にもかかわらず、ヌクレオチド位-22018の単独での解析もまた通常、成人型ラクターゼ欠乏症に対する素因の診断に堅固な基盤を提供する。

【0020】

SNPの存在をスクリーニングする確立された方法が豊富にあるために、現在は、便利に、短時間で、低コストで、高精度で、調査中のヒトに著しく迷惑をかけることなく、成人型ラクターゼ欠乏症に対する遺伝的素因を診断することが可能である。

【0021】

本発明はさらに、腸内ラクターゼ・フロリジン加水分解酵素(LPH)遺伝子の5'部分を含む核酸分子に関し、前記核酸分子は、(a)配列番号3の核酸配列を有する、または含む核酸分子であって、配列番号3のその配列が図6に図示される、核酸分子; (b)配列番号4の核酸配列を有する、または含む核酸分子であって、配列番号4のその配列が図

7に図示される、核酸分子；(c)その相補鎖が(a)または(b)の核酸分子と、厳密な条件下にてハイブリダイズする核酸分子であって、前記ポリヌクレオチド/核酸分子が、LPH遺伝子の-13910位に相当する位置にチミジン残基を有する核酸分子；(d)その相補鎖が(a)または(b)の核酸分子と、厳密な条件下にてハイブリダイズする核酸分子であって、前記ポリヌクレオチド/核酸分子が、LPH遺伝子から-22018位に相当する位置にアデノシン残基を有する核酸分子；からなる群から選択される。

【0022】

本発明のこの実施形態は、ヒトが成人型ラクターゼ欠乏症に罹患していないこと、およびその素因を保持しないことを実証するために便利に用いることができる。さらに、LPH遺伝子の上流にある-13910位または-22018位の「野生型」の状況を反映するこの核酸分子は、成人型ラクターゼ欠乏症の素因を試験する実験において対照手段として使用してもよい。試験のために、本明細書を通じて述べられている方法を用いてもよい。

10

【0023】

本発明の好ましい実施形態において、核酸分子はゲノムDNAである。本発明のこの好ましい実施形態は、研究中のヒトから単離された体液、細胞または組織由来のゲノムDNAに基づいて解析が行われるだろうという事実を反映するものである。

【0024】

本発明の核酸分子のさらに好ましい実施形態において、前記ゲノムDNAは遺伝子の一部である。本発明に従って、LPH遺伝子に対して-13910位または-22018位を有するMCM6遺伝子のイントロンの少なくとも1つが解析されることが好ましい。

20

【0025】

さらに、本発明は、少なくとも14ヌクレオチドを有する、本明細書において上述の核酸分子の断片に関し、前記断片は、LPH遺伝子の(上流の)ヌクレオチド位-13910またはヌクレオチド位-22018を含む。

【0026】

本発明の断片は、天然由来ならびに(半)合成由来であってもよい。このように、断片は例えば、有機化学の従来のプロトコールに従って合成された核酸分子であってもよい。重要なことには、本発明の核酸断片は、LPH遺伝子の上流にあるヌクレオチド位-13910または-22018を含む。これらの位置において、断片は、野生型ヌクレオチド、あるいは成人型ラクターゼ欠乏症の一因となる、またはそれを示すヌクレオチド(「変異」配列とも呼ばれる)のいずれかを有し得る。したがって、本発明の断片は、たとえば野生型と変異配列とを区別するアッセイに使用してもよい。本発明の断片は、少なくとも17ヌクレオチド、さらに好ましくは少なくとも21ヌクレオチド、最も好ましくは少なくとも25ヌクレオチド、例えば30ヌクレオチドからなることがさらに好ましい。

30

【0027】

さらに、本発明は、本明細書において上述の核酸分子と相補的な核酸分子に関する。少なくとも14ヌクレオチドを含み、かつLPH遺伝子の上流にある配列の少なくとも-13910位または-22018位を網羅する、本発明のこの実施形態は、ハイブリダイゼーションアッセイにおける記載の位置の遺伝子構成(genetic setup)の解析に特に有用である。このように、例えば、野生型配列(つまり、-13910位のTまたは-22018位のA)、または成人型ラクターゼ欠乏症の一因となる、もしくはそれを示す変異(-13910位のCまたは-22018位のG)に正確に相補的な15merを使用して、多型変異間を区別することができる。これは、適切なハイブリダイゼーションおよび洗浄条件が選択されたとしても、解析される試料中のDNAに正確に相補的ではない、検出可能な標識で標識された核酸分子は、検出可能なシグナルを生じないだろうと考えられるためである。

40

【0028】

この点に関しては、本発明の核酸分子、その断片、ならびに相補的な核酸分子が検出可能に標識されることに留意することが重要である。検出可能な標識としては、<sup>3</sup>Hまたは<sup>32</sup>Pなどの放射性標識、または蛍光標識が挙げられる。核酸の標識は、当技術分野におい

50

て十分に理解されているが、例えばサンプルック (Sambrook) ら、上記引用文に記述されている。

#### 【0029】

さらに、本発明は、本明細書で上述される核酸分子を含むベクターに関する。本発明のベクターは、野生型配列 (1つまたは複数) を含む核酸分子を含有してもよいし、または変異配列 (1つまたは複数) を含む核酸分子を含有してもよい。そのベクターは特に、本発明の核酸分子を含む、遺伝子工学に従来から使用されているプラスミド、コスミド、ウイルスまたはバクテリオファージであってもよい。前記ベクターは、発現ベクターおよび/または遺伝子導入もしくはターゲティングベクターであることが好ましい。レトロウイルス、ワクシニアウイルス、アデノ随伴ウイルス、ヘルペスウイルス、またはウシ乳頭腫ウイルスなどのウイルスに由来する発現ベクターを、本発明の核酸分子を標的とされる細胞集団に送達するために使用してもよい。当業者によく知られている方法を用いて、組換えウイルスベクターを構築することができる；例えば、サンプルック (Sambrook) ら、上記引用文およびアウスベル (Ausubel) ら、Current Protocol in Molecular Biology, Green Publishing Associates and Wiley Interscience, N.Y. (1989) に述べられている技術を参照のこと。一方、本発明の核酸分子およびベクターは、標的細胞に送達するために、リポソームに再構成することができる。細胞宿主の種類に応じて異なる公知の方法によって、本発明の核酸分子を含有するベクターを宿主細胞に導入することができる。例えば、塩化カルシウムトランスフェクションは一般に、原核細胞に用いられているのに対して、例えば、リン酸カルシウムもしくはDEAE-デキストランによるトランスフェクションまたは電気穿孔法は、他の細胞宿主に使用される；上記のサンプルック (Sambrook) を参照のこと。

#### 【0030】

かかるベクターはさらに、適切な宿主細胞における、適切な条件下での前記ベクターの選択を可能にするマーカー遺伝子などの遺伝子を含み得る。本発明の核酸分子は、原核細胞または真核細胞における発現が可能となるように、発現制御配列に作動可能に連結されることが好ましい。前記ポリヌクレオチドの発現は、ポリヌクレオチドの翻訳可能なmRNAへの転写を含む。真核細胞、好ましくは哺乳動物細胞における発現を保證する調節因子は、当業者にはよく知られている。それらは通常、転写の開始を実現する調節配列、および任意に、転写の終結および転写物の安定化を保證するポリAシグナル、および/または前記ポリヌクレオチドの発現をさらに高めるイントロンを含む。その他の調節因子は、転写エンハンサーならびに翻訳エンハンサー、および/または天然に関連する、または異種のプロモーター領域を含み得る。原核生物の宿主細胞中での発現を可能にする、可能性のある調節因子は、例えば、大腸菌におけるPL、lac、trpまたはtacプロモーターを含み、真核生物の宿主細胞中での発現を可能にする調節因子の例としては、酵母におけるAOX1もしくはGAL1プロモーター、またはCMV-、SV40-、RSV-プロモーター (ラウス肉腫ウイルス)、CMV-エンハンサー、SV40-エンハンサーまたは哺乳動物および他の動物細胞におけるグロブリンイントロンが挙げられる。転写開始を担う因子に加えて、かかる調節因子は、ポリヌクレオチドの下流にあるSV40ポリA部位またはtkポリA部位などの転写終結シグナルも含み得る。任意に、異種配列は、所望の特性、例えば、発現された組換え産物の安定化または簡略化された精製を提供する、C末端もしくはC末端同定ペプチドペプチドを含む融合タンパク質をコードすることができる。このコンテキストにおいて、適切な発現ベクターは、岡山 (Okayama) -ベルグ (Berg) のcDNA発現ベクターpcDV1 (Pharmacia社)、pcDM8、pRc/CMV、pcDNA1、pcDNA3、Echo (登録商標) クローニングシステム (Invitrogen社)、pSPORT1 (GIBCO BRL社) またはpRevTet-On/pRevTet-OffまたはpCI (Promega社) などの当技術分野で公知のベクターである。好ましくは、発現制御配列は、真核生物の宿主細胞を形質転換またはトランスフェクトすることができるベクターにおける真核生物プロモーター系であるが、原核生物宿主の制御配列を使用してもよい。

10

20

30

40

50

## 【0031】

上述のように、本発明のベクターは、遺伝子導入またはターゲティングベクターであってもよい。ex - vivoまたはin - vivo技術による治療遺伝子の導入に基づく遺伝子治療は、遺伝子導入の最も重要な応用のうちの1つである。in - vitroまたはin - vivoでの遺伝子治療に用いられる適切なベクターおよび方法は、文献に記述されており、当業者には公知である；例えば、ジョルダーノ (Giordano), Nature Medicine 2 (1996), 534-539; Schaper, Circ. Res. 79 (1996), 911-919; アンダーソン (Anderson), Science 256 (1992), 808-813; Isner, ランセット (Lancet) 348 (1996), 370-374; ミュールハウゼン (Muhlhauser), Circ. Res. 77 (1995), 1077-1086; ワング (Wang), Nature Medicine 2 (1996), 714-716; W094/29469; W0 97/00957, シャーパー (Schaper), Current Opinion in Biotechnology 7 (1996), 635-640, またはケイ (Kay) ら (2001) Nature Medicine, 7, 33-40) およびそれらに記載の参考文献を参照のこと。本発明のポリヌクレオチドおよびベクターは、直接的に導入するために、またはリポソームもしくはウイルスベクター (例えば、アデノウイルス、レトロウイルス) を介して細胞に導入するために設計することができる。前記細胞は、生殖細胞系細胞、胚細胞、または卵細胞またはそれから誘導される細胞であることが好ましく、最も好ましくは、前記細胞は幹細胞である。遺伝子治療は、野生型核酸分子のみで考えられる。

10

## 【0032】

本発明は、プライマーまたはプライマー対にも関し、そのプライマーまたはプライマー対は、LPHのヌクレオチド位 - 13910または - 22018を含む上述の核酸と、またはその相補鎖と、(非常に) 厳密な条件下にてハイブリダイズする。本発明のプライマーは、17または21ヌクレオチドなど、少なくとも14ヌクレオチドの長さを有することが好ましい。そのプライマーは、24ヌクレオチドの最大長さを有することがさらに好ましい。伸長反応または増幅反応などの適切な検出方法と組み合わせられる、-13910位または-22018位のいずれかを含むゲノム配列との、適切な条件下でのプライマーのハイブリダイゼーションまたはハイブリダイゼーションの欠失を用いて、多型変異間を区別し、次いで、例えば成人型ラクターゼ欠乏症について研究中のヒトの素因に関して結論を導き出すことができる。本発明では、2種類のプライマー/プライマー対を取り上げる。一方の種類のプライマーは変異配列を含む配列とハイブリダイズする。言い換えると、そのプライマーは、-13910位にCまたは-22018位にGを含む配列に、またはその相補鎖に正確に相補的である。もう一方の種類のプライマーは、-13910位にTまたは-22018位にAを有する配列に、またはその相補鎖に正確に相補的である。ハイブリダイゼーション条件は好ましくは、十分厳密であるように選択されるであろうことから、例えば変異配列に正確に相補的なプライマーを野生型対立遺伝子と接触させても、結果としてミスマッチが形成するため、効率的なハイブリダイゼーションは行われない。洗浄後、プライマーが除去されるため、シグナルは検出されないだろう。

20

30

## 【0033】

さらに、本発明は、本明細書において上述される本発明のベクターで形質転換された非ヒト宿主に関する。その宿主は、変異配列または野生型配列を保持し得る。育種などでは、宿主は、一方または両方のSNPに対してヘテロ接合性またはホモ接合性であることができる。本発明の宿主は、ゲノムに一過的にまたは安定的に組み込まれた本発明のベクターを保持し得る。本発明の非ヒト宿主を作製する方法は当技術分野でよく知られている。例えば、サンプルック (Sambrook) ら, 上記引用文に記載の従来のトランスフェクションのプロトコルを用いて、形質転換細菌 (大腸菌など) または形質転換酵母を作製することができる。本発明の非ヒト宿主は、例えば成人型ラクターゼ欠乏症の発症を解明するために使用することができる。

40

## 【0034】

本発明の好ましい実施形態において、非ヒト宿主は、細菌、酵母細胞、昆虫細胞、真菌細胞、哺乳動物細胞、植物細胞、遺伝子導入動物または遺伝子導入植物である。大腸菌が好ましい細菌であるのに対して、好ましい酵母細胞は、S. cerevisiaeまたはPichia pastor

50

is細胞である。好ましい真菌細胞は、アスペルギルス属 (*Aspergillus*) 細胞であり、好ましい昆虫細胞としては、*Spodoptera frugiperda*細胞が挙げられる。好ましい哺乳動物細胞は、LPH酵素の発現を示す結腸癌細胞系であり、CaCo2細胞が挙げられる。

#### 【0035】

遺伝子導入非ヒト動物、例えば遺伝子導入マウスの作製方法は、生殖細胞、胚細胞、幹細胞または卵細胞またはそれから誘導される細胞に上述のポリヌクレオチドまたはターゲティングベクターを導入することを含む。非ヒト動物は、本明細書に記載の本発明のスクリーニング法に従って使用することができる。遺伝子導入胚の作製およびそれらのスクリーニングは、例えばA. L. ジョイナー (Joyner) Ed., *Gene Targeting, A Practical Approach* (1993), Oxford University Pressによって記述されているように行うことができる。胚の胚細胞のDNAは、例えば適切な相補的核酸分子でのサザンブロット法を用いて解析することができる；上記文献を参照のこと。遺伝子導入非ヒト動物を作製するための一般的な方法は、当技術分野において記述されており、例えばWO94/24274を参照のこと。遺伝子導入非ヒト生物（相同的に標的とされる非ヒト動物を含む）を作製するには、胚幹細胞 (ES細胞) が好ましい。本質的に述べられているように(ロバートソン (Robertson), E. J. (1987) in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E. J. Robertson, ed. (Oxford: IRL Press), p. 71-112)、有糸分裂上不活性なSNL76/7細胞のフィーダー層上で成長するAB-1細胞系(マクマホン (McMahon) およびブラッドリー (Bradley), *Cell* 62:1073-1085 (1990))などのマウスES細胞が、相同的遺伝子ターゲティングに用いられる。他の適切なES細胞系としては、限定されないが、E14細胞系 (フーパー (Hooper) ら, *Nature* 326:292-295 (1987))、D3細胞系(ドッチマン (Doetschman) ら, *J. Embryol. Exp. Morph.* 87:27-45 (1985))、CCE細胞系 (ロバートソン (Robertson) ら, *Nature* 323:445-448 (1986))、AK-7細胞系(チワン (Zhuang) ら, *Cell* 77:875-884 (1994))が挙げられる。特異的標的化変異を有するES細胞からのマウス系作製の成功は、ES細胞の多分化能（つまり、胚形成に参与し、得られた動物の胚細胞に寄与する胚盤胞または桑実胚など、胚を発生する宿主に注入された後のそれらの能力）に依存する。注入されたES細胞を含有する胚盤胞は、偽妊娠の非ヒトの雌の子宮内で発達させ、キメラマウスとして生まれる。目的の核酸分子を有する細胞のキメラである、得られた遺伝子導入マウスを戻し交配し、本発明の核酸分子にヘテロ接合性の遺伝子導入マウスを同定するために、子孫の尾(tail)の生検のDNAについてPCRまたはサザンブロット分析することによって、正しく標的化された導入遺伝子（1つまたは複数）の存在をスクリーニングする。

#### 【0036】

遺伝子導入非ヒト動物は、例えば遺伝子導入マウス、ラット、ハムスター、イヌ、サル（類人猿）、ウサギ、ブタ、または雌ウシである。前記遺伝子導入非ヒト動物はマウスであることが好ましい。本発明の遺伝子導入動物は、特に、本発明の核酸およびベクターの表現型発現/結果を研究するのに有用である。さらに、本発明の遺伝子導入動物は、例えばげっ歯類の腸におけるLPH酵素の発達上の発現を研究するのに有用である。本発明の非ヒト遺伝子導入動物を用いて、成人型ラクターゼ欠乏症を改善するのに有用な治療薬/組成物または可能性のある他の治療法について試験することができることがさらに考えられる。

#### 【0037】

さらに、本発明は、本発明の変異核酸分子に特異的に結合するが、対応する野生型核酸分子には特異的に結合しない、抗体またはアプタマーまたはファージに関する。その抗体を、結合について試験し、当技術分野でよく知られているいずれかの血清学的技術、例えばチューブ、ゲルにおける凝集技術、2次抗体を用いる、または用いない固相および捕捉技術において、あるいは免疫蛍光増強 (immunofluorescence enhancement) を用いる、または用いないフローサイトメトリーで使用してもよい(例えば、ハーロー (Harlow) およびレーン (Lane) "Antibodies, A Laboratory Manual", CSH Press, Cold Spring Harbor, USA, 1988 (参照番号53)に記載の技術を参照のこと)。

10

20

30

40

50

## 【0038】

本発明に従って、その抗体は、-13910位（その位のヌクレオチドがC）または-22018位（その位のヌクレオチドがG）を含むエピトープを特異的に認識する。それは、この位置にTを有する-13910位を含むエピトープとも、この位置にGを有する-22018位を含むエピトープとも交差反応しない、または本質的に交差反応しない。標準プロトコールに従って作製される抗体の特異性は、ELISAアッセイなどにおいて、野生型および変異配列を保持するDNA分子と接触させることによって試験することができる。変異配列を有するバックグラウンドではシグナルを生じるが、野生型配列を有するバックグラウンドではシグナルを生じない抗体のみが選択されるだろう。

## 【0039】

本発明の抗体は、ポリクローナル抗血清に由来する、またはポリクローナル抗血清に含まれるモノクローナル抗体もしくは抗体であってもよい。本発明によって使用される「抗体」という用語は、Fab、F(ab')<sub>2</sub>、FvもしくはscFv断片などの前記抗体の断片をさらに含む；例えば、ハーロー（Harlow）およびレーン（Lane）<sup>53</sup>，上記引用文を参照のこと。抗体またはその断片は天然由来であってもよいし、または（半）合成的に作製されたものであってもよい。かかる合成産物は、本発明の抗体と同じまたは本質的に同じ結合特異性を有する、半タンパク性（semi-proteinaceous）物質としての非タンパク性物質も含み得る。かかる産物は、例えば、ペプチドミメティックによって得ることができる。

## 【0040】

「アプタマー」という用語は、当技術分野でよく知られており、例えばオズボーン（Osborne）ら，Curr. Opin. Chem. Biol. 1（1997），5-9（参照番号51）またはストール（Stall）およびショーカ（Szoka），Pharm. Res. 12（1995），465-483（参照番号52）に定義されている。

## 【0041】

さらに、本発明は、本明細書において上述の野生型核酸分子に特異的に結合するが、成人型ラクターゼ欠乏症の一因となる、またはそれを示す、対応する変異配列には特異的に結合しない、抗体またはアプタマーまたはファージに関する。変異配列に特異的な抗体に関して述べられた、特異性等に関する記述には、本明細書において必要に応じて変更が加えられる。

## 【0042】

さらに、本発明は、本明細書において上述される野生型核酸分子を含む医薬組成物に関する。

## 【0043】

本発明の医薬組成物は、遺伝子治療アプローチ、特に体細胞遺伝子治療に用いることができる。上記で示され、かつ本発明の医薬組成物に含有される野生型核酸分子は、薬学的に許容される担体および/または希釈剤と組み合わせてもよい。適切な薬剤担体の例は当技術分野でよく知られており、リン酸緩衝液、水、油/水（O/W）型エマルジョンなどのエマルジョン、様々な種類の湿潤剤、滅菌溶液等が挙げられる。かかる担体を含む組成物は、よく知られている従来の方法によって調製することができる。これらの医薬組成物は、適切な用量で対象に投与することができる。適切な組成物の投与は、様々な方法によって、例えば静脈内、腹腔内、皮下、筋肉内、局所的、皮内、鼻腔内または気管支内投与によって行うことができる。投与計画は、治療を行う医師および臨床因子によって決定されるだろう。医療分野においてよく知られているように、いずれか1人の患者に対する投薬量は、患者のサイズ、体表面積、年齢、投与すべき特定の化合物、性別、時間および投与経路、全身の健康状態、および同時に投与される他の薬物などの多くの因子によって異なる。通常の投与量は例えば、発現のために、または発現の抑制のために、核酸0.001~1000μgの範囲であることができる；しかしながら、この例示的な範囲を下回るまたは超える投与量が、特に上記の因子を考慮して考えられる。投与量はさまざまであるだろうが、DNAの静脈内投与に好ましい投与量は、DNA分子約10<sup>6</sup>~10<sup>12</sup>コピーで

10

20

30

40

50

ある。定期的な診断によって、経過がモニターされる。本発明の組成物は、局所的に投与してもよいし、または全身的に投与してもよい。投与は一般に、非経口的、例えば静脈内投与であるだろう；例えば内部標的もしくは外部標的へのパーティクルガンによる (biolistic) 送達によって、または動脈における部位へのカテーテルによって、標的部位に DNA を直接投与することもできる。非経口投与用の製剤は、滅菌水溶液もしくは非水溶液、懸濁液、およびエマルジョンを含む。非水溶媒の例としては、プロピレングリコール、ポリエチレングリコール、オリーブ油などの植物油、およびオレイン酸エチルなどの注射可能な有機エステルが挙げられる。水性担体としては、水、アルコール/水溶液、エマルジョンまたは懸濁液、例えば生理食塩水および緩衝培地が挙げられる。非経口用の賦形剤としては、塩化ナトリウム溶液、リンゲルのブドウ糖 (Ringer's dextrose)、ブドウ糖および塩化ナトリウム、乳酸リンゲル液、または不揮発性油などが挙げられる。静脈内用の賦形剤としては、液体および栄養補充剤、電解質補充剤 (リンゲルのブドウ糖をベースとするものなど) 等が挙げられる。例えば、抗菌薬、抗酸化剤、キレート剤、および不活性ガス等の保存剤および他の添加剤もまた、存在してもよい。

10

20

30

40

50

#### 【0044】

さらに、本発明は、本明細書において上述される核酸分子、本明細書において上述されるベクター、本明細書において上述されるプライマーまたはプライマー対、および/または本明細書において上述される抗体アプタマーおよび/またはファージを含有する診断組成物に関する。その診断組成物は、成人型ラクターゼ欠乏症を発症する彼または彼女の素因に関して、またはその急性症状の診断に関して、ヒトの遺伝的体質を評価するのに有用である。診断組成物の可能な種々の成分は、溶媒中で、または凍結乾燥状態などの他の状態で1つまたは複数のバイアル中にパッケージされる。溶媒中に溶解される場合には、診断組成物は、少なくとも +8 ~ +4 に冷却することが好ましい。他の場合には、凍結することが好ましい。

#### 【0045】

本発明は、成人型ラクターゼ欠乏症もしくはそれに関連する形質の存在または素因について試験する方法であって、ホモ接合またはヘテロ接合の状態における、本明細書において上述される核酸分子の存在に対するかかる素因を保持する将来の患者またはかかる素因を保持する疑いのあるヒトから採取された試料を試験することを含む方法にも関する。様々な実施形態において、野生型配列 (1つまたは複数) または変異配列 (1つまたは複数) の存在について、それを試験することができる。

#### 【0046】

本発明の方法は、前記ヒト/患者の遺伝的構成 (genetic set-up) を検出し、前記患者の有する症状が成人型ラクターゼ欠乏症かどうかの適切な結論を導き出すのに有用である。代替方法として、症状を有していないヒトが、成人型ラクターゼ欠乏症の素因を保有するかどうかを評価することもできる。LPH 遺伝子の 5' 上流にある -13910 位に関しては、シトシンがホモ接合の状態で見出される場合にのみ、その症状は成人型ラクターゼ欠乏症と診断されるか、または対応する素因が明白であるだろう。一方、チミジンがホモ接合の状態で見出された場合、または個体がヘテロ接合性 (C/T) である場合、次いで、患者が患っている症状は、成人型ラクターゼ欠乏症に関連せず、さらに、その患者はこの症状を発症する素因を保有しないと結論づけられる。しかしながら、ヘテロ接合性遺伝子型を保有するヒトの子供は、C 残基を有する染色体がもう一方の親からの対応する染色体と合致する場合に、その症状を発症する可能性がある」と結論づけられる。

#### 【0047】

その状況は同様であり、本質的に同じ結論が、-22018 位における SNP の解析に当てはまる。ホモ接合的に存在する G 残基は、急性成人型ラクターゼ欠乏症の素因またはその発生を表す。ヘテロ接合の G/A 状態は、その症状を発症しない高い可能性と関連する。ホモ接合状態で A を保有する個体は、その症状を発症するとは予想されないだろう。同様に、ある症状を有する患者は、成人型ラクターゼ欠乏症を患っていないと診断されるだろう。

## 【0048】

本発明の方法の好ましい実施形態において、前記試験は、成人型ラクターゼ欠乏症の一因となる、またはそれを示す核酸分子と相補的な、本明細書において上述される相補的核酸分子、あるいはプローブとしての野生型配列に相補的な本明細書に上述される核酸分子を、前記試料に含まれる核酸分子と（非常に）厳密な条件下でハイブリダイズさせることと、前記ハイブリダイゼーションを検出することと、を含む。

## 【0049】

さらにまた、使用する核酸プローブに応じて、野生型または変異配列（つまり、成人型ラクターゼ欠乏症の一因となる、またはそれを示す）のいずれかが検出されるだろう。ハイブリダイゼーション条件は、野生型配列に相補的な核酸分子が変異配列とハイブリダイズしない、または本質的にハイブリダイズしないように選択されるであろうことは理解されよう。同様に、変異配列に相補的な核酸分子は、野生型配列とハイブリダイズしない、または本質的にハイブリダイズしないだろう。本発明のハイブリダイゼーション法において、ホモ接合性およびヘテロ接合性遺伝子型から得られた結果を区別するために、例えば、ハイブリダイゼーション後のそれぞれの検出シグナルの強さ/強度をモニター/検出することができる。本発明のハイブリダイゼーション法において、ホモ接合性、ヘテロ接合性野生型対立遺伝子および/またはホモ接合性変異対立遺伝子の間を区別するために、対応する遺伝子型の内部コントロール試料が分析に含まれるだろう。

## 【0050】

さらに好ましい実施形態において、本発明の方法はさらに、前記ハイブリダイゼーションの産物を制限エンドヌクラーゼで消化するか、または前記ハイブリダイゼーションの産物を、制限エンドヌクラーゼを用いた消化にかけることと、前記消化の産物を分析することを含む。本発明のこの好ましい実施形態によって、簡便な手段を用いて、有効なハイブリダイゼーションと有効でないハイブリダイゼーションを区別することが可能となる。例えば、-13910位または-22018位に隣接するDNA配列がエンドヌクラーゼ制限部位を含有する場合、ハイブリダイズ産物は、有効なハイブリダイゼーションでは、適切な制限酵素によって切断されるのに対して、ハイブリダイゼーションが不十分であると、二本鎖産物は得られないか、または認識可能な制限部位を含まず、したがって、切断されないだろう。特に、DNA変異体C/T<sub>-13910</sub>の配列に特異的な制限酵素は、CviJIであり、DNA変異体G/A<sub>-22018</sub>の配列に特異的な制限酵素は、HhaIおよびAciIである。rg/cyを切断する前記酵素は、Webcutterプログラムを使用することによって見出される。消化産物の解析は、例えば臭化エチジウムでの核酸の染色と任意に併用されるゲル電気泳動などの、従来の手段によって行うことができる。サザンプロット法などの更なる技術との組み合わせもまた考えられる。

## 【0051】

前記ハイブリダイゼーションの検出は、例えば、抗DNA二本鎖抗体によって、または標識されたオリゴヌクレオチドを用いることによって行うことができる。好都合なことに、本発明の方法は、サザンプロット法もしくはノーザンプロット法およびなどのプロット技術および関連する技術と共に用いられる。標識化は、例えば標準プロトコールによって行われ、放射性マーカー、蛍光標識、リン光標識、化学発光標識、酵素標識等での標識化が含まれる（上記もまた参照のこと）。

## 【0052】

上記に従って、本発明の方法の他の好ましい実施形態では、前記プローブは、例えば本明細書において上述される方法によって、および上述の標識で検出可能に標識される。

## 【0053】

本発明の方法のさらに他の好ましい方法において、前記試験は、本明細書において記述される核酸分子の少なくとも一部の核酸配列を決定することを含み、前記部分は、LPH遺伝子のヌクレオチド位-13910および/またはヌクレオチド位-22018を含む。核酸分子の決定は、サンガー（Sanger）法またはマクサム-ギルバート法（Maxam/Gilbert）のプロトコールなど、従来プロトコールのうちの1つに従って行うことができる（更

10

20

30

40

50

なる参考には、サンプルック (Sambrook) ら, 上記引用文を参照のこと)。

【0054】

本発明のさらに好ましい実施形態において、核酸の決定は、固相ミニシーケンシング (minisequencing) によって行われる。固相ミニシーケンシングは、溶液中の野生型および変異ヌクレオチドの定量分析に基づく。最初に、変異を含有するゲノム領域が、1つのピオチン化および非ピオチン化プライマーを用いてPCRによって増幅される。そのピオチン化プライマーは、ストレプトアビジン (SA) コーティングプレートに結合されている。PCR産物は一本鎖状に変性されて、変異部位のすぐ前にこの鎖にミスーシーケンシングプライマーが結合することが可能となる。非標識化dNTPと共に、トリチウム (H3) または蛍光標識された変異および野生型ヌクレオチドは、ミニシーケンシング反応に添加され、Taqポリメラーゼを用いてシーケンスされる。その結果は、計数器または蛍光光度計によって測定され、かつR比として表される、反応における野生型および変異ヌクレオチドの量に基づく。サイバネン (Syvanen) AC, サジャンティラ (Sajantila) A, ルッカ (Lukka) M. *Am J Hum Genet* 1993: 52, 46-59およびスオマラニエン (Suomalainen) A およびサイバネン (Syvanen) AC. *Methods Mol Biol* 1996; 65: 73-79を参照のこと。本発明の方法の好ましい実施形態はさらに、前記核酸配列を決定する前に、前記核酸分子の少なくとも前記一部を増幅することを含む。増幅は、ポリメラーゼ連鎖反応 (PCR) によって行われることが好ましい。リガーゼ連鎖反応などの他の増幅方法を用いることもできる。

10

【0055】

本発明の方法の好ましい実施形態では、前記試験は、増幅反応を行う段階であって、前記増幅反応で使用されるプライマーの少なくとも1つが、本明細書において上述されるプライマーであるか、本明細書において上述されるプライマー対に属する段階と、増幅産物についてアッセイする段階と、を含む。この実施形態において、かつ研究者/医師が得たいと望む情報に応じて、野生型または変異配列のいずれかとハイブリダイズするプライマーが用いられる。

20

【0056】

本発明の方法によって、前記標的配列が、ハイブリダイゼーションに用いられるプライマーに正確に相補的な配列を有するならば、標的配列のみが増幅されるだろう。これは、どの種類のプライマーが使用されるかに応じて、オリゴヌクレオチドプライマーが、好ましくは (非常に) 厳密なハイブリダイゼーション条件下にて、野生型/変異配列とハイブリダイズしないが (結果として、増幅産物は得られない)、正確に合致する配列のみにハイブリダイズするであろうためである。当然のことながら、両方のSNPとハイブリダイズするプライマー対の組み合わせを用いることもできる。この場合には、期待される増幅産物の解析 (2番目の非識別 (non-differentiating) プライマーが各遺伝子座に対して同じである場合、増幅産物はゼロ、1種、2種、3種または4種であり得る) から、-13910位および-22018位の両方の遺伝的状態についての情報が得られるだろう。

30

【0057】

本発明の方法の好ましい実施形態において、前記増幅は、ポリメラーゼ連鎖反応 (PCR) によって行われるか、または前記増幅は、ポリメラーゼ連鎖反応 (PCR) である。PCRは当技術分野で十分に確立されている。本発明に従って使用される通常の条件は、例えば、93 で3分間の変性段階; 55 で30秒間のアニーリング段階; 72 で75秒間の伸長段階; および72 で10分間の最終伸長段階; で例示される、体積合計50  $\mu$ lでの合計35サイクルを含む。

40

【0058】

本発明はさらに、成人型ラクターゼ欠乏症の存在または素因について試験する方法であって、本明細書において上述される抗体またはアプタマーまたはファージへの特異的結合に関して、ヒトから採取された試料をアッセイすることを含む方法に関する。このコンテキストにおいて、ホモ接合性の野生型対照試料 (2つの存続性対立遺伝子を含む) と比較して、本発明の抗原の存在に対して弱い染色が、ヘテロ接合性野生型 (一方は存続性対立遺

50

伝子であり、もう一方はラクターゼ欠乏 (hypolactasic) の対立遺伝子) に示されるのに対して、ホモ接合のラクターゼ欠乏 (hypolactasic) の個体には、適切な抗体が使用されたとしても、染色は期待されない。本発明の方法は、内部コントロールとして、可能性のある3つすべての対立遺伝子の組み合わせに対応する対照試料の存在下にて行われることが好ましい。試験は、野生型配列に特異的な、または変異配列に特異的な抗体等を用いて行うことができる。さらにまた、結合についての試験は、ELISAなどの標準技術の使用を含み得る；例えば、ハーロー (Harlow) およびレーン (Lane)<sup>53</sup>、上記引用文を参照のこと。

**【0059】**

本発明の方法の好ましい実施形態において、前記抗体またはアプタマーまたはファージは検出可能に標識される。そのアプタマーは好ましくは、<sup>3</sup>Hもしくは<sup>32</sup>Pで放射標識されるか、または上述の蛍光マーカで標識されるのに対して、ファージまたは抗体は、それに対応する手法において (好ましい放射性標識として<sup>131</sup>Iで) 標識するか、またはHisタグ、FLAGタグ、またはmycタグなどのタグで標識することができる。

10

**【0060】**

本発明の方法のさらに好ましい実施形態において、その試験は免疫アッセイである。

**【0061】**

本発明の方法の他の好ましい実施形態において、前記試料は、血液、血清、血漿、胎児組織、唾液、尿、粘膜組織、粘液、腔組織、腔から採取された胎児組織、皮膚、毛、毛嚢または他のヒト組織である。

20

**【0062】**

本発明のその他の好ましい実施形態では、前記試料からの前記核酸分子は、固体担体に固定化される。

**【0063】**

固体担体に核酸分子を固定化することによって、試験アッセイの扱いが容易になり、さらに、チップ、シリカウエハーまたはマイクロタイタープレートなどの少なくとも一部の固体担体によって、大量の試料を同時に分析することができる。理想的には、その固体担体によって、例えばロボット装置を用いた自動化試験が可能となる。

**【0064】**

本発明の特に好ましい実施形態において、前記固体担体は、チップ、シリカウエハー、ビーズまたはマイクロタイタープレートである。

30

**【0065】**

さらに、本発明は、成人型ラクターゼ欠乏症の存在または素因の解析のための、本明細書において上述される核酸分子の使用に関する。その核酸分子によって同時に、本明細書において詳細に述べられているように、その症状が存在しないこと、またはその症状の素因を解析することが可能となる。

**【0066】**

さらに、本発明は、本明細書において上述される核酸分子、本明細書において上述されるプライマーまたはプライマー対、本明細書において上述されるベクター、および/または本明細書において上述される抗体アプタマーおよび/またはファージを1つまたは複数の容器中に含むキットに関する。

40

**【0067】**

本発明は、本明細書において上述される核酸分子または本明細書において上述されるベクターの、遺伝子治療における使用にも関する。遺伝子治療アプローチは、本発明のベクターに関して本明細書において上述されており、同様に本明細書に当てはまる。本発明によれば、本明細書において上記に定義される、特に配列番号3ないし4で示される核酸分子の断片もまた、遺伝子治療アプローチに用いてもよいことに留意のこと。前記断片は、(c) 本明細書において上記で定義される (配列番号3でも示される) - 13910位のヌクレオチドまたは(d) 本明細書において上記で定義される (配列番号4でも示される) - 22018位のヌクレオチドを含む。前記断片は、少なくとも200、少なくとも25

50

0、少なくとも300、少なくとも400および最も好ましくは少なくとも500ヌクレオチドを含むことが好ましい。

【0068】

本発明の使用の好ましい実施形態では、前記遺伝子治療により、成人型ラクターゼ欠乏症が治療または予防される。

【0069】

以下の実施例により、本発明を説明する。

実施例1：連鎖および連鎖不平衡解析

2q21上のLPH遺伝子の両側にあるD2S114とD2S2385の間の7つの多型マイクロサテライトマーカーを、拡大された9つのフィンランド人ラクターゼ欠乏症家系において解析した(図1)。連鎖に関する有意な証拠が、マーカーD2S314、D2S442、D2S2196およびD2S1334で見出され、マーカーD2S2196で得られた $r = 0$ にて7.67の最高ロッドスコアを有した(表1)。絶対的な組換え現象は、ラクターゼ持続性/非持続性遺伝子座の動原体の境界を定義する、マーカーD2S114で検出され(家系B、IV3)、その遺伝子座のテロメアの境界を定義する、マーカーD2S2385で検出された(家族B、IV17)(図1、表1)。棄却域(critical region)を微細にマッピングするために、さらに9つの多型マーカーを解析した。対立遺伝子頻度および組換え率を攪乱母数として扱って、その領域に関する連鎖不平衡(LD)を、検出された連鎖について条件的にモニターした<sup>16-17</sup>。200kb区間に及ぶ、9つのマーカーのうち6つのマーカー(LPH13、LPH2、LPH1、AC3、AC4、およびAC10)は、LDの非常に有意な証拠( $p < 10^{-4}$ )を示すのに対して、LPH遺伝子からの他の3つのマーカーは、LDの証拠を示さなかった(表1)。2つのマーカー、LPHおよびAC3は、対立遺伝子のラクターゼ持続性において最も有意な連鎖不平衡を表した( $p < 10^{-7}$ )。

【0070】

その家系材料は、本来サヒ(Sahi)<sup>5</sup>によって研究された9つの拡大されたフィンランド人家系からなる。1970年代に、すべての家系材料が成人型ラクターゼ欠乏症について試験された。この研究に用いられる家系材料は、若い世代における家族のDNAを収集することによって拡大された。この研究における家系材料は、合計194個体からなった(図1)。すべての家族の表現型状態が、49個体を除くすべてにおいて、エタノールを用いた乳糖耐性検査(LTTE)<sup>4-5</sup>によって確認された。グルテン性腸症は、血清IgA抗組織トランスグルタミナーゼの測定によってすべての罹患者において除外されている<sup>45</sup>。インフォームド・コンセントを得た後、標準プロトコール<sup>46</sup>に従って、関与しているすべての家族から採取した血液試料からDNAを抽出した。症例対照研究として、空腸生検標本から単離された196個のランダムなDNA試料であって、その試料からの2糖類分解酵活性がヘルシンキ大学病院(Helsinki University Hospital)で測定されている47試料をシーケンスした。標準プロトコール<sup>146</sup>に従って、腸内の生検からDNAを単離した。これらの系列は、137個のラクターゼ持続性試料および59個の非持続性試料を含んだ。さらに、M. ロッシ(Rossi), University of Naplesによって親切にも提供されたイタリア人9名からのDNA、M. レンツェ(Lentze), University of Bonnによって親切にも提供されたドイツ人9名のDNA試料、J.K. セオ(Seo), Seoul National Universityによって親切にも提供された韓国人22名のDNA試料、腸内生検試料を分析した(表中:韓国人23名、イタリア人9名、ドイツ人7名(ドイツからの症例のうち1つは、韓国が起源である))。その診断は、二糖類分解酵素活性の測定に基づくものであった。最終的に、フィンランド人集団におけるC/T<sub>-13910</sub>変異の頻度を決定するために、東フィンランドおよび西フィンランドの小さな州からの匿名フィンランド人血液ドナー938名のDNAおよびCEPH家系<sup>19</sup>に属する109名の親のDNAを解析した。さらに、標準プロトコール<sup>48</sup>を用いて、肝生検から単離されたヒビ(Papio hemedryas ussinus)由来のDNAを解析した。この研究は、ヘルシンキ大学病院(Helsinki University Hospital)およびフィンランド赤十字輸血サービス(Finnish Red Cross Blood Transfusion Serv

ice) の倫理委員会 (Ethical Committee) によって承認された。

【0071】

実施例 2 : 拡大されたハプロタイプ解析

第 1 段階において、2q21 上の LPH 遺伝子の両側にある極めて多型の 10 個のマイクロサテライトマーカーが、他に記載されるように<sup>40, 55</sup>解析された。簡単に言えば、The Genethon Resource Center<sup>55</sup>からのラクターゼ遺伝子付近にある、2q 上の極めて多型のマイクロサテライトマーカーを以下のような遺伝距離 : cen - D2S114 - 1cM - D2S1334 - 0cM - D2S2196 - 0cM - D2S442 - 2cM - D2S314 - 2cM - D2S2385 - 1cM - D2S2288 - 1cM - D2S397 - 1cM - D2S150 - 1cM - D2S132 で解析した。マーカーの順序は大部分、Genethon マップで補われる第 2 染色体の物理的 YAC コンテグマップ (チューマコフ (Chumakov) ら, 1995<sup>56</sup>) から得られた。鋳型 DNA 12 ng、プライマー 5 pmol、各ヌクレオチド 0.2 mM、20 mM トリス HCl (pH 8.8)、15 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>、1.5 mM MgCl<sub>2</sub>、0.1% Tween 20、0.01% ゼラチン、0.25 U Taq ポリメラーゼ (Dynazyme, Finnzymes 社) を含有する合計体積 15 μl 中で PCR が行われた。<sup>32</sup>P - ATP を用いて、プライマーの 1 つを 5' 末端で放射標識した。その反応は、マルチウェル・マイクロタイタープレート中で、94 °C で 30 秒間の変性段階 ; プライマーに応じて、様々な温度にて 30 秒間アニーリングする段階 ; 72 °C で 30 秒間伸長する段階 ; 3 分間に設定された変性段階 ; および 5 分での最終伸長段階 ; を含む 35 サイクル行われた。増幅された断片を 6% ポリアクリルアミドゲル上で分離し、オートラジオグラフィを行った。第 2 段階では、LPH 遺伝子にわたり構築されたコンテグ内のその他の 9 個のマイクロサテライトマーカーを、Repeat Masker program (<http://ftp.genome.washington.edu/cgi-bin/RepeatMasker>) を用いて、BAC (NH034L23、NH0318L13、NH0218L22、および RP11-329I1) の公開ゲノム配列から同定した。その反復配列の両側にあるプライマーを合成した。PCR 条件は、他に記載の通りである<sup>40</sup>。増幅された断片を 6% ポリアクリルアミドゲル上で分離し、オートラジオグラフィを行った。

【0072】

LINKAGE プログラムパッケージ<sup>39</sup>の MLINK オプションを使用することによって、対をなすロッドスコアを計算した。完全浸透度を有し、組換え率の性差なく、疾患対立遺伝子頻度 0.4 を有する、成人型ラクターゼ欠乏症の常染色体劣性遺伝が推定された。その症状はフィンランド人集団<sup>5~6</sup>のその年齢に現れたため、20 歳の年齢を超える個体のみが、研究に含まれた。LTT E によって確認されていない個体の疾患状態は、未知であると見なされた。対立遺伝子頻度およびそのマーカーのヘテロ接合性が、パラメトリックな連鎖解析<sup>49</sup>の目的のために、Downfreq プログラムを用いて、家系材料から推定された。さらに、常染色体劣性型の遺伝を想定して<sup>16</sup>、疑似マーカー (pseudomarker) 連鎖および連鎖不平衡解析を行った。対立遺伝子頻度および組換え率を攪乱母数<sup>16, 49</sup>として扱って、検出された連鎖について LD の試験を条件的に行った。これらの解析から得られた P 値を表 1 に示す。ハプロタイプは、この順序 : LPH1 - LPH2 - LPH13 - AC7 - AC3 - AC4 - AC5 (図 3) でマイクロサテライトマーカーに対して手作業で構築された。本発明者らの家系材料における合計 54 個の非存続性染色体および合計 33 個の存続性染色体がハプロタイプ解析に利用可能であった。

【0073】

密に連鎖するマーカーの順序は、棄却域における 4 つの BAC クローン : NH0034L23、NH0218L22、NH0318L13 および 329I10 を 1 つの連続した配列セグメントに構築することによって確認された。このコンテグは、アスパルチル - tRNA シンターゼ (DARS) 遺伝子のマーカー AC8 からエクソン 10 におよび、合計 222.5 kb を網羅した (図 2)。連鎖領域のこの物理地図に基づいて、150 kb 区間を網羅する、7 つのマーカーを有する拡大されたハプロタイプ (cen - LPH13 - LPH2 - LPH1 - AC7 - AC3 - AC4 - AC5 - tel) (図 3) が構築され

た。1つの主要なハプロタイプが、非存続性対立遺伝子のうち3個(5%)に対して、存続性対立遺伝子20個(60%)中に存在し、ハプロタイプの広範な多様性が非存続性対立遺伝子で認められた。存続性対立遺伝子におけるハプロタイプの残り40%は、歴史的組換え現象による名プロタイプの破壊と一致するように、祖先ハプロタイプと異なった。保存ハプロタイプ解析に基づいて、ラクターゼ存続性の遺伝子座は、マーカーLPH1とAC3との間の47kb区間に限定される可能性がある(図3)。

#### 【0074】

実施例3：成人型ラクターゼ欠乏症遺伝子座の配列解析

マーカーLPH1とAC3との間の47kb領域を、9つのラクターゼ欠乏症家系の家族数人のゲノムDNAからの、オーバーラップPCR断片において増幅し、シーケンスした。その領域は、47kb棄却域(critical region)の36kbを網羅する、微小染色体維持(MCM6)遺伝子<sup>18</sup>を含有する(図2)。合計52個の変異体を除いては、MCM6遺伝子のコード領域では変異は検出されなかった；43個のSNPおよび9個の欠失/挿入多型が、47kb棄却域(critical region)において同定された(表2)。変異体のうち2つだけ(C/T<sub>-13910</sub>、G/A<sub>-22018</sub>)が、フィンランド人家系におけるラクターゼ存続性/非存続性形質と関連した(表2および3)。第1の関連する変異体、C/T<sub>-13910</sub>は、LPH遺伝子の最初のATGコドンから-13910bp位にあるMCM6遺伝子のイントロン13に存在する。第2の関連する変異体、G/A<sub>-22018</sub>は、LPH遺伝子の最初のATGコドンから-22018位にあるMCM6遺伝子のイントロン9に位置する。互いに8kb離れた、これらの2つの変異体は完全に、拡大された9つのフィンランド人人家系における成人型ラクターゼ欠乏症で同時分離した。すべてのラクターゼ欠乏の(非存続性)家族は、C<sub>-13910</sub>とG<sub>-22018</sub>の両方にホモ接合であった(表3)。興味深いことに、これらのどちらの変異体も、反復要素中に存在し、L2誘導要素にはC/T<sub>-13910</sub>、Alu要素にはG/A<sub>-22018</sub>が存在する。

#### 【0075】

実験的に、本発明者らの家系材料からの、棄却域にわたり同様のハプロタイプを共有する、非存続性個体3名、ホモ接合の存続性個体2名およびヘテロ接合の存続性個体2名を、第1段階におけるシーケンシングに使用した(図1)。成人型ラクターゼ欠乏症の棄却域を網羅する、BAC:NH0034L23、NH0218L22、NH0318L23、およびRP-329I10の公開ドラフトゲノム配列を用いて、Sequencher 4ソフトウェア(Gene Codes社)を使用して1つのコンティグに構築した。LPH1とAC3との間の棄却域に及ぶオリゴヌクレオチドプライマーを設計した(オリゴヌクレオチドプライマーのリストは本明細書において以下に記載される)。ゲノムDNA(100ng)、プライマー(各20ng)、dNTP(200μM)、Taqポリメラーゼ0.5U(Dynazyme, Finnzymes社)を有する合計体積50μlで、標準バッファー中にてPCR増幅が行われた。PCR産物のサイズが、本発明者らがDynazyme extend kit(伸長キット)を用いた(条件は本明細書において以下に記述される)1kbを超える場合を除いては、大部分のPCRは、以下のPCRサイクル条件：94 で3分間変性を行う初回、次いで、94 で30秒、55 で30秒、および72 で1.25分の35サイクル、72 で10分間の最終伸長を用いて増幅された。精製されたPCR産物(15~40ng)を、BigDye terminator chemistry(PE Biosystems社)を用いてサイクルシーケンスした。ABI Sequencing Analysis 3.3(PE Biosystems社)およびSequencher 4.1(Gene Codes社)を用いて、データを解析した。

#### 【0076】

シーケンシングによるラクターゼ変異体の検出：

ゲノムDNA(100ng)、プライマー(各20ng)、dNTP(200μM)、0.5U Taqポリメラーゼ(Dynazyme, Finnzymes社)を有する合計体積50μlで、標準バッファー中にて、PCR増幅を行った。どちらのPCRも、以下のPCRサイクル条件：94 で3分間変性を行う初回、次いで、94 で30秒、55 で30秒、および72 で1.25分の35サイクル、72 で10分間の最終伸長を用いて増幅された。

PCRは酵素反応によって精製された。精製されたPCR産物(15~40ng)を、BigDye terminator chemistry (PE Biosystems社)を用いてサイクルシーケンスした。ABI Sequencing Analysis 3.3(PE Biosystems社)およびSequencher 4.1(Gene Codes社)を用いて、データを解析した。

#### 【0077】

固相ミニシーケンシングによるラクターゼ変異体のスクリーニング:

C/T<sub>-13910</sub>変異体にわたるDNA断片を、1つのビオチン化(5'-Bio-CCTCGTTAATACCCACTGACCTA-3')プライマーおよび非ビオチン化(5'-GTCACCTTTGATATGATGAGAGCA-3')プライマーを用いて増幅した。G/A<sub>-22018</sub>には、ビオチン化(5'-Bio-TGCTCAGGACATGCTGAATCAA-3')および1つの非ビオチン化(5'-CTACCCCTATCAGTAAAGGCCCTA-3')プライマーを上述の条件下で使用した。PCR産物10μlをストレプトアビジンコーティングのマイクロタイターウェル(Lab systems社, フィンランド)中に捕捉した。そのウェルを洗浄し、サイバネン(Syvanen)ら(Am J Hum Genet. (1993), 52, 46-59)およびサイバネン(Syvanen)およびランデグレン(Landegren) (Hum Mutat. (1994), 3, 172-9)によって記述されているように、結合したDNAを変性させた。C/T<sub>-13915</sub>にミニシーケンシングプライマー(5'-GGCAAATACAGATAAGATAATGTAG-3')10pmol、G/A<sub>-22018</sub>にはミニシーケンシングプライマー(5'-AAAAACAGCATTCCTCAGCTGGGC-3')10pmol、およびラクターゼ非存続性対立遺伝子に対応するH-dCTP、H-dGTP(115Ci/mmol; Amersham社, 英国)またはラクターゼ存続性対立遺伝子に対応するH-dTTP、H-sATPのいずれか0.1μl、およびDNAポリメラーゼ0.05U(Dynazyme II, Finnzymes社)をそのバッファー中に含有する、ミニシーケンシング反応混合物50μlを各ウェルに添加した。マイクロタイタープレート(50)で20分間インキュベートし、ウェルを洗浄した。検出用プライマーを溶出し、溶出された放射能を液体シンチレーション計数器(Rackbeta 1209, Wallac, フィンランド)で測定した。2つの並行なミニシーケンシング反応が各PCR産物に対して行われた。

#### 【0078】

C/T<sub>-13910</sub>変異体に対するPCRプライマーおよび検出プライマー:

フォワードPCRプライマー: GTCACCTTTGATATGATGAGAGCA Tm 58 配列番号8

検出プライマー: GGCAAATACAGATAAGATAATGTAG Tm 58 配列番号10

Bio-リバープライマー: Bio-CCTCGTTAATACCCACTGACCTA Tm 62 配列番号9

またはBio-TAGGTCAGTGGGTATTAAACGAGGT 配列番号7

G/A<sub>-22018</sub>変異体に対するPCRプライマーおよび検出プライマー:

フォワードPCRプライマー: CTACCCCTATCAGTAAAGGCCCTA Tm 58 配列番号12

検出プライマー: AAAAACAGCATTCCTCAGCTGGGC Tm 62 配列番号14

Bio-リバープライマー: Bio-TGCTCAGGACATGCTGATCAA Tm 62 配列番号13

またはBio-TTGATCAGCATGTCTCAGAGCA 配列番号11

#### 【0079】

実施例4: 症例対照研究試料におけるDNA変異体のモニタリング

ラクターゼ欠乏症の診断試験として二糖類分解酵素活性について解析された、合計196個の腸内生検標本から単離されたDNA試料において、C/T<sub>-13910</sub>およびG/A<sub>-22018</sub>変異体の頻度を解析した。合計59個の試料が、原発性ラクターゼ欠乏症(primary lactase deficiency)を示した。59症例のうち6症例が(表3)、G/A<sub>-22018</sub>変異体に対



の分布と相関するのかがどうかを決定するために、CEPH家系<sup>22</sup>のDNAを解析した。CEPH家系は主に、フランスおよびユタ州から集められ、フランスからの家系は報告されたラクターゼ欠乏症有病率約37%<sup>23</sup>を有し、北ヨーロッパ由来のユタ州の集団は5%未満<sup>24</sup>のラクターゼ欠乏症有病率を有した。CEPH家系において親を遺伝子型判定することによって、フランス人家系の41.2% (17試料のうちの7)が遺伝子型CCを有するのに対して、ユタ州家系の7.6%のみ (92試料のうちの7)が遺伝子型CCを有することが明らかとなった (表4)。さらにまた、解析された試料が少数であるにもかかわらず、これらの数値は、これらの集団<sup>23,24</sup>におけるラクターゼ欠乏症の疫学的研究で得られた値と一致する。観察された変異体の有病率が、記述されている乳糖不耐の集団での頻度とよく一致していることが、表8に示されている。

10

## 【0083】

実施例6：ラクターゼ存続性変異体C/T<sub>-13910</sub>の系統学

フィンランド人家系におけるハプロタイプ解析から、フィンランドにおけるラクターゼ存続性対立遺伝子のすべてではないが大部分が、ある共通の祖先に由来していることが示唆された。連鎖不平衡を利用して、存続性対立遺伝子のフィンランド人集団<sup>25</sup>への導入の時間を推定した。世代時間が20年であると仮定すると、この推定から、創始者の変異がフィンランド人集団におよそ9000~11400年前に導入されたことが示されるだろう (表6)。これは、およそ8000~9000年前<sup>26</sup>にフィンランド本土に移住した初期の徴候とよく一致し、理にかなって、紀元前8000~10,000年の酪農業の始まりとかなり一致するだろう。さらに重要なことには、異なる集団における存続性対立遺伝子中に同じDNA変異体が存在することから、この変異体はよりいっそう偶発的であり、かつその変異は、解析した集団の派生前に生じたことが示唆されるだろう。

20

## 【0084】

ラクターゼ対立遺伝子の系統発生的起源を洞察するために、ヒビ (Papio Hamadryas) のMCM6遺伝子のイントロン9およびイントロン13の一部をシーケンスした。遺伝子型GGおよびCCが、G/A<sub>-22018</sub>およびC/T<sub>-13910</sub>の両方でヒビDNAに存在した。このことから、対立遺伝子GおよびCはそれぞれ、非存続型を表す、祖先対立遺伝子の出現を反映しており、変異によりこの対立遺伝子が形質転換され、存続性対立遺伝子が生じたことが示唆される。この仮定は、非存続性対立遺伝子で見出される対立遺伝子の高い多様性に対して、存続性対立遺伝子におけるLDおよび共有されるハプロタイプを同定することによって支持される。

30

## 【0085】

実施例7：C/TおよびG/A変異体の対をなすLD

C/T<sub>-13910</sub>とG/A<sub>-22018</sub>との間の対のLDを、D'統計量<sup>21</sup>を用いて推定した。D'は、最大値 (D/D<sub>max</sub>、D/D<sub>min</sub>)として計算される：不平衡測定法  $D = h_{pq} - pq$  (式中、 $h_{pq}$ は、各遺伝子座に稀な対立遺伝子を有するハプロタイプの頻度であり、pおよびqは、遺伝子座1および2における稀な対立遺伝子の頻度である)、 $D_{max} = \min(p(1-p), q(1-q))$  (D > 0の場合)、および  $D_{min} = -\min(pq, (1-p)(1-q))$  (D < 0の場合)。自由度 (1df) で<sup>2</sup>として分布する統計量：

## 【数1】

40

$$D^2 \sqrt{\frac{N}{p(1-p)q(1-q)}}$$

を用いて、0からのD'の偏差fの有意性を決定した<sup>21</sup>。

## 【0086】

遺伝子受入番号：BACでは、NH0218L22、N0034L34、NH0318L13およびRP11-329I10はそれぞれ、AC012551、AC011893、

50

A C 0 1 1 9 9 9 および C 0 1 6 5 1 6 である。ヒト多型の受入番号は G e n B a n k A F 3 9 5 6 0 7 ~ A F 3 9 5 6 1 5 である。

【 0 0 8 7 】

#### 参考文献

- 1 . フラッツ (Flatz) , G. & ロッソウエ (Rotthauwe) , H. The human lactase polymorphism: physiology and genetics of lactose absorption and malabsorption. *Prog. Med. Genet.* 2, 205-249 (1977).
- 2 . サヒ (Sahi) , T. , イソコスキー (Isokoski) , M. , ジュッシラ (Jussila) , J. & ラウニアラ (Launiala) , K. Lactose malabsorption in Finnish children of school age. *Acta Paediatr Scand.*61, 11-16 (1972). 10
- 3 . ワン (Wang) , Y. ㄆ. The genetically programmed down-regulation of lactase in children. *Gastroenterology.* 114:1230-1236 (1998).
- 4 . サヒ (Sahi) , T. , イソコスキー (Isokoski) , M. , ジュッシラ (Jussila) , J. & ラウニアラ (Launiala) , K. & ピョーララ (Pyorala) , K. Recessive inheritance of adult-type lactose malabsorption. *Lancet.*823-826 (1973).
- 5 . サヒ (Sahi) , T. The inheritance of selective adult-type lactose malabsorption. *Scand. J. Gastroenterol. suppl.* 30, 1-73(1974).
- 6 . サヒ (Sahi) , T. Genetics and epidemiology of adult-type hypolactasia. *Scand. J. Gastroenterol. Suppl.*202, 7-20 (1994).
- 7 . ボール (Bohl) , W. , ワーグナー (Wagner) , P. & マンテイ (Mantei) , N. Structure of the chromosomal gene and cDNAs coding for lactase-phlorizin hydrolase in human with adult-type hypolactasia or persistence of lactase. *Am. J. Hum. Genet.* 48, 889-902 (1991). 20
- 8 . マンテイ (Mantei) , Nㄆ. Complete primary structure of human and rabbit lactase-phlorizin hydrolase: implications for biosynthesis, membrane anchoring and evolution of the enzyme. *EMBO J.* 7, 2705-2713 (1988).
- 9 . ワン (Wang) , Y. ㄆ. The lactase persistence/non-persistence polymorphism is controlled by a cis-acting element. *Hum. Mol. Genet.*4, 657-662 (1995).
- 10 . ハーベイ (Harvey) , C.B. , Pratt, W.S. , Islam, I. , Whitehouse, D.B. & Swallow, D.M. DNA polymorphisms in the lactase gene: linkage disequilibrium across the 70 kb region. *Eur J. Hum. Genet.* 3, 27-41 (1995). 30
- 11 . エッシャー (Escher) , J.C ㄆ. Molecular basis of lactase levels in adult humans. *J. Clin. Invest.* 89, 480-483 (1992).
- 12 . ロイド (Lloyd) , Mㄆ. Regulation of intestinal lactase in adult hypolactasia. *J. Clin. Invest.* 89, 524-529 (1992).
- 13 . ファハルド (Fajardo) , O. , ナイム (Naim) , H.Y. & Lacey, S.W. The polymorphic expression of lactase in adults is regulated at the messenger RNA level. *Gastroenterology* 106, 1233-
- 14 . ルイーヂ (Luigi) , M.ㄆ. Mosaic regulation of lactase in human adult-type *Gastroenterology* 112, 1506-1514 (1997). 40
- 15 . ロッシ (Rossi) , M.ㄆ. Lactase persistence versus decline in human adults: Multifactorial events are involved in down-regulation after weaning. *Gastroenterology* 112, 1506-1514 (1997).
- 16 . ゴーリング (Goring) , H.H.H. & ターウィルガー (Terwilliger) , J.D. Linkage analysis in the presence of errors IV: Joint pseudomarker analysis of linkage and / or linkage disequilibrium on a mixture of pedigrees and singletons when mode of inheritance cannot be accurately specified. *Am. J. Hum. Genet.* 66, 1310-1327 (2000).
- 17 . ターウィルガー (Terwilliger) , J.D. & ゴーリング (Goring) , H.H.H. Gene mapping in the 20th and 21st centuries: Statistical methods, data analysis, and e 50

- xperimental design. *Hum. Biol.* 72, 63-132 (2000).
- 18 .ハーベイ (Harvey), C.B.ら. Regional localization of the lactase-phlorizin hydrolase, LCT, to chromosome 2q21. *Ann. Hum. Genet.* 57, 179-185 (1993).
- 19 .サイバネン (Syvanen), A.-C., サジャンティラ (Sajantila) A., ルッカ (Lukka), M. Identification of individuals by analysis of biallelic DNA markers, using PCR and solid-phase minisequencing. *Am. J. Hum. Genet.* 52, 46-59 (1993).
- 20 .サイバネン (Syvanen), A.-C. & ランデグレン (Landegren), U. Detection of point mutations by solid-phase methods. *Hum. Mutat.* 3, 172-179 (1994)
- 21 .トンプソン (Thompson), E. A., ディーブ (Deeb), S., ウォーカー (Walker), D. & モトルスキー (Motulsky), A. G. The detection of Linkage disequilibrium between closely linked markers: RFLPs at the AI-CIII Apolipoprotein genes. *Am. J. Hum. Genet.* 42, 113-124 (1998). 10
- 22 .ドーセ (Dausset), J.ら. Centre d'étude du polymorphisme humain (CEPH): Collaborative genetic mapping of human genome. *Genomics* 6, 575-577 (1990).
- 23 .カッデネク (Cuddenec), Y., デルブルック (Delbruck), H. & フラッツ (Flatz), G. Distribution of the adult lactase phenotypes- lactose absorber and malabsorber- in a group of 131 army recruit *Gastroenterol.Clin. Biol.* 6, 776-779 (1982)
- .
- 24 .マクレラン (McLellan), T., ジョーデン (Jorde), L.B. & スコルニック (Skolnick), M.H. Genetic distance between the Utah Mormons and related populations. *Am. J. Hum. Genet.* 36, 836-857 (1984). 20
- 25 .ターウィルガー (Terwilliger), J.D. A powerful likelihood method for the analysis of linkage disequilibrium between trait loci and one or more polymorphic marker loci. *Am. J. Hum. Genet.* 56, 777-787 (1995).
- 26 .ヌニェス (Nunez), M.G. A model of the early settlement of Finland. *Fennoscandia archaeologica* IV, 3-18 (1997).
- 27 .シムーンズ (Simoons), F.J. Primary adult lactose intolerance and the milk ing habit: a problem in biological and cultural interrelations.II. A cultural historical hypothesis. *Am. J. Dig. Dis.* 16, 695-710 (1970).
- 28 .バリーロ (Varilo), T.ら. The age of human mutation: genealogical and linkage disequilibrium analysis of the CLN5 mutation in the Finnish population. *Am. J. Hum. Genet.* 58, 506-512 (1996). 30
- 29 .ハストバック (Hastbacka), J.ら. Linkage disequilibrium mapping in isolated founder populations: diastrophic dysplasia in Finland. *Nature Genet.* 2: 204-211 (1992).
- 30 .ハーベイ (Harvey) C.B.ら. Lactase haplotype frequencies in Caucasians: association with the lactase persistence/non persistence polymorphism. *Ann Hum Genet* 62, 215-223 (1998).
- 31 .オータニ (Ohtani), K.ら. Cell growth-regulated expression of mammalian MCM5 and MCM6 genes mediated by the transcription factor E2F. *Oncogene* 18, 2299-2309 (1999). 40
- 32 .スミス (Smith), A.F.A. The origin of interspersed repeats in the human genome. *Curr. Opin. Genet. Dev.* 6, 743-748 (1996).
- 33 .カザジアン (Kazazian), H.H. & モーラン (Moran), J.V. The impact of L1 retrotransposons on the human genome. *Nature Genet.* 19, 19-24 (1998).
- 34 .モーラン (Moran), J.V., デベラルディニス (DeBerardinis), R.J. & カザジアン (Kazazian), H.H. Exon shuffling by L1 retrotransposition. *Science* 283, 1530-1534 (1999).
- 35 .ウェイ (Wei), W.ら. Human L1 retrotransposition: cis preference versus trans complementation. *Mol. Cell. Biol.* 21, 1429-1439 (2001). 50

36. ドネリー (Donnelly), S.R., ホーキンス (Hawkins), T.E. & モス (Moss), S. E. A conserved nuclear element with a role in mammalian gene regulation. *Hum. Mol. Genet.* vol.8, 9,1723-1728 (1999).
37. ボーケ (Boeke), J.D. LINEs and Alus - the polyA connection. *Nature Genet.* 16, 6-7 (1997).
38. ジャーカ (Jurka), J. Sequence patterns indicate an enzymatic involvement in integration of mammalian retroposons. *Proc. Natl. Acad. Sci. U.S.A.* 94, 1872-1877 (1997).
39. サビラティ (Savilahti) E, ラウニアラ (Launiala) K, キートネン (Kuitunen) P. Congenital lactase deficiency. *Arch. Dis. Child.* 58, 246-252 (1983). 10
40. ジャーベラ (Jarvela), I. Assignment of the locus for congenital lactase deficiency to 2q21, in the vicinity of but separate from the lactase-phlorizin hydrolase gene. *Am. J. Hum. Genet.* 63, 1078-1085 (1998).
41. シムーンズ (Simoons), F. J. The geographic hypothesis and lactose malabsorption. A weighing of the evidence. *Am. J. Dig. Dis.* 23, 963-980 (1978).
42. フラッツ (Flatz), G. & ロツソウエ (Rotthauwe), H, W. The human lactase polymorphism: physiology and genetics of lactose absorption and malabsorption. *Prog. Med. Genet.* 2, 205-249 (1977).
43. マクラッケン (McCracken), R.D. Lactase deficiency: an example of dietary evolution. *Curr. Anthropol.* 12, 479-517 (1971). 20
44. アローラ (Arola), H. Diagnosis of hypolactasia and lactose malabsorption. *Scand. J. Gastroenterol. Suppl.* 202, 26-35 (1994).
45. サルカネン (Sulkanen), S. Tissue transglutaminase autoantibody enzyme-linked immunosorbent assay in detecting celiac disease. *Gastroenterology* 115 (6), 1322-1328 (1998).
46. サンプルック (Sambrook), J., フリツチェ (Fritsch), E.F. & マニアティス (Maniatis), T. *Molecular cloning: a laboratory manual*, (2nd ed). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989).
47. メッサー (Messer), M. & ダルクヴィスト (Dahlqvist), A. A one - step ultramicro method for the assay of intestinal disaccharidases. *Anal. Biochem.* 14 (3) 30, 376-92 (1966).
48. コッティングハム (Cottingham), Jr. R.W., イダリー (Idury), R.M. & シェーファー (Schaffer), A.A. Faster sequential genetic linkage computations. *Am. J. Hum. Genet.* 53, 252-263 (1993).
49. ゴーリング (Goring), H.H.H. & ターウィルガー (Terwilliger), J.D. Linkage analysis in the presence of errors III: Marker loci and their map as nuisance parameters. *Am. J. Hum. Genet.* 66,1298-1309 (2000).
50. ターウィルガー (Terwilliger), J.D. & オット (Ott), J. *Hand book of human genetic analysis*. Johns Hopkins University Press, Baltimore (1994).
51. オズボーン (Osborne)ら, *Curr. Opin. Chem. Biol.* 1 (1997), 5-9 40
52. ストール (Stall) およびゾッカ (Szoka), *Pharm. Res.* 12 (1995), 465-483
53. ハーロー (Harlow) およびレーン (Lane) "Antibodies, A Laboratory Manual", CSH Press, Cold Spring Harbor, USA, 1988
54. ヒギンス (Higgins) およびヘイズ (Hames) (eds.), "Nucleic acid hybridization, a practical approach", IRL Press, Oxford 1985
55. ディブ (Dib) C, フォール (Faure) S, フィザムズ (Fizames) C, サムソン (Samson) D, ドルオー (Drouot) N, ビニャル (Vignal) A, ミラシュュー (Millasseau) P, マルク (Marc) S, ハザン (Hazan) J, セボーン (Seboun) E, ラスロップ (Lathrop) M, ギャペイ (Gyapay) G, モリセット (Morissette) J, ワイセンベック (Weissenbach) J. A comprehensive genetic map of the human genome based on 5,264 microsa 50

tellites. Nature. 1996 Mar 14;380(6570):152-4.

56. チューマコフ (Chumakov) IM, リガウルト (Rigault) P, レ・ゴール (Le Gall) I, ベラーネ (Bellanne) -シャンテロット (Chantelot) C, ビラウルト (Billault) A, ギユー (Guillou) S, ソウラリュウ (Soularue) P, グアスコニ (Guasconi) G, ポウリアー (Poullier) E, グロス (Gros) I, 他. A YAC contig map of the human genome. Nature. 1995 Sep 28;377(6547 Suppl):175-297

【 0 0 8 8 】

【 表 1 】

表 1. 成人型ラクターゼ欠乏症家系における連鎖および連鎖不平衡解析 (微細マッピングマーカを太字で示す)

10

マーカー	Θでのロッドスコア (Z)					p値 <sup>a</sup>
	0.0	0.1	0.2	0.3	0.4	
D2S114	-∞	2.44	1.92	1.13	<b>0.41</b>	0.87195
P6112	2.76	2.20	1.45	0.75	<b>0.22</b>	0.66207
D2S1334	3.15	2.45	<b>1.61</b>	0.84	0.25	0.91039
AC8	<b>2.26</b>	1.99	1.36	0.71	0.21	0.53670
LPH13	3.67	2.94	1.96	1.03	0.31	4x10 <sup>-6</sup>
LPH2	4.09	3.07	2.00	1.00	0.26	5.7x10 <sup>-7</sup>
<b>LPH1</b>	5.91	4.52	2.96	1.53	0.46	5x10 <sup>-6</sup>
<b>AC7</b>	3.63	2.60	1.66	0.83	0.23	0.03471
<b>AC3</b>	6.63	4.88	3.16	1.61	0.44	3.2x10 <sup>-8</sup>
<b>AC4</b>	3.07	2.22	1.42	0.71	0.19	4x10 <sup>-5</sup>
<b>AC5</b>	5.33	4.10	2.72	1.39	0.39	0.02166
<b>AC10</b>	6.60	4.99	3.25	1.65	0.46	1x10 <sup>-5</sup>
D2S2196	7.67	5.62	3.62	1.85	0.54	0.00010
D2S442	3.81	3.08	2.08	1.03	0.27	0.22805
D2S314	4.22	3.61	2.50	1.37	0.45	0.27535
D2S2385	-∞	2.79	1.92	1.01	0.28	0.46457

20

30

40

a: 連鎖を仮定して連鎖不平衡試験を用いて得られたp値<sup>16,49</sup>

【 0 0 8 9 】

【 表 2 】

表2. フィンランド人家系における成人型ラクターゼ欠乏症遺伝子座内で  
同定された変異

位置 <sup>a</sup>	変異体	ラクターゼ存続性		ラクターゼ存続性		ラクターゼ非存続性		
		(ホモ接合)		(ヘテロ接合)				
		BIV4	AIV3	BIV8	CIV3	BIV9	DIV4	EIII2 <sup>b</sup>
-694	A→G	AA	AA	AG	AA	GG	N <sup>c</sup>	AA
-1640/50	T <sub>13</sub> →T <sub>12</sub>	T <sub>13/13</sub>	T <sub>13/13</sub>	T <sub>13/13</sub>	T <sub>13/13</sub>	T <sub>13/13</sub>	T <sub>12/12</sub>	T <sub>12/12</sub>
-2131	C→T	CC	CC	CT	CC	TT	CT*	TT
-3058/72	T <sub>15</sub> →T <sub>16</sub>	T <sub>15/15</sub>	T <sub>15/15</sub>	T <sub>15/15</sub>	T <sub>15/15</sub>	T <sub>15/15</sub>	T <sub>16/16</sub>	T <sub>16/16</sub>
-3075	G→T	GG	GG	GG	GG	GG	GG	TT
-4480	T→A	TT	TT	TA	TT	AA	TT	TT
-5440	C→T	CC	CC	CT	CC	TT	CC	CC
-5926	A→T	AA	AA	AA	AA	AA	TA	TT
-8540	G→A	GG	GG	GA	GA	AA	AG	AA
-8630	C→G	CC	CC	CG	CG	GG	GC	GG
-13495	T→C	TT	TT	TC	TT	CC	CT	CC
-13910	T→C	TT	TT	TC	TC	CC	CC	CC
-15239	G→A	GG	GG	GA	GG	AA	AG	AA
-15862	T→C	CC	CC	CT	CC	TT	TC	TT
-16568/79	T <sub>11</sub> →T <sub>12</sub>	T <sub>11/11</sub>	T <sub>11/11</sub>	T <sub>11/12</sub>	T <sub>11/11</sub>	T <sub>12/12</sub>	T <sub>11/11</sub>	T <sub>12/12</sub>
-16888	A→G	AA	AA	GA	AA	GG	GA	GG
-17300	C→T	CC	CC	CC	CC	CC	CT	TT

10

20

30

-19044	T→C	TT	TT	TC	TT	CC	CT	CC	
-19519	T→C	TT	TT	TC	TT	CC	TT	TT	
-20077	C→G	CC	CC	CG	CC	GG	GC	GG	
-20486	G→A	GG	GG	GA	GG	AA	GG	GG	
-21721/28	A <sub>7</sub> →A <sub>6</sub>	A <sub>7/7</sub>	A <sub>7/7</sub>	A <sub>7/7</sub>	A <sub>7/7</sub>	A <sub>7/7</sub>	A <sub>7/A<sub>6</sub></sub>	A <sub>7/7</sub>	
-21731	A→C	AA	AA	AA	AA	AA	CC	AA	10
-21736/43	A <sub>9</sub> →A <sub>8</sub>	A <sub>9/9</sub>	A <sub>9/9</sub>	A <sub>9/A<sub>8</sub></sub>	A <sub>9/9</sub>	A <sub>8/8</sub>	A <sub>8/8</sub>	A <sub>8/8</sub>	
-22018	G→A	AA	AA	AG	AG	GG	GG	GG	
-22741	C→T	CC	CC	CC	CC	CC	N	TT	
-22788	A→G	AA	AA	AG	AA	GG	N	GG	
-23069	A→G	AA	AA	AG	AA	GG	N	GG	
-23442	A→G	AA	AA	AA	AA	AA	N	GG	20
-23771	T→C	TT	TT	TT	TT	TT	N	CC	
-25093/23	Δ30 bp	Δ Δ	Δ Δ	Δ Δ	Δ Δ	Δ Δ	N	II	
-27310	A→G	AA	AA	AG	AA	GG	GA	GG	
-27480	G→A	GG	GG	GA	GG	AA	AG	AA	
-27807	A→C	AA	AA	AA	AA	AA	AC	CC	
-30183	A→G	AA	AA	AG	AA	GG	AA	AA	30
-31268	A→G	AA	AA	AG	AA	GG	AA	AA	
-31342	T→C	TT	TT	TT	TT	TT	CT	CC	
-33645	C→T	CC	CC	CT	CC	TT	CC	CC	
-35176	T→C	TT	TT	TC	TT	CC	CT	CC	
-36254	C→T	CC	CC	CT	CC	TT	TC	TT	

-36296	G→T	TT	TT	TG	TT	GG	TG	N
-36501	A→T	AA	AA	AT	AA	TT	AT	N
-36506/14	Δ 9 bp	ΔΔ	ΔΔ	ΔI	ΔΔ	II	ΔI	N
-36671/77	T7→T6	T <sub>77</sub>	T <sub>77</sub>	T <sub>76</sub>	T <sub>77</sub>	T <sub>66</sub>	T <sub>77</sub>	T <sub>77</sub>
-37565	T→G	TT	TT	TG	TT	GG	GG	TG
-38276	G→C	GG	GG	GC	GG	CC	GG	GG
-39036	G→C	GG	N	GC	N	CC	N	N
-40608	G→C	GG	GG	GG	GG	GG	GC	CC
-41590	T→C	TT	TT	TC	TT	CC	CT	CC
-42081/82	ΔAG	AG	AG	AG/Δ	AG	ΔΔ	AG	AG
-42618	T→C	TT	TT	TC	TT	CC	TT	TT
-42893	G→A	GG	GG	GA	GG	AA	GG	GG

10

20

a : BAC、NH0218L22、N0034L34、NH0318L13およびRP11-329110の対応するゲノム配列を用いた、LPH遺伝子の翻訳開始コドン (ATG)からの数字である。b : 研究され、図1において矢印で示されるフィンランド人家系からシーケンスされた個体。  
c : 決定されていない。

【 0 0 9 0 】

【 表 3 】

表3. ラクターゼ存続性／非存続性対立遺伝子におけるC/T<sub>-13910</sub>およびG/A<sub>-22018</sub> 遺伝子型の分布

遺伝子型	C/T <sub>-13910</sub>			G/A <sub>-22018</sub>			合計	
	CC	CT	TT	GG	GA	AA		
家族のメンバー	ラクターゼ非存続性	45	0	0	45	0	0	45
	ラクターゼ存続性	0	32	13	0	32	13	45
症例対照試料 フィンランド人	ラクターゼ非存続性	59	0	0	53	6	0	59
	ラクターゼ存続性	0	63	74	0	63	74	137
非フィンランド人 <sup>a</sup>	ラクターゼ非存続性	40	0	0	39	1	0	40
	ラクターゼ存続性	0	5	0	0	5	0	5
合計	ラクターゼ非存続性						0	144
	ラクターゼ存続性							187

a : 非フィンランド人試料は、韓国人23個体、イタリア人9個体、ドイツ人8個体からなる。

【 0 0 9 1 】  
【 表 4 】

表 4. 集団試料における C/T -13910 の有病率

DNA試料	遺伝子型			合計	対立遺伝子 頻度 (%)		(CC) 遺伝子型%
	CC	CT	TT		C	T	
I. フィンランド人集団 :							
1. 東領域	108	287	176	571	0.440	0.560	18.9%
2. 西領域	62	159	146	367	0.385	0.615	16.8%
合計	170	446	322	938	0.418	0.582	18.1%
II. CEPHの親 :							
1. ユタ州家系	7	33	52	92	0.255	0.745	7.6%
2. フランス人家系	7	9	1	17	0.676	0.324	41.2%

10

20

フィンランド内の東および西領域に由来する匿名フィンランド人血液ドナー合計938名のDNA試料、およびCEPHの親から109個のDNA試料。集団におけるラクターゼ欠乏症の有病率は、CC対立遺伝子の遺伝子型頻度によって反映されている。

【 0 0 9 2 】

【 表 5 】

表5. ランダムなフィンランド人試料におけるC/T-13910およびG/A-220018変異体間のLD

	C/T <sub>-13910</sub> での遺伝子型			合計	D'	$\chi^2$ (1 df)	P値
	CC	CT	TT				
G/A <sub>-220018</sub> での遺伝子型							
GG	162	2	1	165			
GA	6	440	3	449			
AA	2	4	318	324			
合計	170	446	322	938	0.984	42.41	$7.62 \times 10^{-11}$

LDはD'統計量<sup>18</sup>を用いて計算された。p値は、方法<sup>18</sup>に記載のように0からのD'の有意性である。

【 0 0 9 3 】

【 表 6 】

10

20

30

40

表6. DISLAMBプログラムを用いた、フィンランド人集団へのC/T-13910の導入の推定。

マーカー	AC3		LPH2		
対立遺伝子 非存続性	ラクターゼ 存続性	ラクターゼ 非存続性	ラクターゼ 存続性	ラクターゼ 非存続性	
1	0	1	0	1	
2	31	10	0	20	10
3	0	1	0	14	
4	2	9	32	15	
5	0	31	0	2	
$\lambda^a$	0.838		0.999		
$\Theta^b$	0.00031 (0,000038-0.00099)		0.0000(0.00000-0.00052)		20
$n^c$	570		450		

a:  $\lambda$ は、その集団の頻度(0.60)に対する、疾患染色体における特定の対立遺伝子(ラクターゼ存続性対立遺伝子)の増加の割合である。b:  $\Theta$ は、組換え率であり、 $1\text{cM}=1\text{Mb}$ と仮定し、最も近いマーカーからの変異の距離によって示される。c:  $n$ は、集団への創始者の変異の導入以来の世代数であり、 $\lambda = \alpha(1-\Theta)^n$ の式が当てはまる。d:  $\Theta$ として計算に用いられる仮定の対立遺伝子はゼロであり、 $\alpha$ は1である。

【 0 0 9 4 】

【 表 7 】

30

表7 生化学的に検証された試料における乳糖不体耐症の有病率

集団(人種)	C/T <sub>13910</sub>						G/A <sub>22018</sub>		
	数	CC	CT	TT	GG	GA	AA		
1. フィンランド人									
ラクターゼ持続性	182	0	95	87	0	95	87		
ラクターゼ非持続性	116	116	0	0	110	6	0		
2. イタリア人									
ラクターゼ持続性	7	0	7	0	0	7	0		
ラクターゼ非持続性	23	23	0	0	22	1	0		
3. ドイツ人									
ラクターゼ持続性	0	0	0	0	0	0	0		
ラクターゼ非持続性	8	8	0	0	8	0	0		
4. ソマリア人									
ラクターゼ持続性	0	0	0	0	0	0	0		
ラクターゼ非持続性	42	42	0	0	42	0	0		
6. 韓国人									
ラクターゼ持続性	0	0	0	0	0	0	0		
ラクターゼ非持続性	23	23	0	0	23	0	0		
合計	401	212	102	87	205	109	87		

【 0 0 9 5 】

【 表 8 】

表8 様々な集団の試料における乳糖不耐症の有病率

集団	数	遺伝子型						ラクターゼ存続 性対立遺伝子の 有病率%	*	
		C/T13910			G/A22018					
		CC	CT	TT	GG	GA	AA			
韓国人	23	23	0	0	23	0	0	0	*	
フランス人	17	7	9	1	6	10	1	59	*	
バスク人	85	7	44	34	13	35	37	92	*	
南イタリア人	100	89	11	0	88	12	0	11	*	
ソマリア人	79	74	5	0	78	1	0	6		
ユタ州	92	7	33	52	7	30	55	92	*	10
アフリカ系アメリカ人	96	76	15	5	78	12	5	21	*	
モロッコ人	90	62	25	3	65	22	3	31	*	
サラウィー人(アフリカ)	57	29	26	2	28	26	3	49	*	
サーミ人	30	20	10	0	21	9	0	33	*	
チベット人	23	23	0	0	23	0	0	0		
東フィンランド人	571	108	287	176	107	288	176	81	*	
西フィンランド人	367	62	159	146	58	161	148	83	*	
フィン-ウゴル語族										
Xan	20	19	1	0	19	1	0	5		
Xm	20	19	1	0	19	1	0	5		
マンシ族	22	20	2	0	20	2	0	9		
コミ人	10	7	3	0	7	3	0	30		
Erza	30	17	10	3	19	9	2	43		20
Moksa	30	13	17	0	14	16	0	57	*	
Udmort	30	12	16	2	11	15	4	60	*	
パキスタン部族										
カラーシュ族	30	30	0	0	28	2	0	0		
ブルーショー族	30	29	1	0	27	3	0	3		
ハザラ族	14	13	1	0	11	3	0	7		
カシミール人	20	15	5	0	14	6	0	25		
マクラーニーバルーチ族	29	19	10	0	19	8	1	34		
ブラーフィー族	30	17	10	3	16	11	3	43		
マクラーニー族(黒色人種)	29	16	10	3	16	10	3	45		
パターン族	29	12	16	1	13	14	2	59	*	30
インド人	29	11	13	5	10	12	5	62	*	
合計	2032									

\*ラクターゼ存続性対立遺伝子の有病率は、報告されているラクターゼ存続性対立遺伝子の有病率(シムーンズ(Simooms), F. J. The geographic hypothesis and lactose malabsorption. A weighing of the evidence. Am. J. Dig. Dis. 23, 963-980(1978))と非常によく相関している。

#### 【図面の簡単な説明】

【0096】

【図1】研究されたフィンランド人成人型ラクターゼ欠乏症家系。黒い印は、ラクターゼ欠乏(hypolactasic)の個体を示し、アスタリスク(\*)は、試料が入手できなかったことを示し、疑問符(?)は未知の疾患状態を示す。は、SNPを同定するためのシーケンシングに用いられる個体を示す(表2)。

【図2】成人型ラクターゼ欠乏症遺伝子座の物理地図。BACクローンは、水平な線の上に表示されている。3つの遺伝子、LPH、MCM6およびDARSは、ブラックボックスより上にある、遺伝子の3'末端に対して向けられる先端を有する太い黒色の矢印によって示される。遺伝子座の微細なマッピングに使用される10個の多型マイクロサテライトマーカーの位置が示されている。水平な線にあるバックスラッシュは、コンティグ配列の配列におけるギャップを示す。マーカーD2S2169の位置は、上述されているように<sup>40</sup>、PACライブラリーから単離されたPAC106020でギャップを埋めることによ

って確認された。LPHの最初のATGの13.9 kbおよび22 kb 5'に位置するイントロン9および13におけるラクターゼ持続性表現型関連変異体の位置を含む、MCM6遺伝子の構成が示される。

【図3】密に連鎖する7個のマイクロサテライトマーカを用いた、フィンランド人の成人型ラクターゼ欠乏症家系に由来する持続性染色体の拡大されたハプロタイプ解析。祖先創始者の染色体を表すハプロタイプに陰影がつけられている。持続性染色体にも存在した非持続性染色体のハプロタイプのみが示されている。祖先の組み合わせに基づいて、成人型ラクターゼ欠乏症の遺伝子座は、マーカーLPH1およびAC3の間の47 kb区間に限定することができる。

【図4】ラクターゼ持続性に特異的である、その中のTがCによって置換される、-13910位にSNPを含むMCM6遺伝子のイントロン13の配列に含まれる配列(3220 bp)。前記位置は、小文字で示されている。この配列は配列番号1を指す。 10

【図5】ラクターゼ持続型配列に特異的である、その中のAがGによって置換される、-22018位にSNPを含むMCM6遺伝子のイントロン9の配列に含まれる配列(1295 bp)。前記位置は、小文字で示されている。この配列は配列番号2を指す。

【図6-1】-13910位にTを含む、MCM6遺伝子のラクターゼ持続型イントロン13の配列(3220 bp)。前記位置は、小文字で示されている。この配列は配列番号3を指す。

【図6-2】-13910位にTを含む、MCM6遺伝子のラクターゼ持続型イントロン13の配列(3220 bp)。前記位置は、小文字で示されている。この配列は配列番号3を指す。 20

【図7】-22018位にAを含む、MCM6遺伝子のラクターゼ持続型イントロン9の配列(1295 bp)。前記位置は、小文字で示される。この配列は配列番号4を指す。

【図8-1】ラクターゼ持続型配列に特異的である、その中のTがCによって置換される、-13910位にSNPを含むMCM6遺伝子のイントロン13の配列(3220 bp)。前記位置は、小文字で示される。この配列は配列番号5を指す。

【図8-2】ラクターゼ持続型配列に特異的である、その中のTがCによって置換される、-13910位にSNPを含むMCM6遺伝子のイントロン13の配列(3220 bp)。前記位置は、小文字で示される。この配列は配列番号5を指す。

【図9】ラクターゼ持続型配列に特異的である、その中のAがGによって置換される、-22018位にSNPを含むMCM6遺伝子のイントロン9の配列(1295 bp)。前記位置は、小文字で示される。この配列は配列番号6を指す。 30

【 図 1 】

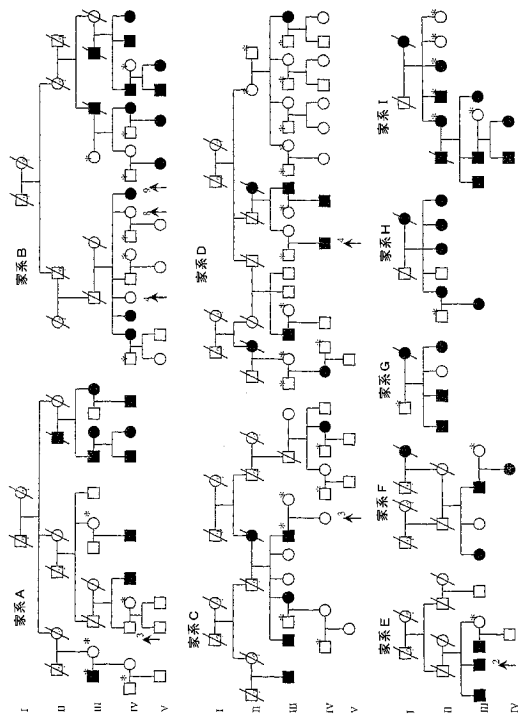


図 1

【 図 2 】

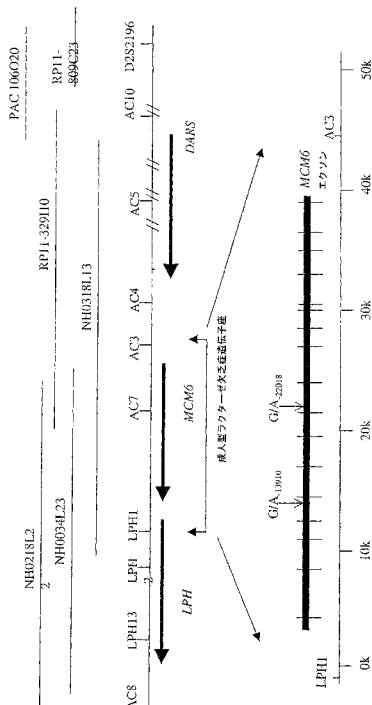


図 2

【 図 3 】

マーカー	ハプロタイプ										合計
	1	2	3	4	5	6	7	8	9	10	
LPH13	6	6	6	6	6	6	6	6	6	6	66
LPH12	4	4	4	4	4	4	4	4	4	4	44
LPH1	3	2	2	1	1	4	4	4	4	4	44
AC7	5	5	5	5	5	5	5	5	5	5	55
AC3	2	2	2	4	2	2	2	2	2	2	22
AC4	5	5	5	5	5	5	5	5	5	5	55
AC5	6	6	5	6	6	3	6	5	3	3	54
ラクター-1含有領域が正常型	20	4	1	2	2	1	1	1	1	1	33
ラクター-2含有領域が正常型	3	2	0	0	0	0	2	0	2	0	24

図 3

【 図 4 】

SEQ ID NO:1

ACCTTCATTTCAGGAAAAATGACCTTAGACCCCTACATCTACTACTAGTAGGCGCTCTGCCCTG  
 GCAATACAGATAAGATAATGTAGCCCTTGGCCCTCAAAGGAACCTCTCTCTTAGGTTGCA  
 TTTGTATAATGTTGATTTTGTAGATTGTTCTTGTAGCCCTGCATTCACGAGGATAGTCT

図 4

【 図 5 】

SEQ ID NO:2

TAGAACAATTTTACACTCTTCAGTATAAAGAAGTCAGAAATACCCCTACCCCTATCAGTAAA  
 GGCCCTTAAGTTACCATTAATAAAGATGTCCTTAAATAACAGCATTCACAGCTGGGCGCGGT  
 GGCTCACACCTTTGTCCCGTACTTTGGGAAGCCGAGGTGGGTGGATCACCTGAGTCCAG

図 5

【 図 6 - 1 】

SEQ ID NO:3

```

ATCAGAGTCACTTTGATATGATGAGAGCAGAGATAAACAAGATTTGTCATGTTTTAAT
CTTTGGTATGGGACATACAGATTCACATGCAAAATACATTTTTATGTAAGTGTAAAGTC
TCATACAGCAGTAAAGTATCTTCCCTTTAAGAGCCTTGTAAGCATTTGAGTGTAGTGT
AGAGGGAGCAGTACAGCTCAGTTTATAGAGTGCATTAAGAGCAGTAAATACCATTTAAT
ACCTTTCCATTCAGGAAAAATGACTTAGACCCACAAATGTAAGTAAAGCCTTCAGCGCTG
GCAATACAGATAAGATAAAGTACTCCCTGGCCCTCAAGAACTCTCCCTCTTAGGTGCA
TTTGTAATAATCTTTGATTTTATGATTTCCTTTGAGCCCTGCATTCACAGGATAGCT
AGTGGCTTTAAGCAGGTAAAGAGGGAGTAGTACGAAAGGGCATTCAAGCCTCCCATCTT
CGCTTCAACAAAGCAGCCCTGCCTTTTCTAGTTTATTAATAGGTTTGTATGTAAGCTG
GTCTTTGAAAGGGGGTTTGGCTTTTTPACAGCTGTACTGAGGTAAATTTAATAAAA
GGAAATATGATGGCAATGGTGTAGTTTTCACATACATCTTTGTGAATACCCAGCTCAAGA
TCCAAAACATTTCCATAAATTCAGAAAGTTCGCAACCCCTGCCTCTTTTCAGTCTTAGCC
CTCTCCCTGAAATTAACCATCTTCCGACTTCACATCTACTTCTTATCCACAGGTAA
TTTTTTGGCTTTTCCACATAAATTTCAAAATTCCTTGCAATGCTACTCTACTACTTAGAG
AAGTACTTTCACACCTAGCTTTTAAATTTCTTTGATTTCAACCAATTTTAGGGAACACC
TGTAGGGACAAAAATGAATGAGAGCCCTGCCTTCCATTTGCTGTAATCTGTTGGAA
CGAGACATGTTTAAATTAAGCATGTAATAAATAGAGTGGGTGATGAATAATCTATATA
CTAAATCCCCACGACACAGTTCACCTATGTAAACAACTGCATGTTGACCCCGAACC
TAAATATAAGTTGGAAATTAATAAATAACAGAGGGGAGATAGAGCATCAACAACAGAG
TCTGAGATGAATTAATTTATACCAAGAGGAGGAGGACTCAGGGAGGTGGCCAGCTT
TAAACCCAGCTCACTGAAGGGTGTGAGCAATTTGATAGGCAAGATACCTGGGACAGGT
CATCTTAAAOCCATGCTAACATTTGTACTTTTTCCTTTCATGTAAGTTCTCTGAAATGA
SCTGCATAAACTGCTACACCTTTAGGGCAGTCTCAATTTGATTTTATTTTGTCTAT
TTTTAAAAATTAGCTCTCAAAATAGCAGATTCACATGATATAAATAATATGCACATAAAT
TATATACAAATAATTTTTCGAACTGAAATTTAGTACTGCATATATATTAGAGCTATT
TCTGTCTATATGTCATATCTTCTTCCATTAATAAATACTTTTGTAGGCTTTCTCAC
TCTAAGATTAATAAATACTTCCCATTTTACTAGCTAGTTTCTAGTTTTCACAAA
CCATTTATGAAACATTCCTCTTTTGTACACCTGTGGCACTGCTTAAATATATATCT
TGTTGTAGTTAGGATCTCTTTTGGACTTTPCCATCTGTCTTCAATGAGCTTTTACAGCTCC
TCTTACATTTGGTACCATGATTTTAAATCTATGGGGCTTTGTAGTTTAAATGTAAGGGTA
GTCCAGGCGCATTTGCTCTTACAGCTGTAGGAACTTGAANTCAGCTTGTCTGTGTTTT
AAAGAAAAACCTGGATTTTATTTTACAGTATACATCTCTATATAATAACTTGAAGAT
TGAAGAACTCTATGATTTTTCCTATTCAGTAACTGATACCTTACAGTATAGGTTAGGTTGTA
CTACTATAAAATCTCAGCTGCATAAAACAAATTTTTTTTGTGTTGTCTACACATCCATTA
GGCTACAAAGGAGCCTCCTTGTAGCTGTACTCAGAGATTCAGGCTGATATAAAGGTTTGT
ATCTTGACATAGCTTTTCAATGATGACAGAACTGGGAGGAGAGAGGCTGGTGGACCATGTG
CTTCTCCCTCTTCTCAGAAATGACAACTACTCACAATTTCACTTCCAGAGAAATTA
ACATGGCCCTCTCTAAGCTCAAAATGAAAGAAAGTCCCTTCTACCAGGTCGCCAGAAAT
TAGAAGAGCAAACTTTGTAAACAGTTCTGAGTACCACAAATACCCGTTATCTTTCCACTT
AAGTCTTCTGCTTCACTGATGAGTCTTTAAACTTTCTCATATGTTTTTTCAGTGTCTT
TTTGTAAATTTCTGATATTTTATCATGTTTTGTTCTGACTGGGAGTAGCCTTTTTTCCA

```

図 6

【 図 6 - 2 】

```

TTTCATTTCTGGCTGGTTTCATTGCTGGTTGTTTTTTTGGTTTGGTTTTGTTTTGAGAT
GGACTCTCACTCTGTCGCCAGGCTGGAGTGCAGCTGCACAACTCCGGCTCACTGCAACC
TCTGCCCTCCAGGTTCAAGCGAATTTCTTCTTCTCAGCCCTCCCTGAGTAGCTGGGATACAG
GCATGTGCCACCAGCCAGCTAATTTTTTATATTTTTTATAGTAGAGTGGGTTCTCCAT
TTTGTGTCAGGCTGGTCTCAAACTCCAACTCTCAGGIGATCCGGCTGCTGCTCTCCAA
AGTGTGGGATATAGACATGAGCCACCTGCTGCCCTGCTTCTTATGGGATGTAATAG
CTTTGGATTCATGATATGATATATGTTTTTATTTTTCTACAAATACACTCTAGAGGT
GGAATGTTGGGCTCAGGTTTATGATGATGTTTTCTTCCCAACAGTGTCTCAATTTCTG
TTTTCAACCCCTGTAATGAGAGTGTCTTCCCTCTTTGCAACAACCTGATATTTGTCAGT
ATTTAGCCATTTCTGGTGAATTTATAGTGTATTTCTGTTGGTGAAGAGAGAAATGAG
AGAGGTTGTTTTGAGAAAAACAAAGCAACCTGTGAGAGTGTCTGTGTTTTGTGAGAAAA
CAAAAATACACTACTCTGTGATTTTCAATGGGAGAAATCTGTGTGGTATATCAAAAAAAG
TAGCTAATTTACTCTCATATTTAGTTTGTAGT

```

【 図 7 】

SEQ ID NO:4

```

TAAGAACATTTTACACTCTTTCAGTATAAAGAAGTCAAGATCCCCCTACCTTACAGTAAA
AATTAGCCATTTCTGGTGAATTTATAGTGTATTTCTGTTGGTGAAGAGAGAAATGAG
GGCTCACACCTTTCTCCAGTACTTTTCCGAGCCGAGGTTGGTGGATCACTGAGTCCAG
GAGTTCGAGACCCCTGGCCCAACTGGCCGAAACCCATTTCTCTACTAAAATAACAAA
AATTAGCCGGGATGTTGGTGGGGTGTCTTGTGTTCCAGCTACTCAAGAGGCTGAGTGGG
AGGATCACTGAGCCAGGAGGTGGAGGCTGCATGAGCCAGATTTGCCACTGCACCTCC
AGCTGGGTGACAGAGGGAGCTGTCTCAAAAAGAACAAAAGAAAACAAACCCAGCAT
CTTTTAGTAAATAATTCATAGTTTTCTTCTACTAGAAATTTAAATTTGTATGATGATCA
GCATGCTTGGCCAGCTGTGTTTTGCTGTACTAGTTTAGTCTGGTATGTTGATATAG
TTATAGGATATAAATCACTCTGAGTGCACAAAGGTTTTGATCTGAGATCAAGTACAGT
TAAGTGTATATTTTAGTTTTGCTCTTAGTTTTAAGTCAATTTGGTGGTTCTTCTTCCAGT
TTAGGATCTGTGAAATTTCTTCTCTAGAAAAGGAGTAAAGATCTTCTACTACCTATC
TTCTACTGTTTTGAGAAATAGAAGAGTCCCTGTGTTAGAGACTTTTGTGATTTACTGAT
AATTTTTCCATCTGAAAGACTGTTCTTTCTTTCCGTAAGTCTTGTCTGCTCCGAG
GCTGGAGTGCATGGTGCACCTTGGCTCACTGCACCTTGCCTCCGGGTTCAGAACAA
TCTCTCCGCTCAGCCCTCCGAGTATCTGGGATACAGGTCAGCCACACACACCTGAGCT
AATTTTGTATTTTTCACTAGAGCCGGGTTTTCAACCTGTTGGCCAGGCTGTCTGAACT
TTTGCCTCATGATCAAGCCACTCAGCTTTCCAAAGTCTGGGATACAGGTTGAGGTCGCC
CCCACTCCGGCTGTGTGTTTTTAAAGAGACGGGTCTCACTCTGTCACTTAACCTGG
AGTACAGTGGCAATCACTGGCTCACTGTAACCTCAAAATGCCGGCTTAGTGAAGCTTCA
TCTCCGCTTGGCTCCCAAGTGTGGGATACAAAGTGGAGCCATGCACTCCAGCTTGA
AGACAGCTTCTTAGGCTTGATTTGTTGGTTACAGG

```

図 7

【 図 8 - 1 】

SEQ ID NO:5

```

ATCAGAGTCACTTTGATATGATGAGAGCAGAGATAAACAAGATTTGTCATGTTTTAAT
CTTTGGTATGGGACATACAGATTCACATGCAAAATACATTTTTATGTAAGTGTAAAGTC
TCATACAGCAGTAAAGTATCTTCCCTTTAAGAGCCTTGGTAAGCAATTTGAGTGTAGTTGT
AGAGGGAGCAGTACAGCTATAGTTTATAGAGTGCATTAAGAGTAAAGTATCCATTTAAT
ACCTTTCTATTCAGGAAAAATGTAAGTACTAGCCCTCAAAATGTAAGTAAAGTAAAGT
GCATACAGATAGATAAATGATAGCCTGGCTCAAGAACTCTTCCCTTTAGGTTGCA
TTTTGTAATATGTTTTGATTTTATGATTTGTTTTTCCAGCCCTGCATTCACAGGATAGCT
AGTGGTATTAAGAGGTAAGAGGGGATTTAGCAAAAGGGCAATCAAGCCTCCCATCTT
CGCTTCAACAAAGCAGCCCTGGGTTTTTCTAGTTTATTAATAGGTTTGTATGTAAGGTC
GCTTTGAAAGGGGGTTTGGCTTTTTTTTACAGTGTACTGAGTATATTTATAAATAA
GGAAATATGATGGCAATGGTGTAGTTTTTTCACATACATCTTTGTGAATACCCAGCTCAAGA
TCCAAAACATTTCCATAAATTTAGAAAGTTCCAAAACCCCTGCTCTTTTCAGTCTTAGCC
CTCTTCCCTGAAATAACTGTTCCGACTTCAATCACTACTTTTATCCCAAGGTTAA
TTTTTTGGCTTTTTTCCACTAAATTTTCAAAATCTCTGATGTTGTTACTTTACTATTGAGC
AAGTACTTTACACTAGCTTATTTAATATCTTTGATTCACCAATATTTAGGGAACACC
TGTAGGGGACAAAAAATGAATGAGAGCCCTTCCCTTCCATTTGCTGCAATCTGTTGGGAA
CGAGACTGTTTTTAAATAGCATGTAATAAATAGAGTGGTGTAAATTAATCTATATA
CTAAATCCCAATGACACAGGTTTACCTATGTAACAAACCTGCATSTGTACCCCGAAAC
TAAATATAAGTTCGAAATTAATAAATAAAGAGAGGAGGATAGAGCATCAACAACAGAG
TGTGAGATGAATTTACTTTTATTCACAAAGAGGAGGAGGACTCAGGGAGGTTCCGCACTT
TAAACCCAGTCACTGAAAGGCTGCAGAAATTTGGATAGGCAAGATACCTGGGACAGGT
CATTTCAAAACCATGCTAAGCATTTGTAATTTTTTTTTCATTTGATAGTTCCTGAAATGA
GTGCAATAAACTGGTACATGCTTTAGGGCAGCTCTAAATGATTTTTATTTTTGTTCTAT
TTTTTAAATTTAGTCTTCAAATAGCAGATTCACATGATATAAATAATATCCACATAAAT
TATATACACAAATATTTTTCTGAAAGAAATTTAGTATCTGCATATATTTAAGAGCTATT
TCTGCTCATATGTTCAATAATCTTCAATCAAAATAAATAAATTTTGTAGGCTTTCTCAC
TCTAAGATATAAATAATTTCCCAATTTTACCTAGCCAGTTTCTAGTGTGTTCCAAA
CCATTTATGAAACATTTCACTTTTGTGACACTGTTTGGCACTGCTTAAATATAATTTCT
TGTGTGTTTAGGATCTCTTTTGTGACTTTTCCACTTCTGTTCTGTTAGGCTCTTATCAGCTCC
TCTTACATTTGTTAGCAGTGTCTTAAATCTATGGGGCTTTGTTAGTTTAAATGTTAGGGCTA
GHTCCAGCCCATGTTTCTTCTTACTGCTTGTAGGACTTAGAAATCACTGCTGCTGCTTPTT
AAGAAAAACCTGGATTTTTTTTATCAGTATAAATCTTAAATTTATTTAACTTGAAGAAAT
TGAAACATCTATATTTTTTCTTATCCAGTAAGTACTCTTAAATAGGTTAGGTTGTA
TACTATAAATCTCAGCTGCATAAACAATTTTTTTTTTGTGCTGCTACACATTA
GGTCACTAAGGGACCTCAGCTTGTCAAGTACTCAGAGATTCAGGCTGATATAAAGGTTTGT
ATCTTGACATACGCTTTCTGATGACAGAAAGAGGAGGAGGAGGTTGGTGAOCCATGCT
CTTCTCCCTCTTCTATCCAGAAATGACACATCTCACAATTTCAATCTGCCAGAGAAATTA
ACATGGCCCTTCCATAAATGAATGACAGAAATGCTTCTTACCAGGTTCCCAAGCTT
TAGAAGAGCAAACTTTGTAAGAGCTTTGAGTACCAAAATACCCGTTACTTTTCCACTT
AAGTCTTCTGCTTCACTGATGAGTCTTTAAACTTTCTCATATGTTTTTTCAGTGTCTT
TGTGTGAATTTCTGATATTTTATCATGTTTTGTTCTGACTGGGAGTAGCCTTTTTTCCA

```

図 8

【 図 8 - 2 】

```

TTTCATTTCTGGCTGGTTTCATTGCTGGTTGTTTTTTTGGTTTGGTTTTGTTTTGAGAT
GGACTCTCACTCTGTCGCCAGGCTGGAGTGCAGTGCACAACTCCGGCTCACTGCAACC
TCTGCCCTCCAGGTTCAAGCGAATTTCTTCTTCTCAGCCCTCCCTGAGTAGCTGGGATACAG
GCATGTGCCACCATGCCCAGCTAATTTTTTATATTTTTTATAGTAGATGGGTTTCTCCAT
GTTGGTCAGGCTGGTCTCAAACCTCCAAATCTCAGGTTGATCCGCTGCTCTGCTCTCCAA
AGCTCTGGGATTTATAGACATGAGCCACCCTGCTGCTTCAAGGATTTCTTATAGT
TCTTTGGATTCATATGATATGATATATGTTTTATATTTCTCAAGTACATACCTAGGAT
GGAAATGTTTGGGTCATAGGTTAATGCTGTTTTTTCTGCAACACAGTGTGTCAATTTCTG
TTTTACCCTGTGAAATGAGAGTTTCTTACTCTTCTGACACACTGATATTTGTGAGT
ATTTTAGCCATTTCTGGTGAATTTATAGTGTCTATTTCTGCTAGATTTAAGAGAGAAAG
AGAGGTTGTTCTGCAAAACCAAAAGCAACCTGTAGAGAGTGTGTGTTTTGTTGAGAAA
CCAAAATACACTACTCTGTGATTTTCAATGGGAGAAAATCTGTTGTGATATCAAAAAAG
TAGAAATAACTACTCTCATATTTAGTTTGTGGTTAGGTT

```

図 9

【 図 9 】

SEQ ID NO:6

```

TAAGAACATTTTACACTCTTTCAGTATAAAGAAGTCAAGATACCCCTACCTTACAGTAAA
GGCCATAAAGTTACCATTAAGAAGATGTCCTTAAAAAACAGCATTTCTCAGCTGGGCGCGG
GGCTCACACCTTTGTCCTCAGTACTTTGGAAAGCCGAGGTTGGGTTGATCACCAGGATCAG
GAGTTCGAGACCCAGCCTGGCCAACTAGGCAAAACCAATTTCTCTACTAAAATAACAAA
AATTAGCCGGGATGTTGGCGGGTGTCTTGTGTTCCAGCTACCTAGAGAGCTGAGGTTGG
AGGATCACTGAGCCAGGAGGTTGGAGGCTGCATTTTCTGCAAGATTTGTGCCACTGACTCC
AGCCTGGGTGACAGCGGAGCTGTCTCAAAAATAAACAATAAATAAATAAACCAGCAT
TCTTACTAATAATTTCACTACTTCTTCTTCACTAGATTTAATAATTTGATTTGATGATCA
GCATGCTCTGAGCAGGTGTGTTGCTGTTACTAGTTTAGATCGGTAGATGTGTATATAAG
TTATAGTATAAATAAATCTGATGTTGACAAAGGTTTGTAGTTTAGTACAGTACAGTACAG
TAAGTGTATATTTTTTATGCTCTTACTGTTTTTAAAGTCAATTTGTGTTCTTCTACTGCT
TAGGACTGTGTAATTTCTCTCTTGAAAAAGGGATTAAGAATCTTCACTTACTTCTTCT
TTCATGTTGTTGGAAATAGAGAGTCCCTGCTGGTAGCAGCTTTGTAGTTTACTTTGT
AATTTCCATCTGAAAGACTGTTCTTGTTTTTGTTGATGAGCTTCTGCTCTGCTGCCAG
GCTGGAGTGCAGTGGTGCACACTTGGCTCACTGCAACCTTCCCTCCCTCCCGGTTCAAGCA
ATTTCTCTGCTGAGCCTCCGAGTATCTGGGATACAGGTCAGCAGCACACCAACCTGGCT
AATTTTGTATTTTCAATTTAGTAGGAGGAGGCTTCCACCATTTGTGGCGAGGAGGAGT
CTTCACTCATGATGATGCCCACTCAGCTTCCBAAGTCTCTGCAATACAGTGTGAGGCT
CCACACTCCGCTGCTGTTTTTAAAGAGCAAGGCTTCACTCTGACCTAACCTGG
AGTACAGTGGCAATCACTGGCTCACTGTAACCTCAAAATGCCGGCTTAGTGAAGCTTCT
TCTGCTTGGCTCCCAAGTGTGGGATACAAAGTGGAGCCATGCACTCCAGCTTGA
AGACAGCTTCTTAGGCTTGATTTGTTGGTTACAGG

```

## 【国際公開パンフレット】

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau(43) International Publication Date  
20 February 2003 (20.02.2003)

PCT

(10) International Publication Number  
WO 03/014342 A2

- (51) International Patent Classification: C12N 15/00 P.O.Box 281, 00029 Hus (FI). TERWILLIGER, Joseph [US/US]; Columbia University, Department of Psychiatry and, Columbia Genome Center, 1150 St. Nicholas Avenue, New York, NY 10032 (US).
- (21) International Application Number: PCT/JP02/08963
- (22) International Filing Date: 9 August 2002 (09.08.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
01119377.8 10 August 2001 (10.08.2001) JP  
01119328.6 14 August 2001 (14.08.2001) JP  
60/315.955 31 August 2001 (31.08.2001) US
- (74) Agent: VOSSIUS & PARTNER; Siebertstrasse 4, 81675 Munich (DE).
- (81) Designated States (national): 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, PI, 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.
- (71) Applicant (for all designated States except US): NATIONAL PUBLIC HEALTH INSTITUTE [FI/FI]; Mannerheimintie 166, FIN-00500 Helsinki (FI).
- (72) Inventors; and  
(75) Inventors/Applicants (for US only): PELTONEN, Leena [FI/FI]; National Public Health Institute, Biomedicum Dept of Molecular Medicine, Haartmaninkatu 8, 00251 Helsinki (FI). ENATTAH, Nabil [LY/ET]; National Public Health Institute, Biomedicum, Dept of Molecular Medicine, Haartmaninkatu 8, 00251 Helsinki (FI). JÄRVELÄ, Irma [FI/FI]; National Public Health Institute, Biomedicum Dept of Molecular Medicine, Haartmaninkatu 8, 00251 Helsinki (FI). SAHL, Timo [FI/FI]; Finnish Defence Forces, P.O.Box 919, 00131 Helsinki (FI). SAVILAHTI, Erkki [FI/FI]; Hospital for Children and Adolescents,
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:  
— without international search report and to be republished upon receipt of that report
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: IDENTIFICATION OF A DNA VARIANT ASSOCIATED WITH ADULT TYPE HYPOLACTASIA

(57) Abstract: The present invention relates to a nucleic acid molecule comprising a 5' portion of an intestinal lactase-pherolizine hydrolase (LPH) gene contributing to or indicative of the adult-type hypolactasia wherein said nucleic acid molecule is selected from the group consisting of (a) a nucleic acid molecule having or comprising the nucleic acid sequence of SEQ ID NO: 1, the sequence of SEQ ID NO:1 is also depicted in the Fig. 4 and comprised in the sequence as depicted in the Fig. 8; (b) a nucleic acid molecule having or comprising the nucleic acid sequence of SEQ ID NO: 2, the sequence of SEQ ID NO:2 is also depicted in Fig.5 and comprised in the sequence as depicted in the Fig. 9; (c) a nucleic acid molecule of at least 20 nucleotides the complementary strand of which hybridizes under stringent conditions to the nucleic acid molecule of (a) or (b), wherein said polynucleotide/nucleic acid molecule has at a position corresponding to position -13910 5' from the LPH gene a cytosine residue; and (d) a nucleic acid molecule of at least 20 nucleotides the complementary strand of which hybridizes under stringent conditions to the nucleic acid molecule of (a) or (b), wherein said polynucleotide/nucleic acid molecule has at a position corresponding to position -22018 5' from the LPH gene a guanine residue. The present invention further relates to methods for testing for the presence of or predisposition to adult-type hypolactasia that are based on the analysis of an SNP contained in the above recited nucleic acid molecule. Additionally, the present invention relates to diagnostic composition and kit useful in the detection of the presence of or predisposition to adult-type hypolactasia.

WO 03/014342 A2

WO 03/014342

PCT/EP02/08963

**Identification of a DNA variant associated with adult type hypolactasia**

5 The present invention relates to a nucleic acid molecule comprising a 5' portion of an intestinal lactase-phlorizine hydrolase (LPH) gene contributing to or indicative of the adult-type hypolactasia wherein said nucleic acid molecule is selected from the group consisting of (a) a nucleic acid molecule having or comprising the nucleic acid sequence of SEQ ID NO: 1, the sequence of SEQ ID NO:1 is also depicted in Fig. 4  
10 and comprised in the sequence as depicted in Fig. 8; (b) a nucleic acid molecule having or comprising the nucleic acid sequence of SEQ ID NO: 2, the sequence of SEQ ID NO:2 is also depicted in Fig. 5 and comprised in the sequence as depicted in Fig. 9; (c) a nucleic acid molecule of at least 20 nucleotides the complementary strand of which hybridizes under stringent conditions to the nucleic acid molecule of  
15 (a) or (b), wherein said polynucleotide/nucleic acid molecule has at a position corresponding to position -13910 5' from the LPH gene a cytosine residue; and (d) a nucleic acid molecule of at least 20 nucleotides the complementary strand of which hybridizes under stringent conditions to the nucleic acid molecule of (a) or (b), wherein said polynucleotide/nucleic acid molecule has at a position corresponding to  
20 position -22018 5' from the LPH gene a guanine residue. The present invention further relates to methods for testing for the presence of or predisposition to adult-type hypolactasia that are based on the analysis of an SNP contained in the above recited nucleic acid molecule. Additionally, the present invention relates to diagnostic composition and kit useful in the detection of the presence of or predisposition to  
25 adult-type hypolactasia.

A variety of documents is cited throughout this specification. The disclosure content of these documents, including manufacturer's manuals and catalogues, is herewith incorporated by reference.

30

Lactase-phlorizin hydrolase enzyme (LPH), which is exclusively expressed by intestinal epithelial cells, hydrolyses lactose, sugar of milk, into glucose and galactose<sup>1</sup>. The expression of the LPH enzyme dramatically declines to very low

WO 03/014342

PCT/EP02/08963

2

levels at the weaning period in mammals when lactose is no longer an essential part of the diet. In humans, the condition known as adult-type hypolactasia or lactase non-persistence, affects most populations and severely limits the use of fresh milk among adults due to lactose intolerance. The age of onset of lactase non-persistence status varies between populations, ranging from 1-2 years of age among the Thais to 10-20 years of age among the Finns<sup>2,3</sup>. However, in Northern European and a few other ethnic groups, LPH activity persists throughout life in the majority of adults, a condition known as lactase persistence. The phenotype lactase persistence/non-persistence has been shown to be genetically determined, the persistent status being dominant over the non-persistent status<sup>4,5</sup>.

The state of the art diagnosis of adult-type hypolactasia is based on the lactose tolerance test (LTT). After overnight fasting (10 hours), 1g/kg of lactose is given as a 12.5% solution, the maximum dose being 50g. Capillary blood samples are taken before and 20 and 30 min after lactose ingestion. The glucose concentration is determined by the glucose oxidase method (Hjelm and de Verdier 1963). Abdominal symptoms on the day of LTT are noted. A maximum rise in blood glucose concentration of 1.1 mmol/l or more was taken as a sign of lactose malabsorption (Gudman-Hoyer and Harnum 1968, Jussila 1970, Sahi 1972). LTT contains a 10% risk for false positive and negative diagnoses, i.e. the sensitivity and specificity of LTT is about 90% (Isokoski et al. 1972, Newcomer et al. 1975, Sahi 1983).

The accuracy of LTT can be improved by giving 0.3 g/kg ethanol that inhibits the metabolism of galactose in the liver (Tygstrup and Lundqvist 1962) and 15 min later 1g/kg lactose as 12.5% solution.

Children with maximum rises of less than 0.2mg/100ml in the first or repeated LTT have been sent for small-intestinal biopsy that is taken through gastroscopy. This is an invasive procedure that needs expertise and is usually performed at university hospitals by specialists in gastroenterology only. Biopsy samples are examined with a dissection microscope and histologically, and the mucosal maltase, sucrase and lactase activities are determined (Launiala et al. 1964). The diagnosis of hypolactasia in children is justified if the histology of the intestinal biopsy is normal and lactase activity is less than 20U/g protein and lactase/sucrase ratio less than 0.30, or in the

WO 03/014342

PCT/EP02/08963

3

LTT with ethanol administration a maximum rise in blood glucose concentration of less than 20mg/100ml and in galactose concentration of 5mg/100 ml or less (Sahi et al, 1972) is demonstrated. As described above, the current methods to diagnose adult-type hypolactasia are laborious. LTT is inexact and therefore, an invasive  
5 procedure, gastroscopy is needed before the diagnosis can be ascertained. Since adult-type hypolactasia is very common and the major cause of nonspecific abdominal symptoms (in one third of patients complaining stomach pain), there is a clear need to improve the diagnostics of this common health problem.

10 Yet, so far no biochemical test that is easy to handle and, at the same time, provides quick and accurate results has been developed. Elucidation of the cause of the disease on the genomic DNA/ expression level has equally been unsuccessful. Thus, the sequencing of the coding and promoter regions of the LPH gene in adults has revealed no DNA-variations which correlate with lactase persistence/ non-  
15 persistence, nor has evidence emerged of splice variants or mRNA editing variants associated with this trait<sup>7-9</sup>. Previous studies have shown that the lactase persistence/non-persistence trait is possibly controlled by cis-acting element(s) residing within or adjacent to the lactase gene, and strong linkage disequilibrium (LD) has been observed across the 70 kb haplotype spanning the lactase gene<sup>9,10</sup>.  
20 Several studies report evidence that the main control of the LPH gene expression operates at the level of transcription regulation<sup>11-13</sup>. However, it has been suggested that variation influencing both transcriptional and posttranscriptional control of expression of the LPH gene may be involved in the etiology of adult-type hypolactasia<sup>14-15</sup>.

25

In view of the above, the technical problem underlying the present invention was to provide means and methods that allow for an accurate and convenient diagnosis of adult-type hypolactasia or of a predisposition to this disease.

30 The solution to said technical problem is achieved by the embodiments characterized in the claims.

Thus, the present invention relates to a nucleic acid molecule comprising a 5' portion

WO 03/014342

PCT/EP02/08963

4

of an intestinal lactase-phlorizine hydrolase (LPH) gene contributing to or indicative of adult-type hypolactasia wherein said nucleic acid molecule is selected from the group consisting of (a) a nucleic acid molecule having or comprising the nucleic acid sequence of SEQ ID NO: 1, the sequence of SEQ ID NO:1 is also depicted in Fig. 4 and comprised in the sequence as depicted in Fig. 8; (b) a nucleic acid molecule having or comprising the nucleic acid sequence of SEQ ID NO: 2, the sequence of SEQ ID NO:2 is also as depicted in Fig. 5 and comprised in the sequence as depicted in Fig. 9; (c) a nucleic acid molecule of at least 20 nucleotides the complementary strand of which hybridizes under stringent conditions to the nucleic acid molecule of (a) or (b), wherein said polynucleotide/nucleic acid molecule has at a position corresponding to position -13910 5' from the LPH gene a cytosine residue; and (d) a nucleic acid molecule of at least 20 nucleotides the complementary strand of which hybridizes under stringent conditions to the nucleic acid molecule of (a) or (b), wherein said polynucleotide/nucleic acid molecule has at a position corresponding to position -22018 5' from the LPH gene a guanine residue.

In accordance with the invention, the term "intestinal lactase-phlorizine hydrolase (LPH) gene" denotes a gene that encodes an enzyme having the activity of hydrolyzing lactose into its components glucose and galactose. The enzyme is characterized by E.C. 3.2.1.23.62.

The term "adult-type hypolactasia" refers to a condition also known as lactose intolerance, which is an autosomal recessive condition resulting from the "physiological" decline of the lactase-phlorizin hydrolase (LPH) enzyme activity in intestinal cells in a significant proportion of the global population.

The term "contributing to or indicative of adult-type hypolactasia", refers to the fact that the SNPs and thus the corresponding nucleic acid molecules found are indicative of the condition and possibly also causative therefore. Accordingly, this term necessarily requires that the recited 5' position is indicative of the condition. Said term, on the other hand, does not necessarily require that the 5' portion is causative or contributes to the condition. Yet, said term does not exclude a causative or contributory role of either or both SNPs.

WO 03/014342

PCT/EP02/08963

5

The term "which hybridizes under stringent conditions" refers to hybridization conditions that are well known to or can be established by the person skilled in the art according to conventional protocols. The term most advantageously refers to highly stringent conditions. Appropriate stringent conditions for each sequence may be established on the basis of well-known parameters such as temperature, composition of the nucleic acid molecules, salt conditions etc.: see, for example, Sambrook et al., "Molecular Cloning, A Laboratory Manual"; CSH Press, Cold Spring Harbor, 1989 or Higgins and Hames (eds.), "Nucleic acid hybridization, a practical approach", IRL Press, Oxford 1985 (reference 54), see in particular the chapter "Hybridization Strategy" by Britten & Davidson, 3 to 15. Typical (highly stringent) conditions comprise hybridization at 65°C in 0.5xSSC and 0.1% SDS or hybridization at 42°C in 50% formamide, 4xSSC and 0.1% SDS. Hybridization is usually followed by washing to remove unspecific signal. Washing conditions include conditions such as 65°C, 0.2xSSC and 0.1% SDS or 2xSSC and 0.1% SDS or 0.3xSSC and 0.1% SDS at 25°C – 65°C.

As disclosed herein above, the present invention also relates to a hybridizing nucleic acid molecules of at least 20 nucleotides; see (c) and (d) herein above. Yet, the present invention also relates to a nucleic acid molecule of at least 50, at least 100, at least 150, or at least 200 nucleotides. Preferably, said hybridizing fragments comprise at least 25, at least 50, or at least 75 nucleotides, at least 100 nucleotides, 5' and 3' of the position –13910 as defined in (c) or of position –22018 as defined in (d) herein above.

25

The term "nucleic acid molecule" refers both to naturally and non-naturally occurring nucleic acid molecules. Non-naturally occurring nucleic acid molecules include cDNA as well as derivatives such as PNA.

30 The term "nucleic acid molecule [...] comprising the nucleic acid sequence of SEQ ID NO:" throughout this specification refers to nucleic acid molecules that are at least 1 nucleotide longer than the nucleic acid molecule specified by the SEQ ID NO. At the same time, these nucleic acid molecules extend, at a maximum, 30000 nucleotides

WO 03/014342

PCT/EP02/08963

6

over the 5' and/or 3' end of the nucleic acid molecule of the invention specified e.g. by the SEQ ID NO: 2 or 1, 3 or 4.

Surprisingly, it was found in accordance with the present invention that the two  
5 hypolactasia-associated variants locate at a considerable distance from the LPH  
gene, positioned in different introns of the *MCM6* gene. *MCM6* is a member of a  
gene family (*MCM 2-7*), required for the initiation of DNA replication ensuring that it  
takes place only once during the cell cycle<sup>31</sup>. *MCM6*, unlike *LPH*, is not restricted in  
its tissue distribution and there is no correlation in the levels of *MCM6* and *LPH*  
10 transcripts<sup>18</sup>. These findings would suggest that these two genes do not share any  
functionally significant cis-acting elements providing tissue specificity or  
developmental regulation<sup>18</sup>. Most probably the identified variants have different  
functional significance for the expression of the *LPH* and *MCM6* genes. Further  
surprisingly, based on complete association to hypolactasia they (or one of them) are  
15 associated to age-dependent down regulation of the transcript level of the *LPH* gene  
in the intestinal epithelium but have little or no effect on the transcription of the  
*MCM6*.

Experimentally, using linkage, allelic association and extended haplotype analysis  
20 carried out in nine extended Finnish families the adult-type hypolactasia locus was  
restricted to a 47 kb interval on 2q21. The sequence analysis of the region revealed a  
single nucleotide polymorphism (SNP), C/T-13910 that completely cosegregated with  
adult-type hypolactasia in all Finnish families and in a sample set of 236 individuals  
from four different populations. Another SNP G/A-22018 residing 8 kb telomeric from  
25 C/T-13910 was associated with the trait in all but 7 cases. The prevalence of C/T-  
13910 SNP in 1047 DNA samples reflected the reported prevalence of adult-type  
hypolactasia in three different populations providing additional evidence for its  
importance for the trait.

30 The surprising finding referred to above for the first time allows the establishment of  
test systems that are based on the molecular analysis of the recited single nucleotide  
polymorphisms upstream of the LPH gene. Whereas both SNPs provide for a solid  
basis for the diagnosis of or the diagnosis of a predisposition to adult-type

WO 03/014342

PCT/EP02/08963

7

hypolactasia, it is preferred that the nucleotide position -13910 is analyzed, either alone or in combination with nucleotide position -22018. This is because the SNP at position -13910 was associated in 100% of the analysed cases with the disease whereas the SNP at position -22018 was associated in only 98% of all cases with adult-type hypolactasia. Nevertheless, analyses of nucleotide position -22018 alone will usually also provide a sound basis for a diagnosis of a predisposition to adult-type hypolactasia.

Due to the abundance of established methods for assessing for the presence of SNPs, it is now possible to conveniently, in a short amount of time, at low cost, with high accuracy and without significant trouble for the person under investigation, diagnose a genetic predisposition to adult-type hypolactasia.

The invention further relates to a nucleic acid molecule comprising a 5' portion of an intestinal lactase-phlorizine hydrolase (LPH) gene wherein said nucleic acid molecule is selected from the group consisting of (a) a nucleic acid molecule having or comprising the nucleic acid sequence of SEQ ID NO:3, the sequence of SEQ ID NO:3 is also depicted in Fig. 6; (b) a nucleic acid molecule having or comprising the nucleic acid sequence of SEQ ID NO:4, the sequence of SEQ ID NO:4 is also depicted in Fig. 7; (c) a nucleic acid molecule the complementary strand of which hybridizes under stringent conditions to the nucleic acid molecule of (a) or (b), wherein said polynucleotide/nucleic acid molecule has at a position corresponding to position -13910 of the LPH gene a thymidine residue; and (d) a nucleic acid molecule the complementary strand of which hybridizes under stringent conditions to the nucleic acid molecule of (a) or (b), wherein said polynucleotide/nucleic acid molecule has at a position corresponding to position -22018 of the LPH gene a adenosine residue.

This embodiment of the present invention may conveniently be used to demonstrate that a person does not suffer from adult-type hypolactasia and has no predisposition therefor. Further, this nucleic acid molecule reflecting the "wild-type" situation of the position -13910 or -22018 upstream of the LPH gene may be used as a control means in experiments where a predisposition to adult-type hypolactasia is tested for.

WO 03/014342

PCT/EP02/08963

8

For testing, methods as described throughout this specification may be used.

In a preferred embodiment of the invention the nucleic acid molecule is genomic DNA.

- 5 This preferred embodiment of the invention reflects the fact that usually the analysis would be carried out on the basis of genomic DNA from body fluid, cells or tissue isolated from the person under investigation.

- 10 In a further preferred embodiment of the nucleic acid molecule of the invention said genomic DNA is part of a gene.

In accordance with the invention, it is preferred that at least one of the introns of the MCM6 gene harboring position -13910 or position -22018 relative to the LPH gene is analyzed.

- 15 In addition, the invention relates to a fragment of the nucleic acid molecule as described herein above having at least 14 nucleotides wherein said fragment comprises nucleotide position -13910 or nucleotide position -22018 (upstream) of the LPH gene.

The fragment of the invention may be of natural as well as of (semi)synthetic origin.

- 20 Thus, the fragment may, for example, be a nucleic acid molecule that has been synthesized according to conventional protocols of organic chemistry. Importantly, the nucleic acid fragment of the invention comprises nucleotide position -13910 or nucleotide position -22018 upstream of the LPH gene. In these positions, the fragment may have either the wild-type nucleotide or the nucleotide contributing to or  
25 indicative of adult-type hypolactasia (also referred to as the "mutant" sequence). Consequently, the fragment of the invention may be used, for example, in assays differentiating between the wild-type and the mutant sequence.

- 30 It is further preferred that the fragment of the invention consists of at least 17 nucleotides, more preferred at least 21 nucleotides, and most preferred at least 25 nucleotides such as 30 nucleotides.

Furthermore, the invention relates to a nucleic acid molecule which is complementary to the nucleic acid molecule as described herein above.

WO 03/014342

PCT/EP02/08963

9

This embodiment of the invention comprising at least 14 nucleotides and covering at least position -13910 or position -22018 of the sequence upstream of the LPH gene is particularly useful in the analysis of the genetic setup in the recited positions in hybridization assays. Thus, for example, a 15mer exactly complementary either to the wild-type sequence (i.e. a T in position -13910 or an A in position -22018) or to the variants contributing to or indicative of adult-type hypolactasia (i.e. a C in position -13910 or a G in position -22018) may be used to differentiate between the polymorphic variants. This is because a nucleic acid molecule labeled with a detectable label not exactly complementary to the DNA in the analyzed sample will not give rise to a detectable signal, if appropriate hybridization and washing conditions are chosen.

In this regard, it is important to note that the nucleic acid molecule of the invention, the fragment thereof as well as the complementary nucleic acid molecule may be detectably labeled. Detectable labels include radioactive labels such as  $^3\text{H}$ , or  $^{32}\text{P}$  or fluorescent labels. Labeling of nucleic acids is well understood in the art and described, for example, in Sambrook et al., loc. cit..

In addition, the invention relates to a vector comprising the nucleic acid molecule as described herein above. The vector of the invention may either contain a nucleic acid molecule comprising the wild-type sequence(s) or it may contain a nucleic acid molecule comprising the mutant sequence(s).

The vectors may particularly be plasmids, cosmids, viruses or bacteriophages used conventionally in genetic engineering that comprise the nucleic acid molecule of the invention. Preferably, said vector is an expression vector and/or a gene transfer or targeting vector. Expression vectors derived from viruses such as retroviruses, vaccinia virus, adeno-associated virus, herpes viruses, or bovine papilloma virus, may be used for delivery of the nucleic acid molecule of the invention into targeted cell population. Methods which are well known to those skilled in the art can be used to construct recombinant viral vectors; see, for example, the techniques described in Sambrook et al., loc. cit. and Ausubel et al., *Current Protocols in Molecular Biology*, Green Publishing Associates and Wiley Interscience, N.Y. (1989). Alternatively, the nucleic acid molecules and vectors of the invention can be reconstituted into

WO 03/014342

PCT/EP02/08963

10

liposomes for delivery to target cells. The vectors containing the nucleic acid molecules of the invention can be transferred into the host cell by well-known methods, which vary depending on the type of cellular host. For example, calcium chloride transfection is commonly utilized for prokaryotic cells, whereas, e.g., calcium phosphate or DEAE-Dextran mediated transfection or electroporation may be used for other cellular hosts; see Sambrook, supra.

Such vectors may comprise further genes such as marker genes which allow for the selection of said vector in a suitable host cell and under suitable conditions. Preferably, the nucleic acid molecule of the invention is operatively linked to expression control sequences allowing expression in prokaryotic or eukaryotic cells. Expression of said polynucleotide comprises transcription of the polynucleotide into a translatable mRNA. Regulatory elements ensuring expression in eukaryotic cells, preferably mammalian cells, are well known to those skilled in the art. They usually comprise regulatory sequences ensuring initiation of transcription and, optionally, a poly-A signal ensuring termination of transcription and stabilization of the transcript, and/or an intron further enhancing expression of said polynucleotide. Additional regulatory elements may include transcriptional as well as translational enhancers, and/or naturally-associated or heterologous promoter regions. Possible regulatory elements permitting expression in prokaryotic host cells comprise, e.g., the PL, lac, trp or tac promoter in *E. coli*, and examples for regulatory elements permitting expression in eukaryotic host cells are the AOX1 or GAL1 promoter in yeast or the CMV-, SV40-, RSV-promoter (Rous sarcoma virus), CMV-enhancer, SV40-enhancer or a globin intron in mammalian and other animal cells. Beside elements which are responsible for the initiation of transcription such regulatory elements may also comprise transcription termination signals, such as the SV40-poly-A site or the tk-poly-A site, downstream of the polynucleotide. Optionally, the heterologous sequence can encode a fusion protein including an C- or N-terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product. In this context, suitable expression vectors are known in the art such as Okayama-Berg cDNA expression vector pcDV1 (Pharmacia), pCDM8, pRc/CMV, pcDNA1, pcDNA3, the Echo™ Cloning System (Invitrogen), pSPORT1 (GIBCO BRL) or pRevTet-On/pRevTet-Off or pCI (Promega). Preferably, the expression control sequences will be eukaryotic promoter systems in

WO 03/014342

PCT/EP02/08963

11

vectors capable of transforming or transfecting eukaryotic host cells, but control sequences for prokaryotic hosts may also be used.

As mentioned above, the vector of the present invention may also be a gene transfer or targeting vector. Gene therapy, which is based on introducing therapeutic genes into cells by ex-vivo or in-vivo techniques is one of the most important applications of gene transfer. Suitable vectors and methods for in-vitro or in-vivo gene therapy are described in the literature and are known to the person skilled in the art; see, e.g., Giordano, *Nature Medicine* 2 (1996), 534-539; Schaper, *Circ. Res.* 79 (1996), 911-919; Anderson, *Science* 256 (1992), 808-813; Isner, *Lancet* 348 (1996), 370-374; Muhlhauser, *Circ. Res.* 77 (1995), 1077-1086; Wang, *Nature Medicine* 2 (1996), 714-716; WO94/29469; WO 97/00957, Schaper, *Current Opinion in Biotechnology* 7 (1996), 635-640, or Kay et al. (2001) *Nature Medicine*, 7, 33-40 and references cited therein. The polynucleotides and vectors of the invention may be designed for direct introduction or for introduction via liposomes, or viral vectors (e.g. adenoviral, retroviral) into the cell. Preferably, said cell is a germ line cell, embryonic cell, or egg cell or derived therefrom, most preferably said cell is a stem cell. Gene therapy is envisaged with the wild-type nucleic acid molecule only.

The invention as well relates to a primer or primer pair, wherein the primer or primer pair hybridizes under (highly) stringent conditions to the nucleic acid as described herein above comprising nucleotide position -13910 or -22018 of the LPH gene or to the complementary strand thereof.

Preferably, the primers of the invention have a length of at least 14 nucleotides such as 17 or 21 nucleotides. It is further preferred that the primers have a maximum length of 24 nucleotides. Hybridization or lack of hybridization of a primer under appropriate conditions to a genome sequence comprising either position -13910 or position -22018 coupled with an appropriate detection method such as an elongation reaction or an amplification reaction may be used to differentiate between the polymorphic variants and then draw conclusions with regard to, e.g., the predisposition of the person under investigation for adult-type hypolactasia. The present invention envisages two types of primers/primer pairs. One type hybridizes to a sequence comprising the mutant sequence. In other words, the primer is exactly complementary to a sequence that contains the C in position -13910 or the G in

WO 03/014342

PCT/EP02/08963

12

position -22018 or to the complementary strand thereof. The other type of primer is exactly complementary to a sequence having a T in position -13910 or an A in position -22018 or to the complementary strand thereof. Since hybridization conditions would preferably be chosen to be stringent enough, contacting of e.g. a primer exactly complementary to the mutant sequence with a wild-type allele would not result in efficient hybridization due to the mismatch formation. After washing, no signal would be detected due to the removal of the primer.

10 Additionally, the invention relates to a non-human host transformed with the vector of the invention as described herein above. The host may either carry the mutant or the wild-type sequence. Upon breeding etc. the host may be heterozygous or homozygous for one or both SNPs.

15 The host of the invention may carry the vector of the invention either transiently or stably integrated into the genome. Methods for generating the non-human host of the invention are well known in the art. For example, conventional transfection protocols described in Sambrook et al., loc. cit., may be employed to generate transformed bacteria (such as *E. coli*) or transformed yeasts. The non-human host of the invention may be used, for example, to elucidate the onset of adult-type hypolactasia.

20 In a preferred embodiment of the invention the non-human host is a bacterium, a yeast cell, an insect cell, a fungal cell, a mammalian cell, a plant cell, a transgenic animal or a transgenic plant.

25 Whereas *E. coli* is a preferred bacterium, preferred yeast cells are *S. cerevisiae* or *Pichia pastoris* cells. Preferred fungal cells are *Aspergillus* cells and preferred insect cells include *Spodoptera frugiperda* cells. Preferred mammalian cells are colon carcinoma cell lines showing expression of the LPH enzyme and include CaCo2-cells.

30 A method for the production of a transgenic non-human animal, for example transgenic mouse, comprises introduction of the aforementioned polynucleotide or targeting vector into a germ cell, an embryonic cell, stem cell or an egg or a cell derived therefrom. The non-human animal can be used in accordance with a screening method of the invention described herein. Production of transgenic

WO 03/014342

PCT/EP02/08963

13

embryos and screening of those can be performed, e.g., as described by A. L. Joyner Ed., *Gene Targeting, A Practical Approach* (1993), Oxford University Press. The DNA of the embryonal membranes of embryos can be analyzed using, e.g., Southern blots with an appropriate complementary nucleic acid molecule; see supra. A general method for making transgenic non-human animals is described in the art, see for example WO 94/24274. For making transgenic non-human organisms (which include homologously targeted non-human animals), embryonal stem cells (ES cells) are preferred. Murine ES cells, such as AB-1 line grown on mitotically inactive SNL76/7 cell feeder layers (McMahon and Bradley, *Cell* 62:1073-1085 (1990)) essentially as described (Robertson, E. J. (1987) in *Teratocarcinomas and Embryonal Stem Cells: A Practical Approach*, E. J. Robertson, ed. (Oxford: IRL Press), p. 71-112) may be used for homologous gene targeting. Other suitable ES lines include, but are not limited to, the E14 line (Hooper et al., *Nature* 326:292-295 (1987)), the D3 line (Doetschman et al., *J. Embryol. Exp. Morph.* 87:27-45 (1985)), the CCE line (Robertson et al., *Nature* 323:445-448 (1986)), the AK-7 line (Zhuang et al., *Cell* 77:875-884 (1994)). The success of generating a mouse line from ES cells bearing a specific targeted mutation depends on the pluripotency of the ES cells (i. e., their ability, once injected into a host developing embryo, such as a blastocyst or morula, to participate in embryogenesis and contribute to the germ cells of the resulting animal). The blastocysts containing the injected ES cells are allowed to develop in the uteri of pseudopregnant nonhuman females and are born as chimeric mice. The resultant transgenic mice are chimeric for cells having the desired nucleic acid molecule are backcrossed and screened for the presence of the correctly targeted transgene (s) by PCR or Southern blot analysis on tail biopsy DNA of offspring so as to identify transgenic mice heterozygous for the nucleic acid molecule of the invention.

The transgenic non-human animals may, for example, be transgenic mice, rats, hamsters, dogs, monkeys (apes), rabbits, pigs, or cows. Preferably, said transgenic non-human animal is a mouse. The transgenic animals of the invention are, inter alia, useful to study the phenotypic expression/outcome of the nucleic acids and vectors of the present invention. Furthermore, the transgenic animals of the present invention are useful to study the developmental expression of the LPH enzyme, for example in

WO 03/014342

PCT/EP02/08963

14

the rodent intestine. It is furthermore envisaged, that the non-human transgenic animals of the invention can be employed to test for therapeutic agents/compositions or other possible therapies which are useful to ameliorate adult-type hypolactasia.

- 5 In addition, the invention relates to an antibody or aptamer or phage that specifically binds to the mutant nucleic acid molecule of the invention but not to the corresponding wild type nucleic acid molecule.  
The antibody may be tested for binding and used in any serologic technique well known in the art, such as agglutination techniques in tubes, gels, solid phase and capture techniques with or without secondary antibodies, or in flow cytometry with or without immunofluorescence enhancement (see, for example, techniques described in Harlow and Lane „Antibodies, A Laboratory Manual“, CSH Press, Cold Spring Harbor, USA, 1988 (see reference 53).
- 10
- 15 In line with the invention, the antibody specifically recognizes an epitope comprising position -13910 (wherein the nucleotide is C) or position -22018 (wherein the nucleotide is G). It does not or essentially does not cross-react with an epitope comprising position -13910 with a T in this position nor with the epitope comprising position -22018 with a G in this position. Specificity of an antibody which may be generated according to standard protocols, may be tested by contacting with DNA
- 20 molecules carrying the wild-type and the mutant sequence such as in an ELISA assay. Only those antibodies will be selected that produce a signal over background with the mutant sequence but not with the wild-type sequence.
- 25 The antibody of the invention may be a monoclonal antibody or an antibody derived from or comprised in a polyclonal antiserum. The term "antibody", as used in accordance with the present invention, further comprises fragments of said antibody such as Fab, F(ab')<sub>2</sub>, Fv or scFv fragments; see, for example, Harlow and Lane<sup>58</sup>, loc. cit. The antibody or the fragment thereof may be of natural origin or may be
- 30 (semi)synthetically produced. Such synthetic products also comprise non-proteinaceous as semi-proteinaceous material that has the same or essentially the same binding specificity as the antibody of the invention. Such products may, for example, be obtained by peptidomimetics.

WO 03/014342

PCT/EP02/08963

15

The term "aptamer" is well known in the art and defined, e.g., in Osborne et al., *Curr. Opin. Chem. Biol.* 1 (1997), 5-9 (see reference 51) or in Stall and Szoka, *Pharm. Res.* 12 (1995), 465-483 (see reference 52).

5 Moreover, the invention relates to an antibody or aptamer or phage that specifically binds to the wild-type nucleic acid molecule as described herein above but not to the corresponding mutant sequence contributing to or indicative of adult-type hypolactasia. The statements with respect to specificity etc. made for the antibody which is specific for the mutant sequence apply *mutatis mutandis* here.

10

Furthermore, the invention relates to a pharmaceutical composition comprising the wild-type nucleic acid molecule as described herein above.

The pharmaceutical composition of the invention may be used in gene therapy approaches, particularly in somatic gene therapy.

15 The wild-type nucleic acid molecule referred to above and contained in the pharmaceutical composition of the invention may be combined with a pharmaceutically acceptable carrier and/or diluent.

Examples of suitable pharmaceutical carriers are well known in the art and include  
20 phosphate buffered saline solutions, water, emulsions, such as oil/water emulsions, various types of wetting agents, sterile solutions etc. Compositions comprising such carriers can be formulated by well known conventional methods. These pharmaceutical compositions can be administered to the subject at a suitable dose. Administration of the suitable compositions may be effected by different ways, e.g.,  
25 by intravenous, intraperitoneal, subcutaneous, intramuscular, topical, intradermal, intranasal or intrabronchial administration. The dosage regimen will be determined by the attending physician and clinical factors. As is well known in the medical arts, dosages for any one patient depends upon many factors, including the patient's size, body surface area, age, the particular compound to be administered, sex, time and  
30 route of administration, general health, and other drugs being administered concurrently. A typical dose can be, for example, in the range of 0.001 to 1000  $\mu\text{g}$  of nucleic acid for expression or for inhibition of expression; however, doses below or above this exemplary range are envisioned, especially considering the

WO 03/014342

PCT/EP02/08963

16

aforementioned factors. Dosages will vary but a preferred dosage for intravenous administration of DNA is from approximately  $10^8$  to  $10^{12}$  copies of the DNA molecule. Progress can be monitored by periodic assessment. The compositions of the invention may be administered locally or systemically. Administration will generally be  
5 parenterally, e.g., intravenously; DNA may also be administered directly to the target site, e.g., by biolistic delivery to an internal or external target site or by catheter to a site in an artery. Preparations for parenteral administration include sterile aqueous or non-aqueous solutions, suspensions, and emulsions. Examples of non-aqueous solvents are propylene glycol, polyethylene glycol, vegetable oils such as olive oil, and injectable organic esters such as ethyl oleate. Aqueous carriers include water,  
10 alcoholic/aqueous solutions, emulsions or suspensions, including saline and buffered media. Parenteral vehicles include sodium chloride solution, Ringer's dextrose, dextrose and sodium chloride, lactated Ringer's, or fixed oils. Intravenous vehicles include fluid and nutrient replenishers, electrolyte replenishers (such as those based on Ringer's dextrose), and the like. Preservatives and other additives may also be  
15 present such as, for example, antimicrobials, anti-oxidants, chelating agents, and inert gases and the like.

Additionally, the invention relates to a diagnostic composition comprising the nucleic acid molecule as described herein above, the vector as described herein above, the  
20 primer or primer pair as described herein above, and/or the antibody aptamer and/or phage as described herein above.

The diagnostic composition is useful for assessing the genetic status of a person with respect to his or her predisposition to develop adult-type hypolactasia or with regard  
25 to the diagnosis of the acute condition. The various possible components of the diagnostic composition may be packaged in one or more vials, in a solvent or otherwise such as in lyophilized form. If dissolved in a solvent, the diagnostic composition is preferably cooled to at least  $+8^{\circ}\text{C}$  to  $+4^{\circ}\text{C}$ . Freezing may be preferred in other instances.

30 The invention also relates to a method for testing for the presence or predisposition of adult-type hypolactasia or associated trait comprising testing a sample obtained from a prospective patient or from a person suspected of carrying such a

WO 03/014342

PCT/EP02/08963

17

predisposition to the presence of the nucleic acid molecule as described herein above in a homozygous or heterozygous state. In varying embodiments, it may be tested either for the presence of the wild-type sequence(s) or of the mutant sequence(s).

5

The method of the invention is useful for detecting the genetic set-up of said person/patient and drawing appropriate conclusions whether a condition from which said patient suffers is adult-type hypolactasia. Alternatively, it may be assessed whether a person not suffering from a condition carries a predisposition to adult-type hypolactasia. With regard to position -13910 upstream of the LPH gene, only if cytosine is found in a homozygous state, a condition would be diagnosed as adult-type hypolactasia or a corresponding predisposition would be manifest. On the other hand, if thymidine is found in a homozygous state or if the individual is heterozygous (C/T), then it may be concluded that a condition from which a patient suffers is not related to adult-type hypolactasia and further, that the patient does not carry a predisposition to develop this condition. It may, however, be concluded that children of persons carrying the heterozygous genotype may develop the condition if chromosome carrying the C residue is matched with a corresponding chromosome from the other parent.

20

The situation is similar and essentially the same conclusions apply for the analysis of the SNP in position -22018. A homozygously occurring G residue marks a predisposition to or the occurrence of acute adult-type hypolactasia. A heterozygous G/A state correlates with a high likelihood to not develop the condition. Individuals carrying A in a homozygous state would not be expected to develop the condition. Similarly, patients suffering from a condition would be diagnosed not to suffer from adult-type hypolactasia.

25

In a preferred embodiment of the method of the invention said testing comprises hybridizing the complementary nucleic acid molecule as described herein above which is complementary to the nucleic acid molecule contributing to or indicative of adult-type hypolactasia or the nucleic acid molecule as described herein above which is complementary to the wild-type sequence as a probe under (highly) stringent

30

WO 03/014342

PCT/EP02/08963

18

conditions to nucleic acid molecules comprised in said sample and detecting said hybridization.

Again, depending on the nucleic acid probe used, either wild-type or mutant sequences (i.e. sequences contributing to or indicative of adult-type hypolactasia) would be detected. It is understood that hybridization conditions would be chosen such that a nucleic acid molecule complementary to wild-type sequences would not or essentially not hybridize to the mutant sequence. Similarly, a nucleic acid molecule complementary to the mutant sequence would not or would not essentially not hybridize to the wild-type sequence. In order to differentiate between results obtained from homozygous and heterozygous genotypes in the hybridization methods of the invention, one can for example monitor/detect the strength/intensity of the respective detection signal after the hybridization. To differentiate between wild-type homozygous, heterozygous and/or mutant homozygous alleles in the hybridization methods of the invention, internal control samples of the corresponding genotypes will be included in the analysis.

In a further preferred embodiment, the method of the invention further comprises digesting the product of said hybridization with a restriction endonuclease or subjecting the product of said hybridization to digestion with a restriction endonuclease and analyzing the product of said digestion.

This preferred embodiment of the invention allows by convenient means, the differentiation between an effective hybridization and a non-effective hybridization. For example, if the DNA sequence adjacent to position -13910 or position -22018 comprises an endonuclease restriction site, the hybridized product will be cleavable by an appropriate restriction enzyme upon an effective hybridization whereas a lack of hybridization will yield no double-stranded product or will not comprise the recognizable restriction site and, accordingly, will not be cleaved. In particular, the restriction enzymes specific for the sequence of the DNA-variant C/T<sub>-13910</sub> is CviI I, for the DNA-variant G/A<sub>-22018</sub> are HhaI and Acl I. Said restriction enzymes which cut rg/cy where found by the use of the program Webcutter. The analysis of the digestion product can be effected by conventional means, such as by gel electrophoresis which may be optionally combined by the staining of the nucleic acid with, for example, ethidium bromide. Combinations with further techniques such as Southern

WO 03/014342

PCT/EP02/08963

19

blotting are also envisaged.

Detection of said hybridization may be effected, for example, by an anti-DNA double-strand antibody or by employing a labeled oligonucleotide. Conveniently, the method  
5 of the invention is employed together with blotting techniques such as Southern or Northern blotting and related techniques. Labeling may be effected, for example, by standard protocols and includes labeling with radioactive markers, fluorescent, phosphorescent, chemiluminescent, enzymatic labels, etc. (see also above).

10 In accordance with the above, in another preferred embodiment of the method of the invention said probe is detectably labeled, e.g. by the methods and with the labels described herein above.

In yet another preferred embodiment of the method of the invention said testing  
15 comprises determining the nucleic acid sequence of at least a portion of the nucleic acid molecule as described herein above, said portion comprising nucleotide position -13910 and/or nucleotide position -22018 of the LPH gene.

Determination of the nucleic acid molecule may be effected in accordance with one  
20 of the conventional protocols such as the Sanger or Maxam/Gilbert protocols (see Sambrook et al., loc. cit., for further guidance).

In a further preferred embodiment of the method of the invention the determination of the nucleic acid sequence is effected by solid-phase minisequencing. Solid-phase minisequencing is based on quantitative analysis of the wild type and mutant  
25 nucleotide in a solution. First, the genomic region containing the mutation is amplified by PCR with one biotinylated and non-biotinylated primer where the biotinylated primer is attached to a streptavidin (SA) coated plate. The PCR-product is denatured to a single stranded form to allow a minisequencing primer to bind to this strand just before the site of the mutation. The tritium (H3) or fluorescence labeled mutated and  
30 wild type nucleotides together with nonlabeled dNTPs are added to the minisequencing reaction and sequenced using Taq-polymerase. The result is based on the amount of wild type and mutant nucleotides in the reaction measured by beta counter or fluorometer and expressed as an R-ratio. See also Syvänen AC, Sajantila

WO 03/014342

20

PCT/EP02/08963

A, Lukka M. Am J Hum Genet 1993: 52,46-59 and Suomalainen A and Syvanen AC. Methods Mol Biol 1996;65:73-79.

A preferred embodiment of the method of the invention further comprises, prior to determining said nucleic acid sequence, amplification of at least said portion of said nucleic acid molecule.

Preferably, amplification is effected by polymerase chain reaction (PCR). Other amplification methods such as ligase chain reaction may also be employed.

In a preferred embodiment of the method of the invention said testing comprises carrying out an amplification reaction wherein at least one of the primers employed in said amplification reaction is the primer as described herein above or belongs to the primer pair as described herein above, comprising assaying for an amplification product. In this embodiment and depending on the information the investigator/physician wishes to obtain, primers hybridizing either to the wild-type or mutant sequences may be employed.

The method of the invention will result in an amplification of only the target sequence, if said target sequence carries a sequence exactly complementary to the primer used for hybridization. This is because the oligonucleotide primer will under preferably (highly) stringent hybridization conditions not hybridize to the wild-type/mutant sequence – depending which type of primer is used – (with the consequence that no amplification product is obtained) but only to the exactly matching sequence. Naturally, combinations of primer pairs hybridizing to both SNPs may be used. In this case, the analysis of the amplification products expected (which may be no, one, two, three or four amplification product(s) if the second, non-differentiating primer is the same for each locus) will provide information on the genetic status of both positions – 13910 and –22018.

In a preferred embodiment of the method of the invention said amplification is effected by or said amplification is the polymerase chain reaction (PCR).

The PCR is well established in the art. Typical conditions to be used in accordance with the present invention include for example a total of 35 cycles in a total of 50 $\mu$ l volume exemplified with a denaturation step at 93° C for 3 minutes; an annealing

WO 03/014342

PCT/EP02/08963

21

step at 55° C for 30 seconds; an extension step at 72° C for 75 seconds and a final extension step at 72° C for 10 minutes.

- The invention furthermore relates to a method for testing for the presence or predisposition of adult-type hypolactasia comprising assaying a sample obtained from a human for specific binding to the antibody or aptamer or phage as described herein above. In this context a weaker staining for the presence of the antigen of the invention compared to homozygous wild type control samples (comprising two persistent alleles) is indicative for the heterozygous wild type (one persistent allele and one hypolactasic allele, whereas for the homozygous hypolactasic individual no staining is expected if the appropriate antibody is used. Preferably, the method of the invention is performed in the presence of control samples corresponding to all three possible allelic combinations as internal controls. Testing may be carried out with an antibody etc. specific for the wild-type or specific for the mutant sequence.
- 15 Testing for binding may, again, involve the employment of standard techniques such as ELISAs; see, for example, Harlow and Lane<sup>65</sup>, loc. cit.

In a preferred embodiment of the method of the invention said antibody or aptamer or phage is detectably labeled.

- 20 Whereas the aptamers are preferably radioactively labeled with <sup>3</sup>H or <sup>32</sup>P or with a fluorescent marker as described above, the phage or antibody may either be labeled in a corresponding manner (with <sup>131</sup>I as the preferred radioactive label) or be labeled with a tag such as His-tag, FLAG-tag or myc-tag.

- 25 In a further preferred embodiment of the method of the invention the test is an immuno-assay.

- In another preferred embodiment of the method of the invention said sample is blood, serum, plasma, fetal tissue, saliva, urine, mucosal tissue, mucus, vaginal tissue, fetal tissue obtained from the vagina, skin, hair, hair follicle or another human tissue.

In an additional preferred embodiment of the method of the invention said nucleic acid molecule from said sample is fixed to a solid support.

WO 03/014342

PCT/EP02/08963

22

Fixation of the nucleic acid molecule to a solid support will allow an easy handling of the test assay and furthermore, at least some solid supports such as chips, silica wafers or microtiter plates allow for the simultaneous analysis of larger numbers of samples. Ideally, the solid support allows for an automated testing employing, for example, roboting devices.

In a particularly preferred embodiment of the method of the invention said solid support is a chip, a silica wafer, a bead or a microtiter plate.

Furthermore, the invention relates to the use of the nucleic acid molecule as described herein above for the analysis of the presence or predisposition of adult-type hypolactasia.

The nucleic acid molecule simultaneously allows for the analysis of the absence of the condition or the predisposition to the condition, as has been described in detail herein above.

In addition, the invention relates to a kit comprising the nucleic acid molecule as described herein above, the primer or primer pair as described herein above, the vector as described herein above, and/or the antibody aptamer and/or phage as described herein above in one or more containers.

The invention as well relates to the use of the nucleic acid molecule as described herein above or the vector as described herein above in gene therapy.

Gene therapy approaches have been discussed herein above in connection with the vector of the invention and equally apply here. It is of note that in accordance with this invention, also fragments of the nucleic acid molecules as defined herein above and as, in particular, depicted in SEQ ID NOs: 3 to 4 may be employed in gene therapy approaches. Said fragments comprise the nucleotide at position -13910 as defined in (c) herein above (and also shown in SEQ ID NO: 3) or position -22018 as defined in (d) herein above (and as shown in SEQ ID NO: 4). Preferably, said fragments comprise at least 200, at least 250, at least 300, at least 400 and most preferably at least 500 nucleotides.

WO 03/014342

PCT/EP02/08963

23

In a preferred embodiment of the use of the invention said gene therapy treats or prevents adult-type hypolactasia.

5 The figures show:

**Fig. 1:** The Finnish adult-type hypolactasia families studied. Blackened symbols indicate hypolactasic individuals, asterisk (\*) indicate that no sample was available, question mark (?) indicates unknown affection status. † indicates the individuals used for sequencing for SNP identification (Table 2).

**Fig 2:** Physical map of adult-type hypolactasia locus. BAC clones are shown above the horizontal line. The three genes LPH, MCM6 and DARS are shown by thick black arrows with the tip pointed toward the 3' end of the gene above the black boxes. The position of ten polymorphic microsatellite markers used for fine mapping of the locus are shown. The backslash in the horizontal line denotes a gap in the sequence of the contig sequence. The position of marker D2S2169 was confirmed by bridging the gap with PAC 106O20 isolated from the PAC library as described before<sup>40</sup>. The organisation of the MCM6 gene is shown including the position of the lactase persistent phenotype-associated variants in introns 9 and 13 located 13.9 kb and 22 kb 5' of the first ATG of LPH.

**Fig. 3:** Extended haplotype analysis of the persistent chromosomes derived from Finnish adult-type hypolactasia families using seven closely linked microsatellite markers. The haplotypes representing the ancestral founder persistent chromosome are shaded. Only the haplotypes of non-persistent chromosomes that were also present in the persistent chromosomes are shown. On the basis of ancestral recombinations, the adult-type hypolactasia locus could be restricted to 47 kb interval between markers LPH1 and AC3.

**Fig. 4:** The sequence comprised in the sequence of intron 13 of the MCM6 gene (3220bp) comprising the SNP at position -13910 in which the T, which is

specific for the lactase persistence, is substituted by a C. Said position is indicated by the use of a small letter. This sequence refers to SEQ ID NO:1.

- 5 **Fig. 5:** The sequence comprised in the sequence of intron 9 of the MCM6 gene(1295bp) comprising the SNP at position -22018 in which the A, which is specific for the lactase persisting-type sequence is substituted by a G. Said position is indicated by the use of a small letter. This sequence refers to SEQ ID NO:2.
- 10 **Fig. 6:** The sequence of the lactase persisting-type intron 13 of the MCM6 gene (3220bp) comprising at position -13910 a T. Said position is indicated by the use of a small letter. This sequence refers to SEQ ID NO:3.
- 15 **Fig. 7:** The sequence of the lactase persisting-type intron 9 of the MCM6 gene(1295bp) comprising at position -22018 an A. Said position is indicated by the use of a small letter. This sequence refers to SEQ ID NO:4.
- 20 **Fig. 8:** The sequence of intron 13 of the MCM6 gene (3220bp) comprising the SNP at position -13910 in which the T, which is specific for the lactase persisting-type sequence is substituted by a C. Said position is indicated by the use of a small letter. This sequence refers to SEQ ID NO:5.
- 25 **Fig. 9:** The sequence of intron 9 of the MCM6 gene(1295bp) comprising the SNP at position -22018 in which the A, which is specific for the lactase persisting-type sequence is substituted by a G. Said position is indicated by the use of a small letter. This sequence refers to SEQ ID NO:6.

The examples illustrate the invention.

**Example 1: Linkage and linkage disequilibrium analysis**

5 Seven polymorphic microsatellite markers between D2S114 and D2S2385 flanking the *LPH* gene on 2q21 were analyzed in nine extended Finnish hypolactasia families (Fig. 1). Significant evidence for linkage was found with markers D2S314, D2S442, D2S2196 and D2S1334, with a maximum lod score of 7.67 at  $\theta = 0$  obtained with marker D2S2196 (Table 1). Obligatory recombination events were detected with  
10 marker D2S114 (family B, IV3), which defines the centromeric boundary for the lactase persistence/non-persistence locus, and with marker D2S2385 (family B, IV17) (Fig. 1, Table 1), which defines the telomeric boundary of the locus. To fine map the critical region, nine additional polymorphic markers were analyzed (Table 1). Linkage disequilibrium (LD) over the region was monitored conditional on the detected linkage  
15 treating the allele frequencies and the recombination fraction as nuisance parameters<sup>16-17</sup>. Six out of nine markers (LPH13, LPH2, LPH1, AC3, AC4, and AC10), spanning over ~200kb interval showed highly significant evidence of LD ( $p < 10^{-4}$ ) whereas markers 3' from the *LPH* gene showed no evidence of LD (Table 1). Two markers, LPH2 and AC3, displayed the most significant linkage disequilibrium in the lactase persistence alleles ( $p < 10^{-7}$ ).  
20

The family material consisted of nine extended Finnish pedigrees originally studied by Sahi<sup>5</sup>. All family material was tested for adult-type hypolactasia in the 1970s. The family material for this study was enlarged by collecting the DNA of the family  
25 members in the younger generations. The family material in this study consisted of 194 individuals in total (Fig. 1). The phenotypic status of all family members was confirmed by lactose tolerance tests with ethanol (LTTE)<sup>4-5</sup> in all but 49 individuals. Gluten enteropathy has been excluded in all affected patients by measurement of the serum IgA anti-tissue transglutaminase<sup>45</sup>. DNA was extracted from blood samples  
30 taken from all participating family members in accordance with standard protocols<sup>46</sup>, after obtaining informed consent. As a case-control study 196 random DNA samples isolated from jejunal biopsy specimens from which disaccharidase activities had been measured<sup>47</sup> at the Helsinki University Hospital were sequenced. DNA was isolated

WO 03/014342

PCT/EP02/08963

26

from intestinal biopsies according to the standard protocol<sup>46</sup>. These series comprised 137 lactase persistent and 59 non-persistent samples. In addition DNA from nine Italian, kindly provided by M. Rossi, University of Naples, nine German DNA samples, kindly provided by M. Lentze, University of Bonn and twenty two South Korean, kindly provided by J.K. Seo, Seoul National University, intestinal biopsy sample specimens were analyzed (In the table: 23 Korean, 9 Italian and 7 Germans (One of the cases from Germany originated from South Korea). The diagnosis was based on the measurement of disaccharidase activities. Finally, to determine the frequency of the C/T<sub>-13910</sub> variant in the Finnish population, the DNA of 938 anonymous Finnish blood donors from small parishes from Eastern and Western Finland and the DNA of 109 parents belonging to the CEPH families<sup>19</sup> were analyzed. In addition, genomic DNA from a baboon (*Papio hamedryas ussinus*) isolated from liver biopsy using standard protocols<sup>45</sup> was analyzed. The study was approved by the Ethical Committees of the Helsinki University Hospital and the Finnish Red Cross Blood Transfusion Service.

#### Example 2: Extended haplotype analysis

In the first stage ten highly polymorphic microsatellite markers flanking the *LPH* gene on 2q21 were analyzed as described elsewhere<sup>40,55</sup>. Briefly, the ten highly polymorphic microsatellite markers on 2q in the vicinity of the lactase gene from The Généthon Resource Center<sup>55</sup> were analyzed with genetic distances as follows: cen - D2S114 - 1cM - D2S1334 - 0cM - D2S2196 - 0cM - D2S442 - 2cM - D2S314 - 2cM - D2S2385 - 1cM - D2S2288 - 1cM - D2S397 - 1cM - D2S150 - 1cM - D2S132. The order of the markers has been mostly obtained from the physical YAC contig map of chromosome 2 (Chumakov et al. 1995<sup>55</sup>) supplemented with the Généthon map. PCR was performed in a total volume of 15 ul containing 12ng of template DNA, 5pmol of primers, 0.2mM of each nucleotide, 20mMTrisHCl (pH 8.8), 15 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 1.5 mM MgCl<sub>2</sub>, 0.1% Tween 20, 0.01% gelatin and 0.25U Taq polymerase (Dynazyme, Finnzymes). One of the primers was radiolabeled at the 5' end with <sup>32</sup>P-γATP. The reactions were performed in a multiwell microtitre plate for 35 cycles with denaturation at 94 °C for 30s, annealing at various temperatures

WO 03/014342

PCT/EP02/08963

27

depending on the primers for 30s and extension at 72 °C for 30s; denaturation was set at 3min and final extension at 5min. The amplified fragments were separated on 6% polyacrylamide gel, and autoradiography was performed.

In the second stage, nine additional microsatellite markers within the contig constructed over the *LPH* gene were identified from the published genomic sequence of the BACs (NH034L23, NH0318L13, NH0218L22, and RP11-32911) using the Repeat Masker program (<http://ftp.genome.washington.edu/cgi-bin/RepeatMasker>). Primers flanking the repeats were synthesized. PCR conditions were as described elsewhere<sup>40</sup>. The amplified fragments were separated on 6% polyacrylamide gel, and autoradiography was performed.

Pairwise lod scores were calculated by use of the MLINK option of the LINKAGE program package<sup>49</sup>. Autosomal recessive inheritance for adult-type hypolactasia with complete penetrance, no sex difference in recombination fractions, and a disease allele frequency of 0.4 was assumed. Only individuals above 20 years of age were included in the study as the condition is manifested by that age in the Finnish population<sup>5-9</sup>. The affection status for individuals not confirmed by LTTE was regarded as unknown. Allele frequencies and heterozygosities for the markers were estimated from family material using the Downfreq program for purposes of the parametric linkage analysis<sup>49</sup>. Additionally, pseudomarker linkage and linkage disequilibrium analyses were performed, assuming autosomal recessive mode of inheritance<sup>16</sup>. A test of LD was performed conditional on the detected linkage treating the allele frequencies and the recombination fraction as nuisance parameters<sup>16,49</sup>. *P*-values from these analyses are shown in Table 1. Haplotypes were constructed manually for the microsatellite markers in this order: LPH1-LPH2-LPH13-AC7-AC3-AC4-AC5 (Fig. 3). A total of 54 non-persistent chromosomes and 33 persistent chromosomes in our family material were available for haplotype analysis.

The order of the closely linked markers was confirmed by assembling four BAC-clones NH0034L23, NH0218L22, NH0318L13 and 329110 in the critical region into one uninterrupted sequence segment. This contig extended from marker AC8 to the exon 10 of the aspartyl-tRNA synthetase (*DARS*) gene and covered a total of 222,5 kb (Fig. 2). Based on this physical map of the linked region, extended haplotypes

WO 03/014342

PCT/EP02/08963

28

with seven markers covering a 150 kb interval (cen-LPH13-LPH2-LPH1-AC7-AC3-AC4-AC5-tel) (Fig. 3) were constructed. One major haplotype was present in 20 persistence alleles (60%) versus 3 of the non-persistence alleles (5%), whereas a wide diversity of haplotypes was observed in non-persistence alleles. The remaining 5 40 % of the haplotypes in the persistence alleles differed from the ancestral haplotype in a manner consistent with a breakdown of the haplotype by historical recombination events. Based on the conserved haplotype analysis, the locus for lactase persistence could be restricted to a 47 kb interval between markers LPH1 and AC3 (Fig.3)

10

**Example 3: Sequence analysis of the adult-type hypolactasia locus**

The 47 kb region between the markers LPH1 and AC3 was amplified in overlapping 15 PCR fragments from genomic DNA of several members of the nine hypolactase families and sequenced. The region contains the minichromosome maintenance (*MCM6*) gene<sup>18</sup>, which covers 36 kb of the critical 47 kb region (Fig. 2). No variations were detected in the coding region of the *MCM6* gene but total of 52 variants; 43 SNPs and 9 deletion/ insertion polymorphisms, were identified in the critical 47 kb 20 region (Table 2). Only two of the variants (C/T<sub>-13910</sub>, G/A<sub>-22018</sub>) were associated with the lactase persistence/ non-persistence trait in the Finnish families (Tables 2 and 3). The first associated variant, C/T<sub>-13910</sub>, resides in intron 13 of the *MCM6* gene at position -13910 bp from the first ATG-codon of the *LPH* gene. The second associated variant, G/A<sub>-22018</sub>, is located in intron 9 of the *MCM6* gene at position - 25 22018 from the first ATG-codon of the *LPH* gene (Fig.2). These two variants, 8 kb apart from each other, completely cosegregated with adult-type hypolactasia in nine extended Finnish families. All hypolactasic (non-persistent) family members were homozygous for both C<sub>-13910</sub> and G<sub>-22018</sub> (Table 3). Interestingly, both these variants reside in repeat elements, C/T<sub>-13910</sub> in an L2-derived element and G/A<sub>-22018</sub> in an Alu 30 element.

Experimentally, three non-persistence, 2 homozygous persistence and 2 heterozygous persistence individuals sharing a similar haplotype across the critical

WO 03/014342

PCT/EP02/08963

29

region from our family material were used for sequencing in the first stage (Fig. 1). Using the published draft genomic sequence of the BACs: NH0034L23, NH0218L22, NH0318L23, and RP-329110 that covered the critical region of adult-type hypolactasia were assembled to one contig using Sequencher 4 software (Gene Codes Corporation). Oligonucleotide primers spanning the critical region between markers LPH1 and AC3 were designed (a list of oligonucleotide primers described herein below). PCR amplifications were carried out in a 50  $\mu$ l volume with genomic DNA (100 ng), primers (20 ng each), dNTPs (200  $\mu$ M), 0.5 U of *Taq* polymerase (Dynazyme, Finnzymes) in a standard buffer. Most PCR were amplified using the following PCR cycle conditions: an initial round of denaturation at 94 °C for 3 min, then 35 cycle at 94°C at 30 s, 55 °C for 30 s, and 72 °C for 1.25 min and a final extension of 72 °C for 10 min, except that in cases where the size of the PCR products were more than 1kb we used the Dynazyme extend kit (conditions are described herein below). Purified PCR products (15-40 ng) were cycle sequenced using BigDye terminator chemistry (PE Biosystems). Data were analyzed using ABI Sequencing Analysis 3.3 (PE Biosystems) and Sequencher 4.1 (Gene Codes).

**Detection of the Lactase variants by Sequencing:**

PCR amplifications were carried out in a 50  $\mu$ l volume with genomic DNA (100 ng), primers (20 ng each), dNTPs (200  $\mu$ M), 0.5 U of *Taq* polymerase (Dynazyme, Finnzymes) in a standard buffer. Both PCRs were amplified using the following PCR cycle conditions: an initial round of denaturation at 94°C for 3 min, then 35 cycles at 94°C at 30 s, 55°C for 30 s, and 72°C for 1.25 min and a final extension of 72°C for 10 min. PCR were purified by enzymatic reaction. Purified PCR products (15-40 ng) were cycle sequenced using BigDye terminator chemistry (PE Biosystems). Data were analyzed using ABI Sequencing Analysis 3.3 (PE Biosystems) and Sequencher 4.1 (Gene Codes).

**Screening of the lactase variants by solid-phase minisequencing:**

The DNA fragment spanning the C/T<sub>-13910</sub> variant was amplified using one biotinylated (5'-Bio-CCTCGTTAATACCCACTGACCTA-3') primer and unbiotinylated (5'-GTCACCTTTGATATGATGAGAGCA-3') primer. For G/A<sub>-22018</sub> biotinylated (5'-Bio-TGCTCAGGACATGCTGATCAA-3') and one unbiotinylated (5'-CTACCCTATCAGTAAAGCCCTA-3') primer were used under conditions described

WO 03/014342

PCT/EP02/08963

30

above. 10  $\mu$ l of the PCR product was captured in a streptavidin coated microtiter well (Lab systems, Finland). The wells were washed, and bound DNA was denaturated as described by Syvänen et al. (*Am J Hum Genet.* (1993), 52, 46-59) and Syvänen and Landegren (*Hum Mutat.* (1994), 3, 172-9). 50  $\mu$ l of the minisequencing reaction mixture contained 10 pmoles of the minisequencing primers for C/T<sub>-13910</sub> (5'-GGCAATACAGATAAGATAAATGTAG-3'), G/A<sub>-22018</sub> (5'-AAAAACAGCATTCCTCAGCTGGGC-3'), and 0.1  $\mu$ l of either H-dCTP, H-dGTP corresponding to the lactase non-persistence allele (115 Ci/mmol; Amersham, UK) or H-dTTP, H-sATP corresponding to the lactase persistence allele and 0.05 units of DNA polymerase (Dynazyme II, Finnzymes) in its buffer was added to each well. The microtiter plates were incubated for 20 min at 50°C, and the wells were washed. The detection was eluted, and the eluted radioactivity was measured in a liquid scintillation counter (Rackbeta 1209, Wallac, Finland). Two parallel minisequencing reactions were carried out for each PCR product.

15 PCR primers and detection primer for the C/T<sub>-13910</sub> variant:

Forward PCR primer:	GTCACCTTTGATATGATGAGAGCA	Tm 58	SEQ ID NO: 8
Detection primer:	GGCAATACAGATAAGATAAATGTAG	Tm 58	SEQ ID NO: 10
Bio-Reverse primer:	Bio-CCTCGTTAATACCCACTGACCTA	Tm 62	SEQ ID NO: 9
or	Bio-TAGGTCAGTGGGTATTAACGAGGT		SEQ ID NO: 7

20 PCR primers and detection primer for the G/A<sub>-22018</sub> variant:

Forward PCR primer:	CTACCCATCAGTAAAGGCCTA	Tm 58	SEQ ID NO: 12
Detection primer:	AAAAACAGCATTCCTCAGCTGGGC	Tm 62	SEQ ID NO: 14
Bio-Reverse primer:	Bio-TGCTCAGGACATGCTGATCAA	Tm 62	SEQ ID NO: 13
or	Bio-TTGATCAGCATGCTCTGAGCA		SEQ ID NO: 11

25

**Example 4: Monitoring the DNA-variants in a case/ control study sample**

The frequency of the C/T<sub>-13910</sub> and G/A<sub>-22018</sub> variants was analyzed in DNA samples isolated from a total of 196 intestinal biopsy samples specimens which had been analyzed for disaccharidase activity as a diagnostic test for hypolactasia. A total of 59 samples showed primary lactase deficiency. Six out of 59 cases (Table 3) were heterozygous GA for the G/A<sub>-22018</sub> variant, the remaining 53 being homozygous for the G allele. All 59 samples were homozygous for the C allele of the variant C/T<sub>-13910</sub>.

30

WO 03/014342

PCT/EP02/08963

31

Among the 137 cases showing lactase persistence, 74 were found to be homozygous for alleles T and A, 63 being heterozygous CT and GA and none being homozygous for alleles C and G at C/T<sub>-13910</sub> and G/A<sub>-22018</sub>, respectively (Table 3).

- 5 To analyze these variants in other populations, DNA samples isolated from intestinal biopsy specimens from 40 non-Finnish cases with established disaccharidase deficiency were sequenced: 23 cases originated from South Korea, 9 from Italy and 8 from Germany. One Italian case was heterozygous GA for G/A<sub>-22018</sub> whereas all remaining 39 cases were homozygous CC and GG for C/T<sub>-13910</sub> and G/A<sub>-22018</sub> respectively (Table 3). An extended study gave rise to the data provided in Table 7 representing data of the complete association of C/T<sub>-13910</sub> variant with the biochemically verified hypolactasia (lactase non-persistence) in 400 individuals for 6 different populations. The G/A<sub>-22018</sub> variant was associated with the lactase non-persistence in 400 out of 401 cases.
- 15

**Example 5: Molecular epidemiology of the lactase persistence variant C/T<sub>-13910</sub>**

- To monitor for the prevalence of the hypolactasia-associated variant in the Finnish population a solid-phase minisequencing method<sup>19,20</sup> was used to screen DNA samples of 938 anonymous Finnish blood donors originating either from the Western early settlement region or the Eastern late settlement region of Finland (Table 4). Experimentally, the DNA fragment spanning the C/T<sub>-13910</sub> variant was amplified using one biotinylated (5'-CCTCGTTAATACCCCTGACCTA-3') primer and unbiotinylated (5'-GTCACCTTTGATATGATGAGAGCA-3') primer. For G/A<sub>-22018</sub> we used one biotinylated (5'-AGTCTGTGGCATGTGTCTTCATG-3') and one unbiotinylated (5'-TGCTCAGGACATGCTGATCAACT-3') primer under conditions described above. 10 µl of the PCR product was captured in a streptavidin coated microtitre well (Lab system, Finland). The wells were washed, and the bound DNA was denatured as described previously<sup>19,20</sup>. 50 µl of the minisequencing reaction mixture contain 10 pmoles of the minisequencing primers for G/A<sub>-22005</sub> (5'-GACAAAGGTGTGAGCCACCG-3'), G/A<sub>-13915</sub> (5'-GGCAATACAGATAAGATAATGTAG-3') and 0,1 µl of either H-dCTP corresponding
- 20
- 25
- 30

WO 03/014342

PCT/EP02/08963

32

to the lactase non-persistence allele (115 Ci/mmol; Amersham, UK) or H-dTTP corresponding to the lactase persistence allele and 0.05 units of DNA polymerase (Dynazyme II, Finnzymes) in its buffer was added to each well. The microtiter plates were incubated for 20 min at 50 °C, and the wells were washed. The detection primer was eluted, and the eluted radioactivity was measured in a liquid scintillation counter (Rackbeta 1209, Wallac, Finland). Two parallel minisequencing reactions were carried out for each PCR product. The overall prevalence of the putative hypolactasia genotype CC<sup>-13910</sup> (170 cases) was 18.1%, with higher prevalence (16.8% versus 18.9%) in the western than in the eastern sample (Table 4). These values are in good agreement with the epidemiological study reporting the prevalence of 17% among Finnish speaking Finns with an increasing gradient from West to East<sup>2</sup>. The same set of samples for the G/A<sup>22018</sup> polymorphism was also genotyped, and the LD between these two SNPs monitored using the D' statistic<sup>21</sup>. They were found to be in almost complete LD (D' = 0.98,  $p = 7.62 \times 10^{-11}$ , Table 5).

15

The prevalence of hypolactasia in different populations is known to vary greatly from less than 5% to almost 100%<sup>3,6</sup>. To determine whether these changes in hypolactasia prevalence would correlate with the distribution of the genotype CC<sup>-13910</sup>, the DNA of the parents of CEPH families<sup>22</sup> was analyzed. CEPH families have been mainly collected from France, with reported prevalence of hypolactasia around 37%<sup>23</sup> and Utah, the Utah populations originating from Northern Europe with prevalence of hypolactasia less than 5%<sup>24</sup>. Genotyping of the parents in CEPH families revealed that 41,2% (7 out of 17 samples) of French families have the genotype CC whereas only 7,6% (7 out of 92 samples) of Utah families have the genotype CC (Table 4). Again, despite the small number of analyzed samples these figures agree with the values obtained in the epidemiological studies of hypolactasia in these populations<sup>23,24</sup>.

Table 8 demonstrates that the observed prevalence of the variants well agrees with the described population frequencies of the lactose intolerance.

30

WO 03/014342

PCT/EP02/08963

33

**Example 6: The genealogy of the lactase persistence variant C/T<sub>-13910</sub>**

Haplotype analysis in the Finnish families suggested that most if not all, lactase persistence alleles in Finland have descended from one common ancestor. Linkage disequilibrium was used to estimate the time of the introduction of the persistence allele into the Finnish population<sup>25</sup>. Assuming 20 years generation time, this estimate would indicate that the founder mutation was introduced into the Finnish population some 9000-11400 years ago (Table 6). This is in good agreement with earliest signs of settlement in the Finnish mainland some 8000-9000 years ago<sup>26</sup> and would reasonably well coincide with the beginning of the dairy farming in 8000-10.000 BC<sup>27</sup>. More importantly, the presence of the same DNA-variant in persistence alleles in different populations would suggest that this variant is even more ancient and the mutation has occurred before differentiation of the analyzed populations.

To get some insight into the phylogenetic origin of the lactase allele, intron 9 and part of intron 13 of the *MCM6* gene of a Baboon (*Papio Hamadryas*) were sequenced. Genotype GG and CC was present in Baboons DNA at both G/A<sub>-22018</sub> and C/T<sub>-13910</sub>. This could suggest that alleles G and C, respectively reflect the appearance of the ancestral allele, presenting the non-persistence type and a mutation has transformed this allele to create the persistence allele. This assumption is supported by the identification of the LD and shared haplotype in the persistence alleles versus a high diversity of alleles found in non-persistence alleles.

**Example 7: Pairwise LD of C/T and G/A variants.**

Pairwise LD between C/T<sub>-13910</sub> and G/A<sub>-22018</sub> was estimated using the D' statistic<sup>21</sup>. Haplotype frequencies were estimated by maximum likelihood using the EH program<sup>50</sup>. D' is calculated as  $\max(D/D_{\max}, D/D_{\min})$ ; where disequilibrium measure  $D = h_{pq} - p q$ , where  $h_{pq}$  is the frequency of the haplotype with rare allele at each locus,  $p$  and  $q$  are frequency of the rare alleles at loci 1 and 2, and  $D_{\max} = \min(p(1-p), q(1-q))$  if  $D > 0$ , and  $D_{\min} = -\min(pq, (1-p)(1-q))$  if  $D < 0$ . The significance of deviation of D' from 0 was determined using the statistic  $D'^2 \sqrt{\frac{N}{p(1-p)q(1-q)}}$  which is distributed as  $\chi^2$  with 1 df<sup>21</sup>

WO 03/014342

PCT/EP02/08963

34

**Gene accessions numbers.** For BACs NH0218L22, N0034L34, NH0318L13, and RP11-329110 are AC012551, AC011893, AC011999 and AC016516 respectively.

5 The accession numbers for human polymorphisms are GenBank AF395607-  
AF395615.

#### References

- 10
1. Flatz, G. & Rothauwe, H. The human lactase polymorphism: physiology and genetics of lactose absorption and malabsorption. *Prog. Med. Genet.* **2**, 205-249 (1977).
  - 15 2. Sahi, T., Isokoski, M., Jussila, J. & Launiala, K. Lactose malabsorption in Finnish children of school age. *Acta Paediatr Scand.* **61**, 11-16 (1972).
  3. Wang, Y. *et al.* The genetically programmed down-regulation of lactase in children. *Gastroenterology.* **114**:1230-1236 (1998).
  - 20 4. Sahi, T., Isokoski, M., Jussila, J., Launiala, K. & Pyörälä, K. Recessive inheritance of adult-type lactose malabsorption. *Lancet.* **823-826** (1973).
  5. Sahi, T. The inheritance of selective adult-type lactose malabsorption. *Scand. J. Gastroenterol. suppl.* **30**, 1-73(1974).
  - 25 6. Sahi, T. Genetics and epidemiology of adult-type hypolactasia. *Scand. J. Gastroenterol. Suppl.* **202**, 7-20 (1994).
  - 30 7. Boll, W., Wagner, P. & Mantei, N. Structure of the chromosomal gene and cDNAs coding for lactase-phlorizin hydrolase in human with adult-type hypolactasia or persistence of lactase. *Am. J. Hum. Genet.* **48**, 889-902 (1991).
  8. Mantei, N. *et al.* Complete primary structure of human and rabbit lactase-phlorizin hydrolase: implications for biosynthesis, membrane anchoring and evolution of the enzyme. *EMBO J.* **7**, 2705-2713 (1988).
  - 35 9. Wang, Y. *et al.* The lactase persistence/non-persistence polymorphism is controlled by a cis-acting element. *Hum. Mol. Genet.* **4**, 657-662 (1995).
  - 40 10. Harvey, C.B., Pratt, W.S., Islam, I., Whitehouse, D.B. & Swallow, D.M. DNA polymorphisms in the lactase gene: linkage disequilibrium across the 70 kb region. *Eur. J. Hum. Genet.* **3**, 27-41 (1995).
  - 45 11. Escher, J.C. *et al.* Molecular basis of lactase levels in adult humans. *J. Clin. Invest.* **89**, 480-483 (1992).

WO 03/014342

PCT/EP02/08963

35

12. Lloyd, M *et al.* Regulation of intestinal lactase in adult hypolactasia. *J. Clin. Invest.* **89**, 524-529 (1992).
- 5 13. Fajardo, O., Naim, H.Y. & Lacey, S.W. The polymorphic expression of lactase in adults is regulated at the messenger RNA level. *Gastroenterology* **106**, 1233-1238 (1994).
14. Luigi, M. *et al.* Mosaic regulation of lactase in human adult-type *Gastroenterology* **112**, 1506-1514 (1997).
- 10 15. Rössli, M. *et al.* Lactase persistence versus decline in human adults: Multifactorial events are involved in down-regulation after weaning. *Gastroenterology* **112**, 1506-1514 (1997).
16. Göring, H.H.H. & Terwilliger, J.D. Linkage analysis in the presence of errors IV: Joint pseudomarker analysis of linkage and *r* or linkage disequilibrium on a mixture of pedigrees and singletons when mode of inheritance cannot be accurately specified. *Am. J. Hum. Genet.* **66**, 1310-1327 (2000).
- 20 17. Terwilliger, J.D. & Göring, H.H.H. Gene mapping in the 20th and 21st centuries: Statistical methods, data analysis, and experimental design. *Hum. Biol.* **72**, 63-132 (2000).
18. Harvey, C.B. *et al.* Regional localization of the lactase-phlorizin hydrolase, LCT, to chromosome 2q21. *Ann. Hum. Genet.* **57**, 179-185 (1993).
- 25 19. Syvänen, A.-C., Sajantila, A., Lukka, M. Identification of individuals by analysis of biallelic DNA markers, using PCR and solid-phase minisequencing. *Am. J. Hum. Genet.* **52**, 46-59 (1993).
- 10 20. Syvänen, A.-C. & Landegren, U. Detection of point mutations by solid-phase methods. *Hum. Mutat.* **3**, 172-179 (1994).
21. Thompson, E. A., Deeb, S., Walker, D. & Motulsky, A. G. The detection of Linkage disequilibrium between closely linked markers: RFLPs at the A1-CIII Apolipoprotein genes. *Am. J. Hum. Genet.* **42**, 113-124 (1988).
- 5 22. Dausset, J. *et al.* Centre d'étude du polymorphisme humain (CEPH): Collaborative genetic mapping of human genome. *Genomics* **6**, 575-577 (1990).
23. Cuddenech, Y., Delbrück, H. & Flatz, G. Distribution of the adult lactase phenotypes- lactose absorber and malabsorber-in a group of 131 army recruit *Gastroenterol. Clin. Biol.* **6**, 776-779 (1982).
- 5 24. McLellan, T., Jorde, L.B. & Skolnick, M.H. Genetic distance between the Utah Mormons and related populations. *Am. J. Hum. Genet.* **36**, 836-857 (1984).
- 25 25. Terwilliger, J.D. A powerful likelihood method for the analysis of linkage

WO 03/014342

PCT/EP02/08963

36

- disequilibrium between trait loci and one or more polymorphic marker loci. *Am. J. Hum. Genet.* **56**, 777-787 (1995).
- 5 26. Nunez, M.G. A model of the early settlement of Finland. *Fennoscandia archaeologica* IV, 3-18 (1997).
27. Simoons, F.J. Primary adult lactose intolerance and the milking habit: a problem in biological and cultural interrelations.II. A cultural historical hypothesis. *Am. J. Dig. Dis.* **16**, 695-710 (1970).
- 10 28. Varilo, T. *et al.* The age of human mutation: genealogical and linkage disequilibrium analysis of the CLN5 mutation in the Finnish population. *Am. J. Hum. Genet.* **58**, 506-512 (1996).
29. Hästbacka, J. *et al.* Linkage disequilibrium mapping in isolated founder populations: diastrophic dysplasia in Finland. *Nature Genet.* **2**: 204-211 (1992).
30. Harvey C.B. *et al.* Lactase haplotype frequencies in Caucasians: association with the lactase persistence/non persistence polymorphism. *Ann Hum Genet* **62**, 215-223 (1998).
- 20 31. Ohtani, K. *et al.* Cell growth-regulated expression of mammalian MCM5 and MCM6 genes mediated by the transcription factor E2F. *Oncogene* **18**, 2299-2309 (1999).
- 25 32. Smith, A.F.A. The origin of interspersed repeats in the human genome. *Curr. Opin. Genet. Dev.* **6**, 743-748 (1996).
33. Kazazian, H.H. & Moran, J.V. The impact of L1 retrotransposons on the human genome. *Nature Genet.* **19**, 19-24 (1998).
- 30 34. Moran, J.V., DeBerardinis, R.J. & Kazazian, H.H. Exon shuffling by L1 retrotransposition. *Science* **283**, 1530-1534 (1999).
- 35 35. Wei, W. *et al.* Human L1 retrotransposition: *cis* preference versus *trans* complementation. *Mol. Cell. Biol.* **21**, 1429-1439 (2001).
36. Donnelly, S.R., Hawkins, T.E. & Moss, S.E. A conserved nuclear element with a role in mammalian gene regulation. *Hum. Mol. Genet.* vol.8, 9, 1723-1728 (1999).
- 40 37. Boeke, J.D. LINEs and Alus – the polyA connection. *Nature Genet.* **16**, 6-7 (1997).
- 45 38. Jurka, J. Sequence patterns indicate an enzymatic involvement in integration of mammalian retrotransposons. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 1872-1877 (1997).
39. Savilahti E, Launiala K, Kuitunen P. Congenital lactase deficiency. *Arch. Dis. Child.* **58**, 246-252 (1983).

WO 03/014342

PCT/EP02/08963

37

40. Järvelä, I. *et al.* Assignment of the locus for congenital lactase deficiency to 2q21, in the vicinity of but separate from the lactase-phlorizin hydrolase gene. *Am. J. Hum. Genet.* **63**, 1078-1085 (1998).
- 5 41. Simoons, F. J. The geographic hypothesis and lactose malabsorption. A weighing of the evidence. *Am. J. Dig. Dis.* **23**, 963-980 (1978).
- 10 42. Flatz, G. & Rothauwe, H. W. The human lactase polymorphism: physiology and genetics of lactose absorption and malabsorption. *Prog. Med. Genet.* **2**, 205-249 (1977).
43. McCracken, R.D. Lactase deficiency: an example of dietary evolution. *Curr. Anthropol.* **12**, 479-517 (1971).
44. Arola, H. *et al.* Diagnosis of hypolactasia and lactose malabsorption. *Scand. J. Gastroenterol. Suppl.* **202**, 26-35 (1994).
- 20 45. Sulkanen, S. *et al.* Tissue transglutaminase autoantibody enzyme-linked immunosorbent assay in detecting celiac disease. *Gastroenterology* **115** (6), 1322-1328 (1998).
- 25 46. Sambrook, J., Fritsch, E.F. & Maniatis, T. *Molecular cloning: a laboratory manual*, (2nd ed). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989).
47. Messer, M. & Dahlqvist, A. A one - step ultramicro method for the assay of intestinal disaccharidases. *Anal. Biochem.* **14** (3), 376-92 (1966).
- 30 48. Cottingham, Jr. R.W., Idury, R.M. & Schaffer, A.A. Faster sequential genetic linkage computations. *Am. J. Hum. Genet.* **53**, 252-263 (1993).
- 35 49. Göring, H.H.H. & Terwilliger, J.D. Linkage analysis in the presence of errors III: Marker loci and their map as nuisance parameters. *Am. J. Hum. Genet.* **66**, 1298-1309 (2000).
50. Terwilliger, J.D. & Ott, J. Hand book of human genetic analysis. *Johns Hopkins University Press*, Baltimore (1994).
- 40 51. Osborne *et al.*, *Curr. Opin. Chem. Biol.* **1** (1997), 5-9
52. Stall and Szoka, *Pharm. Res.* **12** (1995), 465-483
- 15 53. Harlow and Lane „Antibodies, A Laboratory Manual“, CSH Press, Cold Spring Harbor, USA, 1988
54. Higgins and Hames (eds.), "Nucleic acid hybridization, a practical approach", IRL Press, Oxford 1985

WO 03/014342

PCT/EP02/08963

38

55. Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, Marc S, Hazan J, Seboun E, Lathrop M, Gyapay G, Morissette J, Weissenbach J. A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature*. 1996 Mar 14;380(6570):152-4.
- 5 56. Chumakov IM, Rigault P, Le Gail I, Bellanne-Chantelot C, Billault A, Guillou S, Soularue P, Guasconi G, Poullier E, Gros I, et al. A YAC contig map of the human genome. *Nature*. 1995 Sep 28;377(6547 Suppl):175-297

**Table 1. Linkage and Linkage Disequilibrium Analyses in adult-type hypolactasia families (fine mapping markers shown in bold)**

Marker	Lod score(Z) at $\Theta$					<i>p</i> -value <sup>a</sup>
	0.0	0.1	0.2	0.3	0.4	
D2S114	$\infty$	2.44	1.92	1.13	<b>0.41</b>	0.87195
<b>P6112</b>	2.76	2.20	1.45	0.75	<b>0.22</b>	0.66207
<b>D2S1334</b>	3.15	2.45	<b>1.61</b>	0.84	0.25	0.91039
AC8	<b>2.26</b>	1.99	1.36	0.71	0.21	0.53670
LPH13	3.67	2.94	1.96	1.03	0.31	4x10 <sup>-6</sup>
LPH2	4.09	3.07	2.00	1.00	0.26	5.7x10 <sup>-7</sup>
<b>LPH1</b>	5.91	4.52	2.96	1.53	0.46	5x10 <sup>-6</sup>
AC7	3.63	2.60	1.66	0.83	0.23	0.03471
<b>AC3</b>	6.63	4.88	3.16	1.61	0.44	3.2x10 <sup>-8</sup>
AC4	3.07	2.22	1.42	0.71	0.19	4x10 <sup>-5</sup>
<b>AC5</b>	5.33	4.10	2.72	1.39	0.39	0.02166
<b>AC10</b>	6.60	4.99	3.25	1.65	0.46	1x10 <sup>-5</sup>
D2S2196	7.67	5.62	3.62	1.85	0.54	0.00010
D2S442	3.81	3.08	2.08	1.03	0.27	0.22805
D2S314	4.22	3.61	2.50	1.37	0.45	0.27535
D2S2385	$\infty$	2.79	1.92	1.01	0.28	0.46457

a: *p*-values produced using linkage disequilibrium test given linkage<sup>10,49</sup>

**Table 2. The variations identified within adult-type hypolactasia locus in the Finnish Families**

Position <sup>a</sup>	Variant	Lactase persistence (Homozygous)		Lactase persistence (Heterozygous)		Lactase non-persistence		
		BIV4	AIV3	BIV8	CIV3	BIV9	DIV4	EIII2 <sup>b</sup>
-694	A→G	AA	AA	AG	AA	GG	N <sup>c</sup>	AA
-1640/50	T <sub>13</sub> →T <sub>12</sub>	T <sub>13/13</sub>	T <sub>13/13</sub>	T <sub>13/13</sub>	T <sub>13/13</sub>	T <sub>13/13</sub>	T <sub>12/12</sub>	T <sub>12/12</sub>
-2131	C→T	CC	CC	CT	CC	TT	CT <sup>e</sup>	TT
-3058/72	T <sub>15</sub> →T <sub>16</sub>	T <sub>15/15</sub>	T <sub>15/15</sub>	T <sub>15/15</sub>	T <sub>15/15</sub>	T <sub>15/15</sub>	T <sub>16/16</sub>	T <sub>16/16</sub>
-3075	G→T	GG	GG	GG	GG	GG	GG	TT
-4480	T→A	TT	TT	TA	TT	AA	TT	TT
-5440	C→T	CC	CC	CT	CC	TT	CC	CC
-5926	A→T	AA	AA	AA	AA	AA	TA	TT
-8540	G→A	GG	GG	GA	GA	AA	AG	AA
-8630	C→G	CC	CC	CG	CG	GG	GC	GG
-13495	T→C	TT	TT	TC	TT	CC	CT	CC
-13910	T→C	TT	TT	TC	TC	CC	CC	CC
-15239	G→A	GG	GG	GA	GG	AA	AG	AA
-15862	T→C	CC	CC	CT	CC	TT	TC	TT
-16568/79	T <sub>11</sub> →T <sub>12</sub>	T <sub>11/11</sub>	T <sub>11/11</sub>	T <sub>11/12</sub>	T <sub>11/11</sub>	T <sub>12/12</sub>	T <sub>11/11</sub>	T <sub>12/12</sub>
-16888	A→G	AA	AA	GA	AA	GG	GA	GG
-17300	C→T	CC	CC	CC	CC	CC	CT	TT

WO 03/014342

PCT/EP02/08963

41

-19044	T→C	TT	TT	TC	TT	CC	CT	CC
-19519	T→C	TT	TT	TC	TT	CC	TT	TT
-20077	C→G	CC	CC	CG	CC	GG	GC	GG
-20486	G→A	GG	GG	GA	GG	AA	GG	GG
-21721/28	A <sub>7</sub> →A <sub>6</sub>	A <sub>77</sub>	A <sub>77</sub>	A <sub>77</sub>	A <sub>77</sub>	A <sub>77</sub>	A <sub>7/A<sub>6</sub></sub>	A <sub>77</sub>
-21731	A→C	AA	AA	AA	AA	AA	CC	AA
-21736/43	A <sub>9</sub> →A <sub>8</sub>	A <sub>95</sub>	A <sub>95</sub>	A <sub>9/A<sub>8</sub></sub>	A <sub>95</sub>	A <sub>8/8</sub>	A <sub>8/8</sub>	A <sub>8/8</sub>
-22018	G→A	AA	AA	AG	AG	GG	GG	GG
-22741	C→T	CC	CC	CC	CC	CC	N	TT
-22788	A→G	AA	AA	AG	AA	GG	N	GG
-23069	A→G	AA	AA	AG	AA	GG	N	GG
-23442	A→G	AA	AA	AA	AA	AA	N	GG
-23771	T→C	TT	TT	TT	TT	TT	N	CC
-25093/23	Δ30 bp	ΔΔ	ΔΔ	ΔΔ	ΔΔ	ΔΔ	N	II
-27310	A→G	AA	AA	AG	AA	GG	GA	GG
-27480	G→A	GG	GG	GA	GG	AA	AG	AA
-27807	A→C	AA	AA	AA	AA	AA	AC	CC
-30183	A→G	AA	AA	AG	AA	GG	AA	AA
-31268	A→G	AA	AA	AG	AA	GG	AA	AA
-31342	T→C	TT	TT	TT	TT	TT	CT	CC
-33645	C→T	CC	CC	CT	CC	TT	CC	CC
-35176	T→C	TT	TT	TC	TT	CC	CT	CC
-36254	C→T	CC	CC	CT	CC	TT	TC	TT

WO 03/014342

PCT/EP02/08963

42

-36296	G→T	TT	TT	TG	TT	GG	TG	N
-36501	A→T	AA	AA	AT	AA	TT	AT	N
-36506/14	Δ 9 bp	ΔΔ	ΔΔ	Δ I	ΔΔ	II	ΔI	N
-36671/77	T7→T6	T <sub>77</sub>	T <sub>77</sub>	T <sub>76</sub>	T <sub>77</sub>	T <sub>66</sub>	T <sub>77</sub>	T <sub>77</sub>
-37565	T→G	TT	TT	TG	TT	GG	GG	TG
-38276	G→C	GG	GG	GC	GG	CC	GG	GG
-39036	G→C	GG	N	GC	N	CC	N	N
-40608	G→C	GG	GG	GG	GG	GG	GC	CC
-41590	T→C	TT	TT	TC	TT	CC	CT	CC
-42081/82	ΔAG	AG	AG	AG/Δ	AG	ΔΔ	AG	AG
-42618	T→C	TT	TT	TC	TT	CC	TT	TT
-42893	G→A	GG	GG	GA	GG	AA	GG	GG

a: The Number is from initiation translation codon (ATG) of the LPH gene using the compiled genomic sequence of the BACs NH034L23, NH0218L22, NH0318L13 and RP11-329I10, b: the individuals sequenced from the Finnish families studied and showed by arrow in fig.1, c: not determined

Table 3. Distribution of C/T<sub>-13910</sub> & G/A<sub>-22018</sub> genotypes in lactase persistent/non-persistent alleles

	Genotype	C/T <sub>-13910</sub>			G/A <sub>-22018</sub>			Total
		CC	CT	TT	GG	GA	AA	
Family members	Lactase non-persistence	45	0	0	45	0	0	45
	Lactase persistence	0	32	13	0	32	13	45
Case-control samples								
Finnish	Lactase non-persistence	59	0	0	53	6	0	59
	Lactase persistence	0	63	74	0	63	74	137
Non-Finnish <sup>a</sup>	Lactase non-persistence	40	0	0	39	1	0	40
	Lactase persistence	0	5	0	0	5	0	5
Total	Lactase non-persistence						0	144
	Lactase persistence							187

a: non-Finnish samples consist of 23 South Korean, 9 Italian and 7 German individuals

WO 03/014342

PCT/EP02/08963

44

**Table 4. Prevalence of the C/T-13910 variant in population samples**

DNA samples analysed	Genotype			Total	Allele frequency(%)		% (CC) genotype
	CC	CT	TT		C	T	
I. Finnish population:							
1. Eastern regions	108	287	176	571	0.440	0.560	18.9%
2. Western regions	62	159	146	367	0.385	0.615	16.8%
Total	170	446	322	938	0.418	0.582	18.1%
II. CEPH parents:							
1. Utah families	7	33	52	92	0.255	0.745	7.6%
2. French families	7	9	1	17	0.676	0.324	41.2%

A total of 938 DNA samples of anonymous Finnish blood donors from small parishes from Eastern and Western parts within Finland, and 109 DNA samples from CEPH parents. The prevalence of hypolactasia in the populations is reflected by the genotype frequencies of CC alleles.

Table 5. LD between C/T-13910 and G/A-220018 variants in random Finnish samples

	Genotype at C/T-13910			Total	D	$\chi^2(1 \text{ df})$	P-value
	CC	CT	TT				
Genotype at G/A-220018							
GG	162	2	1	165			
GA	6	440	3	449			
AA	2	4	318	324			
Total	170	446	322	938	0.984	42.41	$7.62 \times 10^{-11}$

LD was calculated using D' statistic<sup>18</sup>, p value is the significance of D' from 0 as described in methods<sup>18</sup>.

**Table 6. Estimation of the introduction of the C/T-13910 variant into Finnish population using DISLAMB program.**

Marker	AC3		LPH2	
	Lactase persistence	Lactase non-persistence	Lactase persistence	Lactase non-persistence
1	0	1	0	1
2	31	10	0	20
3	0	1	0	14
4	2	9	32	15
5	0	31	0	2
$\lambda^a$	0.838		0.999	
$\Theta^b$	0.00031 (0.000038-0.00099)		0.0000(0.00000-0.00052)	
$n^c$	570		450	

a:  $\lambda$  is the proportion of increase of a certain allele in disease chromosomes (lactase persistence allele) relative to its population frequency( 0.60). b:  $\Theta$  is the recombination fraction , reflected by the distance of the mutation from the closest marker, assuming 1cM= 1Mb. C: n is the number of generation since the introduction of the founder mutation into a population Applying  $\lambda = \infty (1-\Theta)^n$  formula. d: Hypothetical allele used in the calculations as  $\Theta$  is zero and  $\infty$  is one.

WO 03/014342

PCT/EP02/08963

47

Table 7 Prevalence of lactose intolerance variants in biochemically verified samples

Population	Number	C/T <sub>13910</sub>				G/A <sub>22918</sub>			
		CC	CT	TT	GG	GA	AA		
1. Finnish	Lactase persistence	0	95	87	0	95	87		
	Lactase non-persistence	116	0	0	110	6	0		
2. Italian	Lactase persistence	7	0	7	0	7	0		
	Lactase non-persistence	23	0	0	22	1	0		
3. German	Lactase persistence	0	0	0	0	0	0		
	Lactase non-persistence	8	0	0	8	0	0		
4. Somalian	Lactase persistence	0	0	0	0	0	0		
	Lactase non-persistence	42	0	0	42	0	0		
6. South Koreans	Lactase persistence	0	0	0	0	0	0		
	Lactase non-persistence	23	0	0	23	0	0		
Total	401	212	102	87	205	109	87		

**Table 8** Prevalence of lactose-intolerance variants in various population samples

Population	Number	Genotype				Gi/A22018			%Prevalence of Lactase Persistence allele
		C/T13910		T/T		GG	GA	AA	
		CC	CT	TT					
South Koreans	23	23	0	0	23	0	0	0	
France	17	7	9	1	6	10	1	59	
Basques	85	7	44	34	13	35	37	92	
Southern Italians	100	89	11	0	88	12	0	11	
Somalians	79	74	5	0	78	1	0	6	
Utah	92	7	33	52	7	30	55	92	
African-Americans	96	76	15	5	78	12	5	21	
Marroccans	90	62	25	3	65	22	3	31	
Serawhi (African)	57	29	26	2	28	26	3	49	
Saami	30	20	10	0	21	9	0	33	
Tibet	23	23	0	0	23	0	0	0	
Eastern Finnish	571	108	287	176	107	288	176	81	
Western Finnish	367	62	159	146	58	161	148	83	
Finn-Ligrian tribes									
Xen	20	19	1	0	19	1	0	5	
Xm	20	19	1	0	19	1	0	5	
Imansi	22	20	2	0	20	2	0	9	
Likomi	10	7	3	0	7	3	0	30	
Erza	30	17	10	3	19	9	2	43	
Moksa	30	13	17	0	14	16	0	57	
Udmort	30	12	16	2	11	15	4	60	
Pakistani tribes									
Kalash	30	30	0	0	28	2	0	0	

Table 8 cont

Burusho	30	29	1	0	27	3	0	3
Hazara	14	13	1	0	11	3	0	7
Kashmiri	20	15	5	0	14	6	0	25
Makrani Baluch	29	19	10	0	19	8	1	34
Brajhui	30	17	10	3	16	11	3	43
Makrani (Negroid)	29	16	10	3	16	10	3	45
Pathan	29	12	16	1	13	14	2	59
Indian	29	11	13	5	10	12	5	62
<b>Total</b>	<b>2032</b>							

\*The prevalence of lactase persistence allele is correlated very well with the reported prevalences for the lactase persistence allele (Simoons FJ. The geographic hypothesis and lactose malabsorption Am J Dig Dis 1978 23 (1):963-80)

WO 03/014342

PCT/EP02/08963

50

## Claims

1. A nucleic acid molecule comprising a 5' portion of an intestinal lactase-phlorizine hydrolase (LPH) gene contributing to or indicative of adult-type hypolactasia wherein said nucleic acid molecule is selected from the group consisting of
- 5
- (a) a nucleic acid molecule having or comprising the nucleic acid sequence of SEQ ID NO: 1, the sequence of SEQ ID NO:1 is also depicted in Fig. 4 and comprised in the sequence as depicted in Fig. 8;
- 10
- (b) a nucleic acid molecule having or comprising the nucleic acid sequence of SEQ ID NO: 2, the sequence of SEQ ID NO:2 is also depicted in Fig. 5 and comprised in the sequence as depicted in Fig. 9;
- 15
- (c) a nucleic acid molecule of at least 20 nucleotides the complementary strand of which hybridizes under stringent conditions to the nucleic acid molecule of (a) or (b), wherein said polynucleotide/nucleic acid molecule has at a position corresponding to position -13910 5' from the LPH gene a cytosine residue; and
- 20
- (d) a nucleic acid molecule of at least 20 nucleotides the complementary strand of which hybridizes under stringent conditions to the nucleic acid molecule of (a) or (b), wherein said polynucleotide/nucleic acid molecule has at a position corresponding to position -22018 5' from the LPH gene a guanine residue.
- 25
2. A nucleic acid molecule comprising a 5' portion of an intestinal lactase-phlorizine hydrolase (LPH) gene wherein said nucleic acid molecule is selected from the group consisting of
- (a) a nucleic acid molecule having or comprising the nucleic acid sequence of SEQ ID NO:3, the sequence of SEQ ID NO:3 is also depicted in Fig. 6;
- 30
- (b) a nucleic acid molecule having or comprising the nucleic acid sequence of SEQ ID NO:4, the sequence of SEQ ID NO:4 is also depicted in Fig. 7;
- (c) a nucleic acid molecule the complementary strand of which hybridizes

WO 03/014342

PCT/EP02/08963

51

- under stringent conditions to the nucleic acid molecule of (a) or (b), wherein said polynucleotide/nucleic acid molecule has at a position corresponding to position -13910 of the LPH gene a thymidine residue; and
- 5 (d) a nucleic acid molecule the complementary strand of which hybridizes under stringent conditions to the nucleic acid molecule of (a) or (b), wherein said polynucleotide/nucleic acid molecule has at a position corresponding to position -22018 of the LPH gene a adenosine residue.
- 10 3. The nucleic acid molecule of claim 1 or 2 which is genomic DNA.
4. The nucleic acid molecule of claim 3 wherein said genomic DNA is part of a gene.
- 15 5. A fragment of the nucleic acid molecule of any one of claims 1 to 4 having at least 14 nucleotides wherein said fragment comprises nucleotide position -13910 or nucleotide position -22018 of the LPH gene.
6. A nucleic acid molecule which is complementary to the nucleic acid molecule  
20 of any one of claims 1 and 3 to 5.
7. A nucleic acid molecule which is complementary to the nucleic acid molecule of any one of claims 2 to 5.
- 25 8. A vector comprising the nucleic acid molecule of any one of claim 1 and 3 to 5.
9. A vector comprising the nucleic acid molecule of any one of claims 2 to 4.
- 30 10. A primer or primer pair, wherein the primer or primer pair hybridizes under stringent conditions to the nucleic acid molecule of any one of claims 1 and 3 to 5 comprising nucleotide position -13910 or -22018 of the LPH gene or to the complementary strand thereof.

WO 03/014342

PCT/EP02/08963

52

11. A primer or primer pair, wherein the primer or primer pair hybridizes under stringent conditions to the nucleic acid molecule of any one of claims 2 to 5 comprising nucleotide position -13910 or -22018 of the LPH gene or to the complementary strand thereof.
- 5
12. A non-human host transformed with the vector of claim 6.
13. A non-human host transformed with the vector of claim 7.
- 10 14. The non-human host of claim 12 or 13 which is a bacterium, a yeast cell, an insect cell, a fungal cell, a mammalian cell, a plant cell, a transgenic animal or a transgenic plant.
- 15 15. An antibody or aptamer or phage that specifically binds to the wild-type nucleic acid molecule of any one of claims 1 and 3 to 6 but not to the corresponding wild-type nucleic acid molecule
- 20 16. An antibody or aptamer or phage that specifically binds to the wild-type nucleic acid molecule of any one of claims 2 to 5 and 7 but not to the corresponding mutant sequence contributing to or indicative of adult-type hypolactasia.
7. A pharmaceutical composition comprising the wild-type nucleic acid molecule of claim 2, 3, 4 or the vector of claim 9.
- 25 18. A diagnostic composition comprising the nucleic acid molecule of any one of claims 1 to 7, the vector of claim 8 or 9, the primer or primer pair of claim 11 or 12, and/or the antibody aptamer and/or phage of claim 15 or 16.
- 30 19. A method for testing for the presence or predisposition of adult-type hypolactasia or associated trait comprising testing a sample obtained from a prospective patient or from a person suspected of carrying such a predisposition for the presence of the nucleic acid molecule of any one of claims 1 and 3 to 6 in a homozygous or heterozygous state.

WO 03/014342

PCT/EP02/08963

53

20. A method for testing for the presence or predisposition of adult-type hypolactasia or associated trait comprising testing a sample obtained from a prospective patient or from a person suspected of carrying such a predisposition for the presence of the nucleic acid molecule of any one of claims 2 to 5 and 7 in a homozygous or heterozygous state.
21. The method of claim 19 or 20, wherein said testing comprises hybridizing the complementary nucleic acid molecule of claim 6 which is complementary to the nucleic acid molecule contributing to or indicative of adult-type hypolactasia or the nucleic acid molecule of claim 7 which is complementary to the wild-type sequence as a probe under stringent conditions to nucleic acid molecules comprised in said sample and detecting said hybridization.
22. The method of any one of claims 19 or 21 further comprising digesting the product of said hybridization with a restriction endonuclease or subjecting the product of said hybridization to digestion with a restriction endonuclease and analyzing the product of said digestion.
23. The method of claim 21, wherein said probe is detectably labeled.
24. The method of claim 19 or 20, wherein said testing comprises determining the nucleic acid sequence of at least a portion of the nucleic acid molecule of any one of claims 1 to 7, said portion comprising nucleotide position -13910 and/or nucleotide position -22018 of the LPH gene.
25. The method of claim 24, wherein the determination of the nucleic acid sequence is effected by solid-phase minisequencing.
26. The method of claim 24 further comprising, prior to determining said nucleic acid sequence, amplification of at least said portion of said nucleic acid molecule.

WO 03/014342

PCT/EP02/08963

54

27. The method of claim 19 or 20, wherein said testing comprises carrying out an amplification reaction wherein at least one of the primers employed in said amplification reaction is the primer of claim 10 or belongs to the primer pair of claim 10, comprising assaying for an amplification product.
- 5
28. The method of claim 19 or 20, wherein said testing comprises carrying out an amplification reaction wherein at least one of the primers employed in said amplification reaction is the primer of claim 11 or belongs to the primer pair of claim 11, comprising assaying for an amplification product.
- 10
29. The method of any one of claims 26 to 28 wherein said amplification is effected by or said amplification is the polymerase chain reaction (PCR).
30. A method for testing for the presence or predisposition of adult-type hypolactasia comprising assaying a sample obtained from a human for specific binding to the antibody or aptamer or phage of claim 15
- 15
31. A method for testing for the presence or predisposition of adult-type hypolactasia comprising assaying a sample obtained from a human for specific binding to the antibody or aptamer or phage of claim 16.
- 20
32. The method of claim 30 or 31, wherein said antibody or aptamer or phage is detectably labeled.
- 25
33. The method of any one of claims 30 to 32, wherein the test is an immunoassay.
- 30
34. The method of any one of claims 19 to 33, wherein said sample is blood, serum, plasma, fetal tissue, saliva, urine, mucosal tissue, mucus, vaginal tissue, fetal tissue obtained from the vagina, skin, hair, hair follicle or another human tissue.

WO 03/014342

PCT/EP02/08963

55

35. The method of any one of claims 19 to 34, wherein said nucleic acid molecule from said sample is fixed to a solid support.
36. The method of claim 35, wherein said solid support is a chip, a silica wafer, a bead or a microtiter plate.
37. Use of the nucleic acid molecule of any one of claims 1 to 7 for the analysis of the presence or predisposition of adult-type hypolactasia.
38. Kit comprising the nucleic acid molecule of any one of claims 1 to 7, the primer or primer pair of claim 11 or 12, the vector of claim 8 or 9, and/or the antibody aptamer and/or phage of claim 15 or 16 in one or more containers.
39. Use of the nucleic acid molecule of any one of claims 2 to 4 or the vector of claim 7 in gene therapy.
40. The use of claim 39, wherein said gene therapy treats or prevents adult-type hypolactasia.

WO 03/014342

PCT/EP02/08963

1 / 10

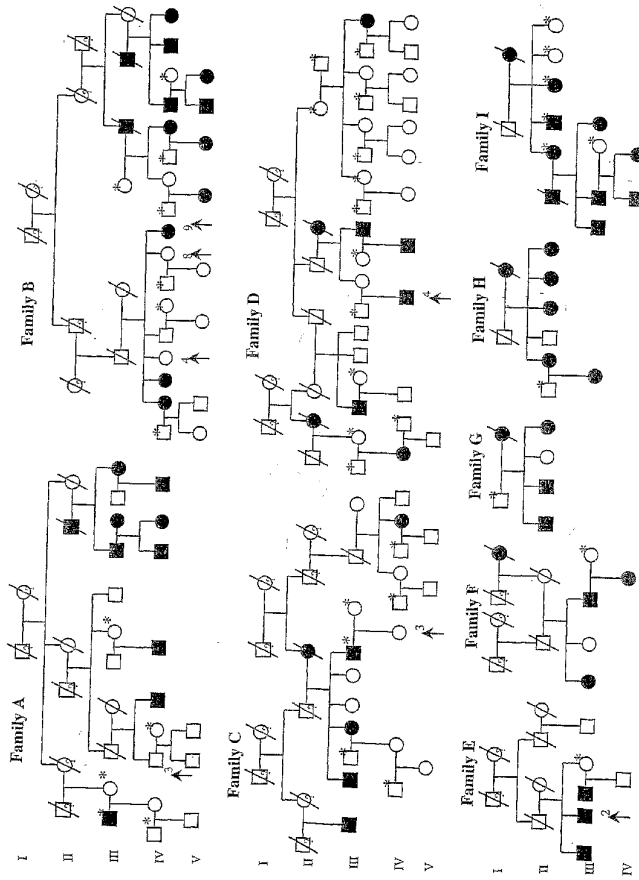


Figure 1

2/10

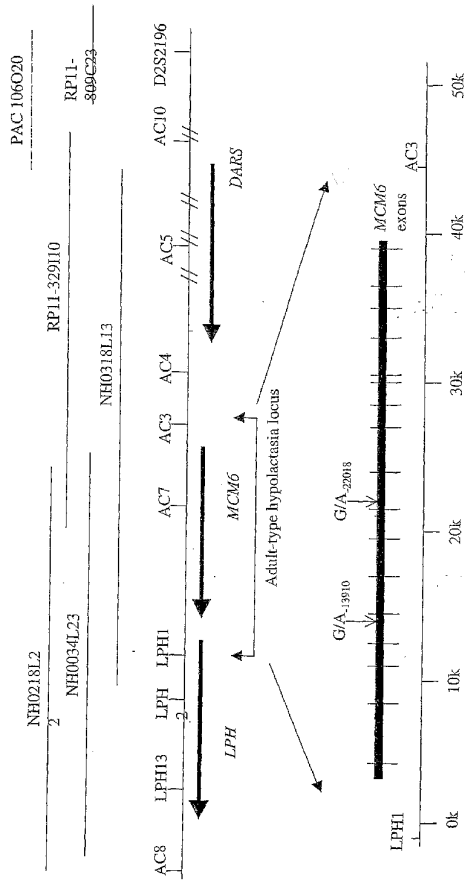


Figure 2

3/10

Marker	Haplotype						Total
	6	6	6	6	6	6	
LPH12	6	6	6	6	6	6	6
LPH2	4	4	4	4	4	4	4
LPH1	3	2	2	1	1	4	4
AC7	5	5	5	5	5	5	5
AC3	2	2	4	2	2	2	2
AC4	5	5	5	5	5	5	5
AC5	6	6	5	6	6	3	5
Lactase persistent alleles	20	4	1	2	2	1	1
Lactase non-persistent alleles	3	2	0	0	0	0	2
							0
							2
							0
							2
							54

Figure 3

WO 03/014342

PCT/EP02/08963

4 / 10

## SEQ ID NO:1

ACCTTTCATTCAGGAAAAATGTACTTAGACCCCTACAATGTACTAGTAGGCCCTCTGCCCTG  
GCAATACAGATAAGATAAATGTAGcCCCTGGCCCTCAAAGGAACCTCCTCCTTAGGTTGCA  
TTTGTATAAATGTTGATTTTAGATTGTTCTTTGAGCCCTGCATTCACGAGGATAGGTC

Figure 4

## SEQ ID NO:2

TAAGAACATTTTACACTCTTCAGTATAAAGAAGTCAGAATACCCCTACCCCTATCAGTAAA  
GGCCTATAAGTTACCATTAATAAGATGTCCTTAAAAACAGCATTCACAGCTGGGCgCGGT  
GGCTCACACCTTTGTCCAGTACTTTGGGAAGCCGAGGTGGGTGGATCACCTGAGGTCAG

Figure. 5

WO 03/014342

PC1/EP02/08963

5 / 10

SEQ ID NO:3

ATCAGAGTCACCTTTGATATGATGAGAGCAGAGATAAACAGATTTGTTGCATGTTTTTAAT  
 CTTTGGTATGGGACATACTAGAAATTCACCTGCAANTACATTTTTATGTAACCTGTTGAATGC  
 TCATACGACCCATGGAAATTCCTCCCTTTAAAGAGCTTGGTAAGCATTTGAGTGTAGTTGTT  
 AGACGGAGACGATCACGTCATAGTTTATAGAGTGCATAAAGACCTAAGTTACCATTTAAT  
 ACCTTTCATTCAGGAAAATGTAAGTACTTAGACCCPACAAATGTAAGTACTAGGCTCTGCGCTG  
 GCAATACAGATAAGATAATGTAAGTCCCTGGCTCAAAGGAACTCCCTCCTTAGGTTGCA  
 TTTGTAATATTTGATTTTTAGATTTGTTCTTTGAGCCCTGCATPCCACGAGGATAGGTC  
 AGTGGGTATTAACGAGGTAAAAGGGGAGTAGTACGAAAAGGGCATPCAAGCGTCCCATCTT  
 CGCTTCAACCAAAGCAGCCCTGCGTTTTTCCTAGTTTTAATAAAGGTTTGATTAAGGTC  
 GTCCTTCAAAGGGGGTTGGCTTTTTTTACAGTGTGACTGAGGTATAATTTATAAAAA  
 GGGAAATGTAAGGCAATGGTGGTTTTTTTACATACATCCCTTGTGAATACCAGCTCAAGA  
 TCCAAAACATTTCCATAATTTAGAAAGTTCCAAACCCCTGCTCTTTTCAGTCTTAGCC  
 CTCTTCCCTGAAGTAACCACTGTTCCGACTTCAATCACTACTTTTTATCCCACAGGTTAA  
 TTTTTTGGCTTTTTTCCACTAAATTTTCAAATTCCTTGATATGGTACTTTACTATTGAGG  
 AAGTACTTTCACACTAGGTTAATTAATATCTTTGATTCACCCAAATTTAGGGAAACACC  
 TGTAGGGGACAAAAATGAATGAGAGCCCTGCTTCCATTTGCTGCTAATCTGGTGGAA  
 CGAGACATGTAATTAATTAAGCATGTAAAAAATAGAGTGGGTGATGAATTAATCTATATA  
 CTAATCCCCTGACACACAGTTTACCTATGTAACAAACCTGCATGTTGACCCCGAAC  
 TAAAAATATAAGTTGGAAATAAAAAAGAGAGGGGAAATAGAGCATCACAAACAGAG  
 TGCTGAGATGAATTAATTTATACCAAGAGGAGGAGGACTCAGGGAGGTGCCAGCTT  
 TAAACCCAGTCACTGAAGGGTGTGCAGAAATTTGGATAGGCRAGRTACCCCTGGGACAAGGT  
 CATTTCAAACCAATGCTAACATTTTACTTTTTTTTTCAFTGTGATAGTTCCGAAATGA  
 GTTGCATAAACTGGTACATCTCTTAGGGCAGTCTCAATGATTTTTATTTTGTCTAT  
 TTTTAAAAAATAGTCTTCAAATAGCAGATTCACATGATATAAATATATGCACATAAAT  
 TATATACACAAATATATTTCTGAATGAATTTAGTATCTGCATATATTTAAGAGCTATT  
 TCTGTCTCATATGTTCAATACTTCACTCCATTAAAAAAATTTTGTAGGCCCTTCTCAC  
 TCTAAGATTAAAAAATTTCTCCCATTAATTTACCTAGCTAGTTTTCTAGTTGTTCCAAAA  
 CCATTTATGAAACAACTCATCTTTTGCACCTGGTTTGGCCTGCTTAATATATATTTCT  
 TGTGTGTGTAGGATCTCCTTTTGGACTTTCCATTTCTGTCTATGAGTCTTATCAGCTCC  
 TCTTACATGGTACCATGATGTTTTAATCTATGGGCTTTGTAGTTTAAATGTAGGGCTA  
 GTTCCAGGCAATGTTCTCTATCAGCTGTAGGAACTTAGAAATCAGCTTGTCTGTTTT  
 AAGAAAAAACCCTGGTATTTTTTTTACAGTATAACATTTCTATTTATTAACTTGAAGAAT  
 TGAACCAATCTATGATTTTTCTTATTCAGTAACTGATCACTTGAATAGGTTAGGTTGTA  
 CTACTATAAAATCTCAGCTGCATAAAACAATTTTTTTTTGCTTGTCTACACATCCATTA  
 GGTCAATCAAGGACTCACCTTGTCAAGTTACTCAGAGATTCAGGCTGATATAAAGGTTTG  
 ATCTTGACATACGCTTTCTATGATGACAGAAAGCAGGGAAGAGAAGGTGCTGAGCCATGTG  
 CTTTCTCCCCCTTCTATCCAGAAATGACACATACTCACATTTTCATTCGCCAGAGAAATTA  
 ACATGGCCCCCTTAAGTTCAAATGGATAGAGAAATGCTTCCCTACCAGGTGCCAGAAAT  
 TAGAGAGCAAAACATTTGTGAACAGTTCTGAGTACCACAAATACCGTTATCTTCCACTT  
 AAGTCTTCTGTTCACTCAGTAGTCTTTAAACTTTCTTCATATGTTTTTCAGTCTTTC  
 TGTGTAATTTCTTGATATTTTATCATGTTTTGTTCTACTGGGAGTAGCCTTTTTTCCA

Figure. 6

WO 03/014342

PC1/EP02/08963

6 /10

TTTCATTTTCGGCTGGTTTCATTGCTGGTTGTTTTTTGTTTTGTTTTGTTTTGAGAT  
GGAGTCTCACTCTGTCGCCAGGCTGGAGTGCAGTGCACAATCTCGGCTCACTGCRAAC  
TCTGCCCTCCAGGTTCAAGCGATTCTTTCTCAGCCTCCTGAGTAGCTGGGATTACAG  
GCATGTGCCACCATGCCAGCTAATTTTATATATTTTACTAGAGATGGGGTTCTCCAT  
GTTGGTCAGGCTGGTCTCAACTCCCAATCTCAGGTGATCCGCTGCCCTGCTCCAA  
AGTGTGGGATTATAGACATGAGCCACCTGCCCTGGCTAGTTCTTATGGGATGTATAG  
TCTTTGGATTTCATATGATATGATATATGTTTATATTTCTACAGTACATACCTAGGAGT  
GGAAATTGTTGGGTCATAGGTTAATGCATGTTTTCTGCCAAAACAGTTGTGTCRAATTTCTG  
TTTTCACCGCTGTGAATGAGAGTTGTTCTACCTTCTTGACAACACTTGATATGTCAGTC  
ATTTTAGCCATTCTGGTGAAATTTATAGTGTATTTCTGTGTGTATAGAGAGAGAAATGAG  
AGAGGGTGTGTTGTGAGAAAACCAAGCAACACTGTGAGAGTGTGTGTTGTTGAGAAAA  
CCAAAATACATACTACTGTGATTTTCATTCGGCAGAAAATCTGTTTGGTATATCAAAAAG  
TAGCTTAATTACTTCATCATATATTGGTTTAGGT

Figure. 6 .cont

7 / 10

SEQ ID NO:4

TAAGAACATTTTACACTCTTCAGTATAAAGAAGTCAGAATACCCCTACCCPATCAGTAAA  
GGCCTATAAGTTACCATTA AAAAGATGTCCTTAAAAACAGCATTCTCAGCTGGGCACGGT  
GGCTCACACCTTTGTCCAGTACTTTGGGAAGCCGAGTGGGTGGATCACCTGAGGTCAG  
GAGTTCGAGACCAGCCTGGCCAACATGGCGAAAAACCCATTTCTCTACTAAAAATACAAA  
AATTAGCCCGGCATGGTGGCGGGTGTCTGTGGTCCCAGCTACTCAAGAGGCTGAGGTGGG  
AGGATCACTGAGCCCAGGAGGTGGAGGCTGCATTGAGCCAGATTGTGCCACTGCACATCC  
AGCCTGGGTGACAGCCGAGACTCTGTCTCAAAAAAACCAAAACAAAAAACCCAGCAT  
TCTTTAGTAAATAATTCATAGTTTCTTCATCTAGAATTTAAATTTGATAGTTGATCA  
GCATGCTCTGAGCACGTGTGTGTGCTTACTAGTTTAGATCGGTAGATGTGTATATAAG  
TTATAGGTATAAAATCAATCCCTGAGTTCACACAAGGTTTTGATGTTGATACAAGTACAG  
TAAGTGTATATTTTAGTTTATGCTCTTAGTTTAAAGTCAATTTGTGTGGTCTTTCTAGCT  
TTAGGATCTGTTGAATATCTTCTCTTAGAAAAGGGAGTTAAGAACTTCACTTACCTATC  
TTCTACTTGTTTGGAGAATAGAAGATCCCTGTGGTAGCAGACTTTGTGAGTTTACTTGT  
AATTTTCCATCTGAAGACTCTTCTTCTTTTCTGATGAAGTCTTGCCTCTGTCGCCAG  
GCTGGAGTGCAGTGGTGCACCTTGGCTCAGTGCACCTCTGCCTCCCGGTTCAAGCAA  
TTCTCCTGCCCTCAGCCTCCCGAGTATCTGGGATTAAGAGTGCACACCCACACCTTGGCT  
AATTTTGTATTTTTCAGTAGAGACGGGGTTTCCACCATGTTGGCCAGGCTGGTCTCGAAT  
CTTGACCTCATGATCAGCCACCTCAGCCTTCCAAAGTGTGGGATTACAGGTGTGAGCC  
CCCACACTCGGCCGTTGTTGTTTTTTAAGAGACAGGGTCTCACTCTGTCACCTAACCTGG  
AGTACAGTGGCAATCATGGCTCACTGTAACCTCAAAATGCCCGGCTTAGTGAAGCTTCT  
TCCGCTTGGCTCCCAAAGTGTGGGATTAAGTGTGAGCCATGCATCCAGCTTGAA  
AGACAGCTTCTTAGGCTTGATTTGTTTGGTTACAGG

Figure. 7

WO 03/014342

PC1/EP02/08963

8 / 10

SEQ ID NO:5

ATCAGAGTCACTTTGATATGATGAGAGCAGAGATAAACAGATTTGTTGCATGTTTTAAT  
 CTTTGGTATGGGACATACTAGAATTCACATGCAAAATACATTTTTATGTAAGTGTGAAATGC  
 TCATACGACCATTGGAAATCTTCCCTTTAAAGAGCTTGGTAAGCAATTTGAGTGTAGTTGTT  
 AGACGGAGACCATCACGTCATAGTTTATAGACTGCATAAAGACGTAAGTTACCATTTAAT  
 ACCTTTCATTCAGGAAAAATGTACTTACACCTACAATGTACTAGTAGGCCCTCTCGCGCTG  
 GCAATACAGATAAGATAANTGTAGCCCTGGCCCTCAAAGGAACTCTCCCTCCTTAGGTTGCA  
 TTTGTAATAATGTTTGATTTTTAGATTGTTCTTTGAGCCCTGCATCCACGAGGATAGGTC  
 AGTGGTATTTAACGAGGTTAAAGGGGAGTAGTACCAAAGGGCATTCRAGCGTCCCATCTT  
 CGCTTCAACCAAAGCAGCCCTGCGTTTTCCTAGTTTTAATAATAGGTTTCAATGTAAGGTC  
 GTCCTTGAAGGGGGTTTTGGCTTTTTTTTACAGTGTGACTGAGGTATAATTTATAAAAA  
 GGGAAATGTATGGCATGGTGTAGTTTTTTCACATACATCTTTGTGAATACCCAGCTCAAGA  
 TCCAAAACATTTCCATAATTTCCAGAAAGTTCCAAACCCCTGCTCTTTTCAGTCTTAGCC  
 CTCTTCCCTGAAGTAAACCTGTTCCGACTTCAATCACTACTTTTATCCCAAGGTTAA  
 TTTTTTGGCTTTTTTCCACTAAATTTTCAATTTCTTTGATATGGTACTTTACTATTGACG  
 AAGTACTTTACACACTAGGTTTATTAATTTCTTTGATTCACCCAATTTTAGGGAACACC  
 TGTAGGGGACAAAATGAATGAGAGCCCTGCTTCCATGCTGCTAACTCTGGTGGGAA  
 CGAGACATGTATTTAATTAAGCATGTAAAAAATAGRGTTGGTGAAGAAATATCTATATA  
 CTAATCCCAATGACACACAGTTTACCTATGTARCAAACTGCATGTGTACCCCGAACC  
 TAAAATATAAGTTGGAATTAATAAAAAAACGAGAGGGGAAATAGAGCATCACAAACAGAG  
 TGCTGAGATGAATTAATTTATACCAAAGAGGAGGAGGACTCAGGGAGGTGCCGACTT  
 TAAACCCAGTCACTGAAGGGTGTGAGAAATTTGGATAGGCAAGTACCCCTGGGACAAAGT  
 CATTTCAAAACCAAGCTAACATTTCTACTTTTTTTTTCATTGTGATAGTTCCTGAAGTGA  
 GTTGCATAAAACTGGTACATGCTTAGGGCAGTCTCTAATTTGATTTTTATTTTTTCTAT  
 TTTTAAAAATAGTCTTCAAAATAGCAGATTCACATGATATTAATAATATGACACATAAAT  
 TATATACACAAATATATTTTTCTGAATGAAATTTAGTATCTGCATATATTTAAGAGCTATT  
 TCTGTCTCATATGTTTCATATCTTCATCCATTAATAAACTCTTTTGTAGGCCCTTCTCAC  
 TCTAAGATTATAAAAAATTTCCCATTTTACCTAGCTAGTTTTCTAGTTGTTCCAAAA  
 CCATTTTATTGAACAATCCATCTTTTTGACACTGGTTTTGGCATGCCCTTAATTAATATATCT  
 TGTGTGTGTTAGGATCTCCCTTTGGACTTTCCATTTCTGTTCAATGAGTCTTATCAGCTCC  
 TCTTACATTTGTTACCATGATGTTTTAACTTATGGGCTTTGTTAGTTTAAATGTAGGGCTA  
 GTTCCAGCCCATTTGTTCTTATCAGCTTTAGGAACCTTAGAAATCAGCTTGCTCTGTTTT  
 AAAGAAAACCTGGTATTTTTTATCAGTATAACATTTCTATTTATATTAACCTGAGAAAT  
 TGAAAACATCTATGATTTTTCCATATTCAGTAACGTATCACTTAGAATAGGTTAGGTTGTA  
 CTACTATAAAATCTCAGCTGCATAAAACAATTTTTTTTGGCTTGTGTACACATCCATTA  
 GGTCAATCAAGGGACTCACCTTGTCAAGTTACTCAGAGATTCAGGCTGATATAAAGGTTG  
 ATCTTGACATACGCTTTTCATGATGACAGAAAGCAGGGAAGAAAGGTGGTGAACCATGTG  
 CTCTTCCCTCTCTTATCCAGAAATGACACATACTCACATTTCAATTCGCCAGAGAAATTA  
 ACATGGCCCTTCCCTAAGTTCAAAATGGATAGAGAAATGCCCTTCCATCCAGTGCACAGAAAT  
 TAGAGAGCAAACTTTGTGAACAGTTCTGAGTACCACAAAATACCGTTATCTTTCCACTT  
 AAGTCTTCTGTTTCACTCAGTAGTCTTTAAACTTTCTTCATATGTTTTTCAAGTGTTC  
 TTGTTGAATTTCTGATATTTATCATGTTTGTTCGTACTGGGAGTAGCCTTTTTTTCCA

Figure. 8

WO 03/014342

PCT/EP02/08963

9/10

TTTCATTTTCTGGCTGGTTTCAFTGCTGGTTGTTTTTTTGTTTTGTTTTGTTTTTGAGAT  
GGAGTCTCACTCTGTTCGCCAGGCTGGAGTGCAGTGTCAACAATCTCGGCTCACTGCAACC  
TCTGCCTCCCAGGTTCAAGCGATTCTTCTTTCTCAGCCTCCTGAGTAGCTGGGATTACAG  
GCATFTGCCACCATGCCAGCTAATTTTTATATTTTTTAGTAGAGATGGGGTTTCTCCAT  
GTTGGTCAGGCTGGTCTCAAACGCCAATCTCAGGTGATCCGCCTGCCTCTGCCTTCCAA  
AGTGCCTGGGATTATAGACATGAGCCACCGTGCCTGGCCTAGTTCCTTATGGGATGATATG  
TCTTTGGATTATATGATATGATATATGTTTATATTTCTACAAGTACATACCTAGGAGT  
GGAATTGTTGGGTCAATAGGTTAATGCATGTTTTTCTGCCAAACAGTTGTGTCATTTCTG  
TTTTACCGCTGTGAATGAGAGTTTCTTACCCTTCTGACAAACACTTGATATGTCAGTC  
ATTTAGCCATTCGGTGAATTTATAGTCTATTTCTGTGTGTAAAGAGAGAGAAATGAG  
AGAGGCTCTTTCTGAGAAAACCAAGCAACACTGTGAGACTGTGTGTCTTTGTGAGAAA  
CCAAAATACATACTACTGTGATTTTCATGGGAGAAAATCTGTTTGGTATATCAAAAAAAG  
TAGCTTAATTACTTCATCATTATTGGTTTAGGT

Figure. 8 cont.

WO 03/014342

PCT/EP02/08963

10 / 10

SEQ ID NO:6

TAAGAACATTTTACACTCTTCAGTATAAAGAAGTCAGAATACCCCTACCCTATCAGTAAA  
GGCCTATAAGTTACCATTAAAAAGATGTCCTTAAAAACAGCATTCTCAGCTGGGCgCGGT  
GGCTCACACCTTTGTCCCAGTACTTTGGGAAGCCGAGGTGGGTGGATCACCTGAGGTCAG  
GAGTTCGAGACCAGCCTGGCCAAACATGGCGAAAACCCATTTCTCTACTAAAAATACAAA  
AATTAGCCGGGCAATGCTGGCGGGTGTGTGGTCCCAGCTACTCAAGAGGCTGAGGTGGG  
AGGATCACTGAGCCAGGAGGTGGAGGCTGCATTTGAGCCAAGATTGTGCCACTGCCTCC  
AGCCTGGGTGACAGAGCGAGACTCTGTCTCAAAAACCAAAAACAAAAAACCCAGCAT  
TCTTTAGTAAATAATTCCATAGTTTCTCTCATCTAGAATTTAAAAATGTGATAGTTGATCA  
GCATGCTCTGAGCACCTGTGTTCCTGTACTAGTTAGATCGGTAGATGTGTATATAG  
TTATAGGTATAAAAATCAATCCGAGTTGACACAAAGTTTGTGATGTTGAGTACAAGTACAG  
TAAGTGTATATTTTATGTTATGCTCTTAGTTTTAAGTCAATGTGTGTTCTTCTAGCT  
TTAGCATCTGTGAATATCTTCCCTAGAAAAGGGAGTTAAGAACTTCACTTACCTATC  
TTCTACTGTGTTTGGAGAAATAGAAGATCCCTGTGGTACGAGACTTTGTGAGTTACTTGT  
AATTTTCCATCTGAAAGACTGTCTTGTTTTTCGTGATGAAGTCTTGCTCTGTGCCCCAG  
GCTGAGTGCAGTGGTGCACACCTTGGCTCACTGCAACCTCTGCCCTCCCGGTTCAAGCAA  
TTCTCCTGCCCTCAGCCTCCCGAGTATCTGGGATTACAGGTGCACACCACCACCTGGCT  
AATTTTGTATTTTCAGTAGACACGGGTTTACCATGTTGGCCAGGCTGGTCTCGAACT  
CTTGACCTCATGATCAGCCACCTCAGCCTTCCAAAGTGTGGGATTACAGGTGTGAGCC  
CCCACACTCGGCCGTGTGTGTTTTTAAAGAGACAGGGTCTCACTCTGTACCTAACCTGG  
AGTACAGTGGCAATCATGGCTCAGTGAACCTCAAATGCCCGCCTTAGTGAAGGCTTCT  
TCTGCCCTTGGCTCCCAAAGTGTGGGATTACAAGTGTGAGCCATGCATCCAGCTTGAA  
AGACAGCTTCTTAGGCTTGATTTGTTTGTACAGG

Figure. 9

WO 03/014342

PCT/EP02/08963

1

## SEQUENCE LISTING

<110> National Public Health Institute  
PELTONEN, Leena  
ENATTAH, Nabil  
JÄRVELÄ, Irma  
SAHI, Timo  
SAVILAHTI, Erkki  
TERWILLIGER, Joseph

<120> Identification of a DNA variant associated with adult type hypolactasia

<130> F 2034 PCT

<150> EP 01 11 9377.8  
<151> 2001-08-10

<150> EP 01 11 9528.6  
<151> 2001-08-14

<150> US 60/315,955  
<151> 2001-08-31

<160> 14

<170> PatentIn version 3.1

<210> 1  
<211> 180  
<212> DNA  
<213> Homo sapiens

<400> 1  
accttcaat caggaaaaat gtacttagac cctacaatgt actagtaggc ctctgogctg 60  
gcaatacaga taagataatg tagccctgg cctcaaagga actctctcc ttaggttgca 120  
tttgtataat gtttgatttt tagattgttc tttagccct gcattccacg aggataggtc 180

<210> 2  
<211> 180  
<212> DNA  
<213> Homo sapiens

<400> 2  
taagaacatt ttacactctt cagtataaag aagtcagaat accctaccc tatcagtaaa 60  
ggcctataag ttaccattaa aaagatgtcc ttaaaaacag cattctcage tgggcgcegt 120  
ggctcacacc ttgtcccag tactttgga agccgaggtg ggtggatcac ctgaggtcag 180

<210> 3  
<211> 3213  
<212> DNA  
<213> Homo sapiens

<400> 3

WO 03/014342

PCT/EP02/08963

2

```

atcagagtc ctttgatag atgagagcag agataaacag atttggtgca tgtttttaa 60
ctttggtag ggacatacta gaattcactg caaatacatt tttatgtaac tgttgaatgc 120
tcatacagcc atggaattct tccottttaa gagottggta agcaittgag tgtagttggt 180
agacggagac gatcacgtca tagtttatag agtgcataaa gaogtaagtt accatttaat 240
acotttcatt caggaaaaat gtaactagac cctacaatgt actagtaggc ctctgcgctg 300
gcaatacaga taagataatg tagtcoctgg cctcaaaagg actctcctcc ttaggttgoa 360
tttgataaat gtttgatttt tagattgttc tttgagcctt gcattccacg aggataggtc 420
agtgggtatt aacgaggtaa aaggggagta gtaogaaagg gcattcaagc gtccatctt 480
cgottcaacc aaagcagccc tgcgttttcc tagttttatt aataggtttg atgaaggtc 540
gtctttgaaa agggggttg gctttttttt acagtgtgac tgaggataaa tttataaaaa 600
gggaaatgta tggcatggty agttttttca catacatcct tgtgaatacc cagctcaaga 660
tccaaaaaat ttocataatt tcagaaagtt ccaaacccct gctcttttcc agtcttagcc 720
ctcttcccct gaagtaacca ctgttccgac ttcaatcact actttatcc cacaggttaa 780
ttttttggot tttttocact aaattttcaa attctttgat atggtacttt actattgacg 840
aagtaacttcc acaactaggtt atttaaatatt ctttgattca ccaaatattt agggaacacc 900
ttagggggac aaaaaatgaa tgagagcccc tgccttccat tgcgtgtaat ctgggtggaa 960
cgagacatgt atttaattaa gcatgtaaaa aatagagtgg gtagtgaat aatctatata 1020
ctaaatcccc atgacacaca gtttacctat gtaacaaacc tgcattgta cccccaacc 1080
taaaatataa gttggaattt aaaaaaaaaa gagagggaga atagagcacc acaaccagag 1140
tgcctgagatg aattacttta ttaccaaaga aggaggagga ctacgggag tgcgacggt 1200
taaacccagt cactgaaggy tgtgcagaat ttgataggc aagataccct gggcaaggt 1260
cattotaaaa ccatgctaac atttgtaactt tttttttcat tgtgatagt cctgaaatga 1320
gttgcataaa actggtacat gtcttagggc agtctctaat tgatttttat ittgttctat 1380
ttttaaaaat tagtctcaa atagcagatt cacatgatat taaaatatat gcacataaat 1440
tatatacaca aatatatttt ctgaatgaaa tttaglatct gcataatatt aagagctatt 1500
ctgtctccat atgttcataa tcttccatcca ttaaaaaaac tttgttagg cctttctcac 1560
tctaagatta taaaaaatc tccattattt tacctagcta gttttctagt tgttccaaaa 1620
ccatttatg aacaatccat ctttttgaca ctggtttgge atgcottaat tatatattot 1680
tgtgtgtgtt aggatctcct tttggacttt ccattctggt cattgagctc taccagctcc 1740
tcttccattg gtaccatgat gttttaatct atggggcttt gtagtttaaa tbtagggcta 1800

```

WO 03/014342

3

PCT/EP02/08963

```

gttcacagc attgtctct atcagctgtt aggaacttag aatcagctt getctgttt 1860
aaagaaaaac ctggtattht ttatocagta taacattota ttatattaa ctgagaat 1920
tgaaaaatc tatgatttt cotattcagt aacgtatcac ttagaatagg ttaggttga 1980
ctactataaa atctcagctg cataaaacaa ttttttttg cttgtgctec acatocatta 2040
ggtcacaag ggaactcact tgtcaagta ctcagagatt oaggctgata taaaggttg 2100
atcttgact aagctttoat gatgacagaa agcagggaa agaaggtgg gagccatgtg 2160
ctttctccc ctctatcca gaaatgacac atactacat ttcattgcc agagaaatta 2220
aatggccc tectaagttc aatggatag agaattgct tcotaccagg tgocagaat 2280
tagaagagca aacatttgg aacagttctg agtaccacaa ataccgttat cttccactt 2340
aagctctctg ttctactag tagtcttta aacttttct catatgttt tcagtgttc 2400
ttgttgaatt tctgataat ttatcatggt tgtctgact gggagtgc ttttttcca 2460
tttcttttc tggctgggtt cttgtctgt tgttttttg tttgttttg ttttgagat 2520
ggagtctcac tctgtgcc aggctggagt gcagtgtcac aatctggct cactgcaac 2580
ctgtctccc aggtcaagc gattcttctt tctcagctc ctgagtact gggattacag 2640
gcattgcca caatgcccag ctaattttt atatttttag tagagatggg gttctocat 2700
gttggtcagg ctggtctcaa actcccaatc tcaagtgatc cgcctgctc tgccttcaa 2760
agtctggga ttatagacat gagccaccgt gctggccta gttctatgg gatgtatg 2820
cttttgatt catatgat gtatatalgt ttatattct acaagtacat acctaggat 2880
ggaattgtg ggtcataggt taatgatgt tttctgcca aacagttgt tcaattctg 2940
tttccccc tgtgaatgag agttgttcta cttcttgac aacactgat atgtcagtc 3000
attttagcca ttctgggtaa ttatagtc tttctgtg tgttaagag agagaatgag 3060
agaggtgtt tgtgagaaa ccaagcaac actgtgagag tgtgtgtt tgtgagaaa 3120
ccaaaataca tactactgt atttcattgg gagaaatct gttgtgata tcaaaaaag 3180
tagcttaatt acttaoat tattggttta ggt 3213

```

```

<210> 4
<211> 1296
<212> DNA
<213> Homo sapiens

```

```

<400> 4
taagancatt ttacactctt cagtataaag aagtcagaat accctaccc taccagtaaa 60
ggocataaag ttaccattaa aaagatgtcc ttaaaaaag cttctcagc tgggcacggt 120

```

WO 03/014342

PCT/EP02/08963

4

```

ggctcacacc ttgtccag taacttggga agccgaggtg ggtggatcac ctgaggtcag 180
gagttcgaga ccagcctggc caacatggcg aaaaccatt ttctctacta aaaatacaaa 240
aattagccgg gcatggggc gggtgottgt ggtccagct actcaagagg ctgaggtggg 300
aggatcaacty agcccaggag gtggaggctg cattgagcca agattgtgco actgcactcc 360
agcctgggtg acagagcgag actctgtctc aaaaaacca aaacaaaaa aaccagcat 420
tctttagtaa ataattoata gtttttctca totagaatth aaaattgtga tagttgatca 480
gcatgtctcg agcaogtgtg ttgtctgta ctagttaga tggtagatg tgtatataag 540
ttataggat aaatcaatc ctgagttgac acaagtttt gatgttgagt acaagtacag 600
taagtgtata tttttagta tgccttagt ttaagtcaa ttgtgtggt cttctagct 660
ttaggatctg ttgaattatc ttccttagaa aagggagta agaatttca cttacctatc 720
ttctactgt ttggagaata gaagagtcoc tgtgtagca gactttgtga gtttactgt 780
aattttccat ctgaaagact gttctgttt ttctgtatga agtcttgctc tgtcggccag 840
gctggagtc agtggtgcaa ccttggctca ctgcaacctc tgcctcccg gttcaagcaa 900
ttctctgccc tcagcctccc gagtatctgg gattacaggt gcacaccacc acacctggct 960
aatttttcta ttttcagtag agacggggtt tcaccatgtt ggccaggctg gtcctgaact 1020
cttgacctca tgatcagccc acctcagcct tccaaagtgc tgggattaca ggtgtgagcc 1080
cccacactcg gcogttgtt tttttaaga gacagggctc cactctgtca cctaacctgg 1140
agtacagtag caateatggc tnaactgtaac ctcaaatgccc cggccttagt gaagcgttct 1200
tootgcttg gctcccaaaa gtgctgggat tacaagtgtg agccatgcat ccagcttgaa 1260
agacagcttc ttaggcttga ttgtttggt tacagg 1296

```

```

<210> 5
<211> 3213
<212> DNA
<213> Homo sapiens

```

```

<400> 5
atcagagtca ctttgatag atgagagcag agataaacag atttgttga tgtttttaat 60
ctttggtatg ggacatacta gaattoactg caatacatt tttatgtaac tgttgaatgc 120
tcatacgacc atggaattct tcccttataa gagcttggtg agcatttgag tgtagttggt 180
agacggagac gatcacgtoa tagtttatag aytgcataaa gacgtaagtt accatttaat 240
acotttcatt caggaaaaat gtacttagac cctacaatgt actagtaggc ctctgctgctg 300
gcnatacaga taagataatg tagcccctgg cctcaaagga actctctccc ttaggttgca 360

```

WO 03/014342

PCT/EP02/08963

5

```

ttgtataat gtttgattt tagattgtt tttgagcoo gcattccaag aggataggto 420
aetgggtatt aacgaggtaa aaggggagta gtaoaaagg gcattcaagc gtcccatctt 480
cgottcaacc aaagcagccc tgcgttttcc tagttttatt aataggtttg atgtaaggtc 540
gtctttgaaa agggggtttg gotttttttt acagtgtgac tgaggataaa tttataaaa 600
gggaaatgta tggcatggtg agttttttta cacaatcctc tgtgaatacc cagctcaaga 660
tccaaaacat ttccataatt tcagaaagt ccaaacccct gctcttttcc agtcttagcc 720
ctcttcccc gaagtaacca ctgttcogac ttcaatcact acttttatcc cacaggtaa 780
ttttttggct tttttccact aaattttcaa attctttgat atggtacttt actattgacg 840
aagtacttto acaactaggtt atttaaatatt ctttgattca cccaatattt agggaacacc 900
tgtaggggag aaaaaatgaa tgagagcccc tgccttccat tgotgcta atctgtggaa 960
cgagacatgt atttaattaa gcatgtaaaa aatagatggg gtgatgaaat aatctatata 1020
ctaaatcccc atgacacaca gtttacctat gtaacaaacc tgcattgtga cccccaagc 1080
taaaatataa gttggaaatt aaaaaaaaa gagagggaga atagagcaco acaaccagag 1140
tgotgagatg aattacttta ttaccaaga agggaggaga ctgaggaggg tgcgaagtt 1200
taaacocagt caotgaaggg tgtgcagaat ttggatagcc aagataccct gggcaaggt 1260
cattctaaaa ccattgtaac atttctactt ttttttccat tgtgatagtt cctgaaatga 1320
gttgcataaa actggtacat gctttaggcc agtctctaat tgattttat tttgtctat 1380
ttttaaaaat tagtttcaa atagcagatt cacatgata taaatataat gcacataaat 1440
tatatacaca aatatatttt ctgaatgaaa tttagtatct gcatatattt aagagctatt 1500
tctgtctcat atgttcataa tcttcatcca ttaaaaaaac tttgttagg cctttctac 1560
tctaaagatta taaaaaatc tccattattt taccatageta gtttctagt tgttcaaaa 1620
ccatttattg aacaatccat ctttttgaca ctggtttggc atgccttaat tatatattct 1680
tgtgtgtgtt aggatctctt tttggacttt ccattctgtt cattgagctc tctcagctcc 1740
tcttaccattg taccatgat gttttaatct atggggcttt gtagtttaa ttaggggcta 1800
gttccagcgc attgttctct atcagctggt aggaacttag aaatcagctt gctctgtttt 1860
aaagaaaaac ctggtatttt tttatcagta taacattcta tttatattaa cttgaagaat 1920
tgaaaacatc tatgattttt cctattoagt aacgtatcac ttagaatagg ttaggttgta 1980
ctactataaa atctcagctg caataaaaca tttttttttg cttgtgctac caatccatta 2040
ggtoalcaag ggactcaact tgtcaagtta ctcaagatt caggctgata taaaggtttg 2100

```

WO 03/014342

6

PCT/EP02/08963

```

atcttgacat acgctttcat gatgacagaa agcaggggaag agaaggtggt gagccatgtg 2160
ctttctcccc cttctatcca gaaatgacac atactcacat ttcattogcc agagaaatta 2220
acatggcccc tcttaagttc aaatggatag agaaatgccc tcttaccagg tgcccagaat 2280
tagaagagca aacatttgtg aacagttctg agtaccacaa ataccgttat ctttccactt 2340
aagtctctctg ttccactcag tagtgcctta aacttttctt catatgtttt tccagtgtttc 2400
ttgttgaaatc tottgatatt ttatcatggt tgttcgtaet gggagtagcc tttttttcca 2460
tttcattttc tggctggttt cattgctggt tgttttttg ttttgtttg tttttgagat 2520
ggagctcac tctctgccc aggtctggagt gcagtgtcac aatctcgctt cactgcaacc 2580
tctgctcccc aggtccaagc gattcttctt tctcagcttc ctgagtagct gggattacag 2640
gcagtgtcca ccattcccag ctaatttttt atatttttag tagagatggg gttctccat 2700
gttggtcagg ctggtctcaa actcccaatc tccagtgatc cgcctgcctc tgccttccaa 2760
agtgtgggga ttatagacat gagccaccgt gcctggcota gttcttatgg gatgtatatg 2820
tctttggatt catatgatat gtatatatgt ttatatttct acaagtacat acctaggagt 2880
ggaattgttg ggtcataggt taatgcatgt tttctgcca aacagttgtg tcaatttctg 2940
ttttcacogc tgtgaatgag agttgttota ccttcttgac aacacttgat atgtgcagtc 3000
atthtagcca ttctgggtgaa tttatagtc ttttctgtg tgtgtaagag agagaatgag 3060
agaggggtgt tgtgagaaaa ccaaagcaac actgtgagag tgtgtgtgtt tgtgagaaaa 3120
ccaaaataca tactactgtg atttctgtg gagaaaatct gtttggata tcaaaaaaag 3180
tagcttaatt acttcatcat tattggttta ggt 3213

```

```

<210> 6
<211> 1296
<212> DNA
<213> Homo sapiens

```

```

<400> 6
taagaacatt ttacactctt cagtataaag aagtoagaat accoctaacc tatcagtaaa 60
ggcctataag ttaccattaa aaagatgtcc ttaaaaacag cattctcagc tggggcgggt 120
ggctcacacc tttgtcccag tactttggga agccgagggt ggtggatcac ctgaggtcag 180
gagttctgaga ccagcctggc caacatggcg aaaaaccatt ttctctacta aaaatacaaaa 240
aattagcccg gcattggtgc ggggtcctgt ggtcccagct actcaagagg ctgaggtggg 300
aggatcactg agcccaggag gtggaggctg cattgagcca agattgtgcc actgcactcc 360
agcctgggtg acagagcgag actctgtctc aaaaaaacca aaaaaaaa aaccagcat 420

```

WO 03/014342

7

PCT/EP02/08963

```

tctttagtaa ataattcata gttttcttca totagaattt aaaattgtga tagttgatca 480
gcattgtcctg agcaactgtg tttgtctgta ctagttaga tcggttagatg tgtatataag 540
ttataggtat aaaatcaatc ctgagttgac acaaggtttt gatgttgagt acaagtacag 600
taagtgtata tttttagtta tgccttagt ttttaagtcaa ttgtgtggtt cttctagct 660
ttaggatctg ttgaattatc ttccttagaa aaggaggtta agaattctca cttactatc 720
ttctactgtt ttgagaata gaagagtccc tgtggtagca gactttgta gttacttgt 780
aattttccat ctgaaagact gttcttgttt ttcgtgatga agtcttgctc tgtcgcacag 840
gctggagtgc agtgggtgcaa ccttggctca ctgcaacctc tgcctccggg gttcaagcaa 900
ttctcctgcc tcagcctccc gattatctgg gattacaggt gcacaccacc acactcggct 960
aatttttcta ttttcagtag agacggggtt tcacctggtt ggcaggtg gtctogaact 1020
cttgacctca tgatcagccc acctcagcct tccaaagtgc tgggattaca ggtgtgagcc 1080
cccactctg gcctgtgtg ttttttaaga gacagggtct cactctgta cctaacctgg 1140
agtacagtgg caatcattgc tcactgtaac ctcaaatgcc cggccttagt gaaggttct 1200
tcctgccttg gcctcccaaa gtgctgggat tacaagtgtg agccatgcat ccagctttaa 1260
agacagcttc ttaggcttga tttgtttggt tacagg 1296

```

```

<210> 7
<211> 24
<212> DNA
<213> Artificial sequence

```

```

<220>
<223> Primer

```

```

<400> 7
taggtcagtg ggtattaacg aggt 24

```

```

<210> 8
<211> 23
<212> DNA
<213> artificial sequence

```

```

<220>
<223> Primer

```

```

<400> 8
gtcactttga tatgatgaga gca 23

```

```

<210> 9
<211> 23
<212> DNA
<213> artificial sequence

```

WO 03/014342

8

PCT/EP02/08963

<220>  
 <223> Primer  
 <400> 9  
 cctcgttaat acccaactgac cta 23

<210> 10  
 <211> 24  
 <212> DNA  
 <213> artificial sequence

<220>  
 <223> Primer  
 <400> 10  
 ggcaatacag ataagataat gtag 24

<210> 11  
 <211> 21  
 <212> DNA  
 <213> artificial sequence

<220>  
 <223> Primer  
 <400> 11  
 ttgatcagca tgcctgagc a 21

<210> 12  
 <211> 22  
 <212> DNA  
 <213> artificial sequence

<220>  
 <223> Primer  
 <400> 12  
 ctaccctatc agtaaaggcc ta 22

<210> 13  
 <211> 21  
 <212> DNA  
 <213> artificial sequence

<220>  
 <223> Primer  
 <400> 13  
 tgotcaggac atggtgatca a 21

<210> 14  
 <211> 23  
 <212> DNA

WO 03/014342

PCT/EP02/08963

9

<213> artificial sequence

<220>

<223> Primer

<400> 14

aaaaacagca ttctcagctg ggc

23

## 【国際公開パンフレット(コレクトバージョン)】

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau(43) International Publication Date  
20 February 2003 (20.02.2003)

PCT

(10) International Publication Number  
WO 03/014342 A3

- (51) International Patent Classification: C12N 15/00 [US/US]: Columbia University, Department of Psychiatry and, Columbia Genome Center, 1150 St. Nicholas Avenue, New York, NY 10032 (US).
- (21) International Application Number: PCT/JP02/08963
- (22) International Filing Date: 9 August 2002 (09.08.2002) (74) Agent: VOSSIUS & PARTNER; Siebertstrasse 4, 81675 Munich (DE).
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
01119577.8 10 August 2001 (10.08.2001) JP  
01119528.6 14 August 2001 (14.08.2001) JP  
60/315.955 31 August 2001 (31.08.2001) US
- (71) Applicant (for all designated States except US): NATIONAL PUBLIC HEALTH INSTITUTE [FI/FI]; Mannerheimintie 166, FIN-00500 Helsinki (FI).
- (72) Inventors; and  
(75) Inventors/Applicants (for US only): PELTONEN, Leena [FI/FI]; National Public Health Institute, Biomedicum Dept of Molecular Medicine, Haartmaninkatu 8, 00251 Helsinki (FI). ENATTAH, Nabil [LY/FE]; National Public Health Institute, Biomedicum, Dept of Molecular Medicine, Haartmaninkatu 8, 00251 Helsinki (FI). JÄRVELÄ, Irma [FI/FI]; National Public Health Institute, Biomedicum Dept of Molecular Medicine, Haartmaninkatu 8, 00251 Helsinki (FI). SAHL, Timo [FI/SE]; Finnish Defence Forces, P.O.Box 919, 00131 Helsinki (FI). SAVILAHTI, Erkki [FI/FI]; Hospital for Children and Adolescents, P.O.Box 281, 00029 HUS (FI). TERWILLIGER, Joseph
- (81) Designated States (national): 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, GR, GM, HR, HU, 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, PI, 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.
- (84) Designated States (regional): ARIPO patent (GI, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:  
with international search report
- (88) Date of publication of the international search report:  
18 December 2003
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: IDENTIFICATION OF A DNA VARIANT ASSOCIATED WITH ADULT TYPE HYPOLACTASIA

(57) Abstract: The present invention relates to a nucleic acid molecule comprising a 5' portion of an intestinal lactase-phlorizin hydrolase (LPH) gene contributing to or indicative of the adult-type hypolactasia wherein said nucleic acid molecule is selected from the group consisting of (a) a nucleic acid molecule having or comprising the nucleic acid sequence of SEQ ID NO: 1, the sequence of SEQ ID NO:1 is also depicted in the Fig. 4 and comprised in the sequence as depicted in the Fig. 8, (b) a nucleic acid molecule having or comprising the nucleic acid sequence of SEQ ID NO: 2, the sequence of SEQ ID NO:2 is also depicted in Fig.5 and comprised in the sequence as depicted in the Fig. 9, (c) a nucleic acid molecule of at least 20 nucleotides the complementary strand of which hybridizes under stringent conditions to the nucleic acid molecule of (a) or (b), wherein said polynucleotide/nucleic acid molecule has at a position corresponding to position -13910 5' from the LPH gene a cytosine residue; and (d) a nucleic acid molecule of at least 20 nucleotides the complementary strand of which hybridizes under stringent conditions to the nucleic acid molecule of (a) or (b), wherein said polynucleotide/nucleic acid molecule has at a position corresponding to position -22018 5' from the LPH gene a guanine residue. The present invention further relates to methods for testing for the presence of or predisposition to adult-type hypolactasia that are based on the analysis of an SNP contained in the above recited nucleic acid molecule. Additionally, the present invention relates to diagnostic composition and kit useful in the detection of the presence of or predisposition to adult-type hypolactasia.

WO 03/014342 A3

## 【国際調査報告】

INTERNATIONAL SEARCH REPORT		International Application No. PCT/EP 02/08963
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N15/00		
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) IPC 7 C12N		
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 practical, search terms used) BIOSIS, SEQUENCE SEARCH, CHEM ABS Data, EMBASE, EPO-Internal, WPI Data		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL 'Online! 18 October 1999 (1999-10-18) SULSTON J.E. ET AL.: "Homo sapiens BAC clone RP11-34L23 from 2, complete sequence." Database accession no. AC011893 XP002244247 100% identity with SEQ ID NOs1 and 2;99,9% identity with SEQ ID NOs 3 and 4. the whole document --- --/--	1
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input type="checkbox"/> Patent family members are listed in 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 document but published on or after the international filing date *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another claim or other special reason (see specification) *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 *I* 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. *Z* document member of the same patent family		
Date of the actual completion of the international search 13 June 2003		Date of mailing of the international search report 02/07/2003
Name and mailing address of the ISA European Patent Office, P.O. Box 5518 Patentlaan 2 NL - 5250 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl Fac. (+31-70) 340-3016		Authorized officer: Dumont, E

INTERNATIONAL SEARCH REPORT		International Application No PCT/EP 02/08963
C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document with indication where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL 'Online! 5 May 1999 (1999-05-05) MAHAIRAS G.G. ET AL.: "HS_5237_A1_G08_SP6E RPCI-11 Human Male BAC Library Homo sapiens genomic clone" Database accession no. A0515834 XP002244248 80% identity with SEQ ID NO: 1 in 116 nt overlap (nt 24-139) the whole document ---	1
X	DATABASE EMBL 'Online! 26 April 2001 (2001-04-26) HARRINGTON J.J. ET AL.: "RST8055 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence." Database accession no. B6189020 XP002244249 86% identity with SEQ ID NO: 2 in 93 nt overlap (89-180) the whole document ---	1
X	DATABASE EMBL 'Online! 11 November 1999 (1999-11-11) MAHAIRAS G.G. ET AL.: "HS_3081_B2_E03_T7C CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone" Database accession no. A0892176 XP002244250 75% identity with SEQ ID NO: 3 in 56 nt overlap (302-357) the whole document ---	2
X	DATABASE EMBL 'Online! 3 August 1999 (1999-08-03) MAHAIRAS G.G. ET AL.: "HS_3106_A2_D07_T7C CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone" Database accession no. A0781670 XP002244251 81% identity with SEQ ID NO: 4 in 69 nt overlap (83-150) the whole document ---	2
Y	JARVELA IRMA ET AL: "Assignment of the locus for congenital lactase deficiency to 2q21, in the vicinity of but separate from the lactase-phlorizin hydrolase gene." AMERICAN JOURNAL OF HUMAN GENETICS, vol. 63, no. 4, October 1998 (1998-10), pages 1078-1085, XP002244245 ISSN: 0002-9297 the whole document ---	1-40
	-/--	

INTERNATIONAL SEARCH REPORT		Internal Application No PCT/EP 02/08963
C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	BOLL W ET AL: "STRUCTURE OF THE CHROMOSOMAL GENE AND COMPLEMENTARY DNAS CODING FOR LACTASE PHLORIZIN HYDROLASE IN HUMANS WITH ADULT-TYPE HYPOLACTASIA OR PERSISTENCE OF LACTASE" AMERICAN JOURNAL OF HUMAN GENETICS, vol. 48, no. 5, 1991, pages 889-902, XP009012268 ISSN: 0002-9297 abstract Discussion	1-40
Y	WANG YANGXI ET AL: "The lactase persistence/non-persistence polymorphisms is controlled by a cis-acting element." HUMAN MOLECULAR GENETICS, vol. 4, no. 4, 1995, pages 657-662, XP009012156 ISSN: 0964-6906 the whole document	1-40
P, X	ENATTAH NABIL SABRI ET AL: "Identification of a variant associated with adult-type hypolactasia." NATURE GENETICS, vol. 30, no. 2, February 2002 (2002-02), pages 233-237, XP002244246 February, 2002 ISSN: 1061-4036 the whole document	1-40

## フロントページの続き

(51) Int.Cl. <sup>7</sup>	F I	テーマコード(参考)
C 0 7 K 16/18	C 0 7 K 16/18	4 H 0 4 5
C 1 2 N 1/15	C 1 2 N 1/15	
C 1 2 N 1/19	C 1 2 N 1/19	
C 1 2 N 1/21	C 1 2 N 1/21	
C 1 2 N 5/10	C 1 2 N 7/00	
C 1 2 N 7/00	C 1 2 Q 1/34	
C 1 2 Q 1/34	C 1 2 Q 1/68	A
C 1 2 Q 1/68	C 1 2 Q 1/70	
C 1 2 Q 1/70	G 0 1 N 33/53	M
G 0 1 N 33/53	G 0 1 N 33/566	
G 0 1 N 33/566	C 1 2 N 5/00	A

(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, N O, 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 エナッター ネイビル

フィンランド 0 0 2 5 1 ヘルシンキ ハールトマニンカテュ 8 ナショナル パブリック  
ヘルス インスティテュート バイオメディカム デパートメント オブ モレキュラー メディ  
シン

(72) 発明者 イェルヴェレ イルマ

フィンランド 0 0 2 5 1 ヘルシンキ ハールトマニンカテュ 8 ナショナル パブリック  
ヘルス インスティテュート バイオメディカム デパートメント オブ モレキュラー メディ  
シン

(72) 発明者 サーヒ ティモ

フィンランド 0 0 1 3 1 ヘルシンキ フィーニッシュ ディフェンス フォーシス ペーオー  
ボックス 9 1 9

(72) 発明者 サヴィラーティ エルッキ

フィンランド 0 0 0 2 9 フス ホスピタル フォア チルドレン アンド アドレッセンツ  
ペーオーボックス 2 8 1

(72) 発明者 ターウィリガー ジョセフ

アメリカ合衆国 ニューヨーク州 1 0 0 3 2 ニューヨーク セント ニコラス アベニュー  
1 1 5 0 コロンビア ユニヴァーシティー デパー トメント オブ サイカトリー アンド  
コロンビア ゲノム センター

F ターム(参考) 4B024 AA01 AA11 BA12 CA01 CA09 CA20 DA01 DA02 DA05 DA11  
EA01 EA02 EA03 EA04 GA11 HA11 HA12 HA15 HA17 HA19  
4B063 QA01 QA13 QA17 QA19 QQ02 QQ35 QQ44 QR55 QR59 QR62  
QS05 QS25 QS32 QS33 QS34  
4B065 AA91X AB01 AC14 BA01 CA27 CA44 CA46  
4C084 AA13 NA14 ZB212 ZC212  
4C087 AA02 BC83 NA14 ZB21 ZC21  
4H045 AA11 AA30 DA75 DA89 EA50 FA74

专利名称(译)	鉴定与成人乳糖酶缺乏相关的DNA变体		
公开(公告)号	<a href="#">JP2005502336A</a>	公开(公告)日	2005-01-27
申请号	JP2003519472	申请日	2002-08-09
申请(专利权)人(译)	国家公共卫生研究所		
[标]发明人	ペルトーネンレーナ エナッターネイビル イエルヴェレイルマ サーヒティモ サヴィラーティエルツキ ターウイリガージョセフ		
发明人	ペルトーネン レーナ エナッター ネイビル イエルヴェレ イルマ サーヒ ティモ サヴィラーティ エルツキ ターウイリガー ジョセフ		
IPC分类号	A61K48/00 A61K35/76 A61K38/00 A61P3/00 A61P43/00 C07H21/04 C07K16/18 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N7/00 C12N9/24 C12N9/38 C12N15/00 C12N15/09 C12Q1/34 C12Q1/68 C12Q1/70 G01N33/53 G01N33/566		
CPC分类号	A01K2217/05 A61K38/00 A61K48/00 A61P1/14 A61P3/00 C12N9/2402 C12N9/2468 C12N9/2471 C12Q1/6883 C12Q2600/156 C12Y302/01023 C12Y302/01062 C12Y302/01108		
FI分类号	C12N15/00.ZNA.A A61K35/76 A61K48/00 A61P3/00 A61P43/00.105 C07K16/18 C12N1/15 C12N1/19 C12N1/21 C12N7/00 C12Q1/34 C12Q1/68.A C12Q1/70 G01N33/53.M G01N33/566 C12N5/00.A		
F-TERM分类号	4B024/AA01 4B024/AA11 4B024/BA12 4B024/CA01 4B024/CA09 4B024/CA20 4B024/DA01 4B024 /DA02 4B024/DA05 4B024/DA11 4B024/EA01 4B024/EA02 4B024/EA03 4B024/EA04 4B024/GA11 4B024/HA11 4B024/HA12 4B024/HA15 4B024/HA17 4B024/HA19 4B063/QA01 4B063/QA13 4B063 /QA17 4B063/QA19 4B063/QQ02 4B063/QQ35 4B063/QQ44 4B063/QR55 4B063/QR59 4B063/QR62 4B063/QS05 4B063/QS25 4B063/QS32 4B063/QS33 4B063/QS34 4B065/AA91X 4B065/AB01 4B065 /AC14 4B065/BA01 4B065/CA27 4B065/CA44 4B065/CA46 4C084/AA13 4C084/NA14 4C084/ZB212 4C084/ZC212 4C087/AA02 4C087/BC83 4C087/NA14 4C087/ZB21 4C087/ZC21 4H045/AA11 4H045 /AA30 4H045/DA75 4H045/DA89 4H045/EA50 4H045/FA74		
优先权	2001119377 2001-08-10 EP 2001119528 2001-08-14 EP 60/315955 2001-08-31 US		
其他公开文献	JP4521184B2		
外部链接	<a href="#">Espacenet</a>		

#### 摘要(译)

本发明有助于成人乳糖酶缺乏或成人乳糖酶缺乏，一个核酸分子，其包含肠乳糖酶-根皮苷水解酶的5'部分(LPH)基因中，(a)序列一个或具有ID NO的核酸序列的核酸分子：1，SEQ ID NO：1的序列：1，还如图4所示，也包含在图中所示的序列中的核酸分子8；(b)序列。一个或具有ID NO的核酸序列的核酸分子：2，SEQ ID NO：1的序列：2，也如图5所示，所述核酸分子也被包括在图9所示的序列中；(c)部分互补链杂交从LPH基因的(一个)或(b)在严格条件下核酸分子，至少20个核苷酸的核酸分子，其中所述多核苷酸/核酸分子，-13910 5' (D) 在与其互补链对应的位置具有胞嘧啶残基的核酸分子在严格条件杂交与(a)或(b)的核酸分子，至少20个核苷酸的核酸分子，其中所述多核苷酸/核酸分子从LPH基因对应于-22018 5'核苷酸分子，其位置具有鸟嘌呤残基；此外，本发明是基于包含在上面描述的核酸分子的SNP的分析，其涉及用于检测在存在或倾向成人型乳糖酶缺乏的方法。

此外，本发明涉及用于检测成人乳糖酶缺乏的存在或倾向的诊断组合物和试剂盒。

(5) Int. Cl. 7		F I		テ-マコード (参考)	
C 1 2 N	15/09	C 1 2 N	15/00	Z N A A	4 B 0 2 4
A 6 1 K	35/76	A 6 1 K	35/76		4 B 0 6 3
A 6 1 K	48/00	A 6 1 K	48/00		4 B 0 6 5
A 6 1 P	3/00	A 6 1 P	3/00		4 C 0 8 4
A 6 1 P	43/00	A 6 1 P	43/00	I O 5	4 C 0 8 7
		審査請求	未請求	予備審査請求	有 (全 119 頁) 最終頁に続

(21) 出願番号	特願2003-519472(P2003-519472)	(71) 出願人	504051445
(86) (22) 出願日	平成14年8月9日(2002.8.9)		ナショナル バブリック ヘルス インフ
(85) 翻訳文提出日	平成16年2月9日(2004.2.9)		ティテュート
(86) 国際出願番号	PCT/EP2002/008963		フィンランド エフイーエン-0030C
(87) 国際公開番号	W02003/014342		ヘルシンキ マネルヘイミンティエ
(87) 国際公開日	平成15年2月20日(2003.2.20)		I G G
(31) 優先権主張番号	01119377.8	(74) 代理人	110000109
(32) 優先日	平成13年8月10日(2001.8.10)		特許業務法人特許事務所サイクス
(33) 優先権主張国	欧州特許庁(EP)		ベルト-ネン レーナ
(31) 優先権主張番号	01119528.6	(72) 発明者	フィンランド 00251 ヘルシンキ
(32) 優先日	平成13年8月14日(2001.8.14)		ハールトマニカテュ 8 ナショナル
(33) 優先権主張国	欧州特許庁(EP)		バブリック ヘルス インスティテュート
(31) 優先権主張番号	60/315,955		バイオメディカム デハートメント オ
(32) 優先日	平成13年8月31日(2001.8.31)		ブ モレキュラー メディシン
(33) 優先権主張国	米国(US)		

最終頁に続く