

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 391 173**

51 Int. Cl.:
C12N 15/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **02794596 .3**
96 Fecha de presentación: **09.08.2002**
97 Número de publicación de la solicitud: **1417305**
97 Fecha de publicación de la solicitud: **12.05.2004**

54 Título: **Identificación de una variante de ADN asociada a hipolactasia de tipo adulto**

30 Prioridad:
10.08.2001 EP 01119377
14.08.2001 EP 01119528
31.08.2001 US 315955 P

45 Fecha de publicación de la mención BOPI:
22.11.2012

45 Fecha de la publicación del folleto de la patente:
22.11.2012

73 Titular/es:
NATIONAL PUBLIC HEALTH INSTITUTE (100.0%)
MANNERHEIMINTIE 166
00300 HELSINKI, FI

72 Inventor/es:
PELTONEN, LEENA;
ENATTAH, NABIL;
JÄRVELÄ, IRMA;
SAHI, TIMO;
SAVILAHTI, ERKKI y
TERWILLIGER, JOSEPH

74 Agente/Representante:
ARIAS SANZ, Juan

ES 2 391 173 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

Identificación de una variante de ADN asociada a hipolactasia de tipo adulto

La presente invención se refiere a una molécula de ácido nucleico que comprende o consiste en una porción 5' de un gen de lactasa-florizina hidrolasa (LPH) intestinal que contribuye a, o es indicativa de, hipolactasia de tipo adulto, en la que dicha molécula de ácido nucleico se caracteriza por las reivindicaciones adjuntas.

La presente invención se refiere adicionalmente a procedimientos para ensayar la presencia de o la predisposición a hipolactasia de tipo adulto que están basados en el análisis de un SNP contenido en la molécula de ácido nucleico anteriormente indicada. Adicionalmente, la presente invención se refiere a una composición de diagnóstico y a un kit útiles para la detección de la presencia de o la predisposición a hipolactasia de tipo adulto.

La enzima lactasa-florizina hidrolasa (LPH), que se expresa exclusivamente por células epiteliales intestinales, hidroliza la lactosa, o azúcar de la leche, a glucosa y galactosa¹. La expresión de la enzima LPH se reduce drásticamente a niveles muy bajos en el periodo de destete de mamíferos cuando la lactosa ya no es una parte esencial de la dieta. En seres humanos, la afección conocida como hipolactasia de tipo adulto o no persistencia de lactasa, afecta a la mayoría de poblaciones y limita gravemente el uso de leche fresca entre los adultos debido a la intolerancia a la lactosa. La edad de inicio del estado de no persistencia de lactasa varía entre las poblaciones, estando en el intervalo de 1-2 años de edad entre los tailandeses hasta 10-20 años de edad entre los finlandeses²⁻³. Sin embargo, en el norte de Europa y otros pocos grupos étnicos, la actividad de LPH persiste a lo largo de la vida en la mayoría de los adultos, una afección conocida como persistencia de lactasa. El fenotipo de persistencia/no persistencia de lactasa se ha mostrado que está determinado genéticamente, siendo el estado persistente dominante frente al estado no persistente⁴⁻⁶.

El diagnóstico del estado de la técnica de hipolactasia de tipo adulto está basado en el ensayo de tolerancia a lactosa (ETL). Después de un ayuno de una noche (10 horas), se administra 1 g/kg de lactosa en forma de una solución al 12,5%, siendo la dosis máxima de 50 g. Se toman muestras de sangre capilar antes y 20 y 30 min después de la ingestión de lactosa. Se determina la concentración de glucosa mediante el procedimiento de glucosa oxidasa (Hjelm y de Verdier, 1963). Se observan los síntomas abdominales el día del ETL. Se tomó una elevación máxima de la concentración sanguínea de glucosa de 1,1 mmol/l o más como signo de malabsorción de lactosa (Gudman-Hoyer y Hamum 1968, Jussila 1970, Sahi 1972). El ETL contiene un riesgo de un 10% de diagnósticos falsos positivos y negativos, concretamente, la sensibilidad y especificidad del ETL es de aproximadamente un 90% (Isokoski *et al.* 1972, Newcomer *et al.*, 1975, Sahi 1983). La exactitud del ETL puede mejorarse administrando 0,3 g/kg de etanol, que inhibe el metabolismo de galactosa en el hígado (Tygstrup y Lundqvist, 1962), y 15 min después 1 g/kg de lactosa en forma de una solución al 12,5%.

Los niños con elevaciones máximas de menos de 0,2 mg/100 ml en el primer ETL o repetido se han enviado a biopsia de intestino delgado, que se toma mediante gastroscopia. Este es un procedimiento invasivo que requiere experiencia y se efectúa habitualmente en hospitales universitarios solo por especialistas en gastroenterología. Las muestras de biopsia se examinan con un microscopio de disección e histológicamente, y se determinan las actividades maltasa, sacarasa y lactasa de mucosa (Launiala *et al.*, 1964). El diagnóstico de hipolactasia en niños se justifica si la histología de la biopsia intestinal es normal y la actividad lactasa es menor de 20 U/g de proteína y la relación de lactasa/sacarasa es menor de 0,30, o en ETL con administración de etanol, si se demuestra una elevación máxima de la concentración sanguínea de glucosa de menos de 20 mg/100 ml y de la concentración de galactosa de 5 mg/100 ml o menos (Sahi *et al.*, 1972). Como se describe anteriormente, los procedimientos actuales para diagnosticar hipolactasia de tipo adulto son laboriosos. El ETL es inexacto y, por lo tanto, se requiere un procedimiento invasivo, la gastroscopia, antes de establecer el diagnóstico. Puesto que la hipolactasia de tipo adulto es muy común y la causa principal de síntomas abdominales no específicos (en un tercio de los pacientes que se quejan de dolor de estómago), existe la clara necesidad de mejorar el diagnóstico de este problema sanitario común.

Sin embargo, hasta ahora no se ha desarrollado ningún ensayo bioquímico que sea fácil de manejar y que, al mismo tiempo, proporcione resultados rápidos y exactos. La elucidación de la causa de la enfermedad al nivel de expresión de ADN genómico ha sido igualmente fallida. Por tanto, la secuenciación de las regiones codificantes y promotoras del gen *LPH* en adultos no ha revelado variaciones de ADN que se correlacionen con la persistencia-no persistencia de lactasa, ni han surgido evidencias de variantes de corte y empalme ni variantes de edición de ARNm asociadas a este rasgo⁷⁻⁸. Estudios previos han mostrado que el rasgo de persistencia/no persistencia de lactasa está posiblemente controlado por un elemento o elementos de acción en cis que residen en o son adyacentes al gen de lactasa, y se ha observado un fuerte desequilibrio de ligamiento (DL) en el haplotipo de 70 kb que se extiende por el gen de lactasa⁹⁻¹⁰. Varios estudios reseñan evidencias de que el control principal de la expresión del gen *LPH* opera al nivel de regulación de la transcripción¹¹⁻¹³. Sin embargo, se ha sugerido que puede estar implicada en la etiología de hipolactasia de tipo adulto una variación que influye tanto en el control transcripcional como postranscripcional de la expresión del gen *LPH*¹⁴⁻¹⁵.

A la vista de lo anterior, el problema técnico subyacente de la presente invención era proporcionar medios y procedimientos que permitieran un diagnóstico exacto y conveniente de hipolactasia de tipo adulto o de una predisposición a esta enfermedad.

La solución a dicho problema técnico se logra mediante las realizaciones caracterizadas en las reivindicaciones.

Por tanto, la presente invención se refiere a una molécula de ácido nucleico que comprende o consiste en una porción 5' de un gen de lactasa-florizina hidrolasa (LPH) intestinal que contribuye a o es indicativa de hipolactasia de tipo adulto, en la que dicha molécula de ácido nucleico se caracteriza por las reivindicaciones adjuntas.

De acuerdo con la invención, el término "gen de lactasa-florizina hidrolasa (LPH) intestinal" denota un gen que codifica una enzima que tiene actividad de hidrólisis de lactosa en sus componentes glucosa y galactosa. La enzima se caracteriza como E.C. 3.2.1.23.62.

El término "hipolactasia de tipo adulto" designa una afección, también conocida como intolerancia a la lactosa, que es una afección recesiva autosómica resultante de la reducción "fisiológica" de la actividad de la enzima lactasa-florizina hidrolasa (LPH) en las células intestinales de una proporción significativa de la población mundial.

El término "contribuye a o es indicativo de hipolactasia de tipo adulto" designa el hecho de que los SNP, y por tanto las correspondientes moléculas de ácido nucleico, encontrados son indicativos de la afección y posiblemente también causantes de la misma. Por consiguiente, este término requiere necesariamente que la porción 5' indicada sea indicativa de la afección. Por otro lado, dicho término no requiere necesariamente que la porción 5' sea causante de o contribuya a la afección. Sin embargo, dicho término no excluye un papel causante o contribuyente de cualquiera o ambos SNP.

El término "que hibrida en condiciones rigurosas" designa condiciones de hibridación que son bien conocidas o que pueden establecerse por el especialista en la materia según protocolos convencionales. El término designa lo más ventajosamente las condiciones altamente rigurosas. Pueden establecerse las condiciones rigurosas apropiadas para cada secuencia basándose en parámetros bien conocidos tales como temperatura, composición de las moléculas de ácido nucleico, condiciones salinas, etc., véanse, por ejemplo: Sambrook *et al.*, "Molecular Cloning, A Laboratory Manual"; CSH Press, Cold Spring Harbor, 1989 o Higgins and Hames (eds.), "Nucleic acid hybridization, a practical approach", IRL Press, Oxford 1985 (referencia 54), y véase en particular el capítulo "Hybridization Strategy" de Britten & Davidson, 3 a 15. Las condiciones típicas (de alto rigor) comprenden hibridación a 65°C en 0,5xSSC y 0,1% de SDS o hibridación a 42°C en formamida al 50%, 4xSSC y 0,1% de SDS. La hibridación es seguida habitualmente de lavado para retirar la señal inespecífica. Las condiciones de lavado incluyen condiciones tales como 65°C, 0,2xSSC y 0,1% de SDS o 2xSSC y 0,1% de SDS o 0,3xSSC y 0,1% de SDS a 25-65°C.

Como se da a conocer en la presente memoria, la presente invención se refiere también a moléculas de ácido nucleico hibridantes de al menos 20 nucleótidos, (véanse las reivindicaciones adjuntas). Sin embargo, la presente invención se refiere también a una molécula de ácido nucleico de al menos 50, al menos 100, al menos 150 o al menos 200 nucleótidos. Preferiblemente, dichos fragmentos hibridantes comprenden al menos 25, al menos 50 o al menos 75 nucleótidos y al menos 100 nucleótidos en 5' y 3' de la posición -13910 como se define en las reivindicaciones adjuntas, como se define en las reivindicaciones adjuntas.

El término "molécula de ácido nucleico" designa tanto moléculas de ácido nucleico de origen natural como no natural. Las moléculas de ácido nucleico de origen no natural incluyen ADNc así como derivados tales como PNA.

El término "molécula de ácido nucleico [...] que comprende la secuencia de ácido nucleico de SEQ ID NO:" a lo largo de esta memoria descriptiva designa moléculas de ácido nucleico que son al menos 1 nucleótido más largas que la molécula de ácido nucleico especificada por la SEQ ID NO. Al mismo tiempo, estas moléculas de ácido nucleico se extienden, como máximo, 30.000 nucleótidos más allá del extremo 5' y/o 3' de la molécula de ácido nucleico de la invención especificada, por ejemplo, por la SEQ ID NO: 2 o 1, 3 o 4.

Sorprendentemente, se ha encontrado de acuerdo con la presente invención que las dos variantes asociadas a hipolactasia se localizan a una distancia considerable del gen *LPH*, colocadas en diferentes intrones del gen *MCM6*. El *MCM6* es un miembro de una familia génica (*MCM* 2-7) necesaria para la iniciación de la replicación de ADN que asegura que tiene lugar solo una vez durante el ciclo celular³¹. El *MCM6*, al contrario que el *LPH*, no está limitado en su distribución de tejido y no hay correlación en los niveles de transcritos de *MCM6* y *LPH*¹⁸. Estos hallazgos sugerían que estos dos genes no comparten ningún elemento de acción en cis funcionalmente significativo que proporcione especificidad de tejido o regulación de desarrollo¹⁸. Lo más probablemente, las variantes identificadas tienen diferente significado funcional para la expresión de los genes *LPH* y *MCM6*. Sorprendentemente además, basándose en la asociación total a hipolactasia, están asociados (o uno de ellos) a una regulación negativa dependiente de la edad del nivel de transcrito del gen *LPH* en el epitelio intestinal, pero tienen poco o ningún efecto sobre la transcripción de *MCM6*.

Experimentalmente, usando ligamiento, asociación alélica y análisis de haplotipo extendido llevados a cabo en 9 familias finlandesas extensas, se limitó el locus de hipolactasia de tipo adulto a un intervalo de 48 kb en 2q21. El análisis de secuencia de la región reveló un polimorfismo mononucleotídico (SNP), C/T-13910, que se cosegregaba completamente con hipolactasia de tipo adulto en todas las familias finlandesas y en un conjunto de muestras de 236 individuos de 4 poblaciones diferentes. Otro SNP, G/A-22018 que reside a 8 kb en dirección telomérica de C/T-13910, estaba asociado al rasgo en todos menos en 7 casos. La prevalencia del SNP C/T-13910 en 1047 muestras de ADN reflejaba la prevalencia reseñada de la hipolactasia de tipo adulto en 3 poblaciones diferentes, proporcionando evidencias adicionales de su importancia para el rasgo.

- El sorprendente hallazgo designado anteriormente por primera vez permite el establecimiento de sistemas de ensayo que están basados en el análisis molecular de los polimorfismos mononucleotídicos indicados en dirección 5' del gen *LPH*. Aunque que ambos SNP proporcionan una base sólida para el diagnóstico de, o el diagnóstico de la predisposición a, hipolactasia de tipo adulto, se prefiere analizar la posición nucleotídica -13910, sola o en combinación con la posición nucleotídica -22018. Esto es debido a que el SNP en posición -13910 estaba asociado en un 100% de los casos analizados a la enfermedad, mientras que el SNP en posición -22018 estaba asociado solo en un 98% de todos los casos con hipolactasia de tipo adulto. No obstante, los análisis de la posición nucleotídica -22018 sola proporcionarán habitualmente también una base firme para el diagnóstico de predisposición a hipolactasia de tipo adulto.
- Debido a la abundancia de procedimientos establecidos para evaluar la presencia de SNP, ahora es posible diagnosticar una predisposición genética a hipolactasia de tipo adulto convenientemente, en un corto espacio de tiempo, a bajo coste, con gran exactitud y sin molestias significativas para la persona bajo investigación.
- La invención se refiere adicionalmente a una molécula de ácido nucleico que comprende o consiste en una porción 5' de un gen de lactasa-florizina hidrolasa (*LPH*) intestinal, en la que dicha molécula de ácido nucleico está caracterizada por las reivindicaciones adjuntas.
- Esta realización de la presente invención puede usarse convenientemente para demostrar que una persona no padece hipolactasia de tipo adulto y no tiene predisposición a la misma. Adicionalmente, esta molécula de ácido nucleico que refleja la situación "de tipo silvestre" de la posición -13910 o -22018 en dirección 5' del gen *LPH* puede usarse como medio de control en experimentos en que se ensaya la predisposición a hipolactasia de tipo adulto.
- Para el ensayo, pueden usarse procedimientos como se describen a lo largo de esta memoria descriptiva.
- En una realización preferida de la invención, la molécula de ácido nucleico es ADN genómico. Esta realización preferida de la invención refleja el hecho de que habitualmente el análisis se llevaba a cabo basándose en el ADN genómico de fluido corporal, células o tejido aislado de la persona bajo investigación.
- En una realización preferida adicional de la molécula de ácido nucleico de la invención, dicho ADN genómico es parte de un gen. De acuerdo con la invención, se prefiere analizar al menos uno de los intrones del gen *MCM6* albergado en posición -13910 o posición -22018 respecto al gen *LPH*.
- Además, la invención se refiere a un fragmento de molécula de ácido nucleico como se describe anteriormente en la presente memoria que tiene al menos 14 nucleótidos, en el que dicho fragmento comprende la posición nucleotídica -13910 o la posición nucleotídica -22018 (en dirección 5') del gen *LPH*. El fragmento de la invención puede ser de origen natural así como (semi)sintético. Por tanto, el fragmento puede ser, por ejemplo, una molécula de ácido nucleico que se ha sintetizado de acuerdo con protocolos convencionales de la química orgánica. De forma importante, el fragmento de ácido nucleico de la invención comprende la posición nucleotídica -13910 o la posición nucleotídica -22018 en dirección 5' del gen *LPH*. En estas posiciones, el fragmento puede tener el nucleótido de tipo silvestre o el nucleótido que contribuye a o es indicativo de hipolactasia de tipo adulto (también designado como la secuencia "mutante"). En consecuencia, el fragmento de la invención puede usarse, por ejemplo, en ensayos para diferenciar entre la secuencia de tipo silvestre y mutante. Se prefiere adicionalmente que el fragmento de la invención consista en al menos 17 nucleótidos, más preferiblemente al menos 21 nucleótidos y lo más preferiblemente al menos 25 nucleótidos, tal como 30 nucleótidos.
- Además, la invención se refiere a una molécula de ácido nucleico que es complementaria de la molécula de ácido nucleico como se describe anteriormente en la presente memoria.
- Esta realización de la invención que comprende al menos 14 nucleótidos y cubre al menos la posición -13910 o la posición -22018 de la secuencia en dirección 5' del gen *LPH* es particularmente útil en el análisis de la configuración genética en las posiciones indicadas en ensayos de hibridación. Por tanto, por ejemplo, puede usarse un pentadecámero exactamente complementario de la secuencia de tipo silvestre (concretamente, una T en posición -13910 o una A en posición -22018) o de las variantes que contribuyen a o son indicativas de hipolactasia de tipo adulto (concretamente, una C en posición -13910 o una G en posición -22018) para diferenciar entre las variantes polimórficas. Esto es debido a que una molécula de ácido nucleico marcada con un marcador detectable no exactamente complementario con el ADN de la muestra analizada no dará lugar a una señal detectable, si se eligen condiciones de hibridación y lavado apropiadas.
- A este respecto, es importante observar que la molécula de ácido nucleico de la invención, el fragmento de la misma, así como la molécula de ácido nucleico complementaria pueden marcarse detectablemente. Los marcadores detectables incluyen marcadores radiactivos tales como ^3H o ^{32}P o marcadores fluorescentes. El marcaje de ácidos nucleicos es bien entendido en la materia y se describe, por ejemplo, en Sambrook *et al.*, *loc. cit.*
- Además, la invención se refiere a un vector que comprende la molécula de ácido nucleico como se describe anteriormente en la presente memoria. El vector de la invención puede contener una molécula de ácido nucleico que comprende la secuencia o secuencias de tipo silvestre o puede contener una molécula de ácido nucleico que comprende la secuencia o secuencias mutantes. Los vectores pueden ser particularmente plásmidos, cósmidos, virus o

bacteriófagos usados convencionalmente en ingeniería genética que comprenden la molécula de ácido nucleico de la invención. Preferiblemente, dicho vector es un vector de expresión y/o un vector de transferencia génica u direccionador. Los vectores de expresión derivados de virus tales como retrovirus, virus Vaccinia, virus adenoasociados, herpesvirus o papilomavirus bovinos pueden usarse para el suministro de la molécula de ácido nucleico de la invención a la población celular diana. Pueden usarse procedimientos que son bien conocidos por los especialistas en la materia para construir vectores víricos recombinantes; véanse, por ejemplo, las técnicas descritas en Sambrook *et al.*, *loc. cit.* y Ausubel *et al.*, "Current Protocols in Molecular Biology", Green Publishing Associates and Wiley Interscience, N.Y. (1989). Como alternativa, las moléculas de ácido nucleico y vectores de la invención pueden reconstituirse en liposomas para suministro a células diana. Los vectores que contienen las moléculas de ácido nucleico de la invención pueden transferirse a la célula hospedadora mediante procedimientos bien conocidos, que varían dependiendo del tipo de hospedador celular. Por ejemplo, la transfección con cloruro de calcio se utiliza comúnmente para células procarióticas, mientras que pueden usarse, por ejemplo, transfección mediada por fosfato de calcio o DEAE-dextrano o electroporación para otros hospedadores celulares; véase Sambrook, *supra*.

Dichos vectores pueden comprender genes adicionales tales como genes marcadores que permitan la selección de dicho vector en una célula hospedadora adecuada y en condiciones adecuadas. Preferiblemente, la molécula de ácido nucleico de la invención está ligada operativamente con secuencias de control de la expresión en células procarióticas o eucarióticas. La expresión de dicho polinucleótido comprende la transcripción del polinucleótido a un ARNm traducible. Los elementos reguladores que aseguran la expresión en células eucarióticas, preferiblemente células de mamífero, son bien conocidos por los especialistas en la materia. Comprenden habitualmente secuencias reguladoras que aseguran la iniciación de la transcripción y, opcionalmente, una señal de poli-A que asegura la terminación de la transcripción y la estabilización del transcrito, y/o un intrón que potencia adicionalmente la expresión de dicho polinucleótido. Los elementos reguladores adicionales pueden incluir potenciadores transcripcionales así como traduccionales y/o regiones promotoras asociadas naturalmente o heterólogas. Los posibles elementos reguladores que permiten la expresión en células hospedadoras procarióticas comprenden, por ejemplo, el promotor PL, lac, trp o tac en *E. coli*, y son ejemplos de elementos reguladores que permiten la expresión en células hospedadoras eucarióticas el promotor AOX1 o GAL1 en levadura o el promotor CMV, SV40 o RSV (virus del sarcoma de Rous), potenciador CMV, potenciador SV40 o intrón de globina en células de mamíferos y otros animales. Aparte de los elementos que son responsables de la iniciación de la transcripción, dichos elementos reguladores pueden comprender también señales de terminación de la transcripción tales como el sitio poli-A de SV40 o el sitio poli-A de tk, en dirección 3' del polinucleótido. Opcionalmente, la secuencia heteróloga puede codificar una proteína de fusión que incluya un péptido de identificación C- o N-terminal que confiera características deseadas, por ejemplo, estabilización o purificación simplificada del producto recombinante expresado. En este contexto, son conocidos en la materia vectores de expresión adecuados tales como el vector de expresión de ADNc de Okayama-Berg pcDV1 (Pharmacia), pCDM8, pRc/CMV, pcDNA1, pcDNA3, el sistema de clonación Echo™ (Invitrogen), pSPORT1 (GIBCO SRL) o pRevTet-On/pRevTet-Off o pCI (Promega). Preferiblemente, las secuencias de control de la expresión serán sistemas promotores eucarióticos en vectores capaces de transformar o transfectar células hospedadoras eucarióticas, pero pueden usarse también secuencias de control para hospedadores procarióticos. Como se menciona anteriormente, el vector de la presente invención puede ser también un vector de transferencia génica u direccionador. La terapia génica, que está basada en introducir genes terapéuticos en células mediante técnicas *ex vivo* o *in vivo*, es una de las aplicaciones más importantes de la transferencia génica. Se describen en la bibliografía vectores y procedimientos adecuados para la terapia génica *in vitro* o *in vivo*, y sin conocidos por el especialista en la materia; véanse, por ejemplo, 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; W094/29469; WO 97/00957, Schaper, *Current Opinion in Biotechnology* 7 (1996), 635-640 o Kay *et al.* (2001) *Nature Medicine*, 7, 33-40 y las referencias citadas en los mismos. Los polinucleótidos y vectores de la invención pueden diseñarse para la introducción directa o para la introducción mediante liposomas o vectores víricos (por ejemplo, adenovíricos o retrovíricos) en la célula. Preferiblemente, dicha célula es una célula de línea germinal, célula embrionaria u óvulo o derivada de las mismas, lo más preferiblemente dicha célula es un citoblasto. La terapia génica se prevé solo con la molécula de ácido nucleico de tipo silvestre.

La invención se refiere también a un cebador o par de cebadores, en la que el cebador o par de cebadores hibrida en condiciones (altamente) rigurosas con el ácido nucleico como se describe anteriormente en la presente memoria que comprende la posición nucleotídica -13910 o -22018 del gen *LPH* o con la hebra complementaria del mismo. Preferiblemente, los cebadores de la invención tienen una longitud de al menos 14 nucleótidos, tal como 17 o 21 nucleótidos. Se prefiere adicionalmente que los cebadores tengan una longitud máxima de 24 nucleótidos. La hibridación o falta de hibridación de un cebador en condiciones apropiadas con una secuencia genómica que comprende la posición -13910 o la posición -22018, acoplada con un procedimiento de detección apropiado tal como una reacción de alargamiento o una reacción de amplificación, puede usarse para diferenciar entre las variantes polimórficas, y extraer entonces conclusiones con respecto, por ejemplo, a la predisposición de la persona bajo investigación a hipolactasia de tipo adulto. La presente invención prevé dos tipos de cebadores/pares de cebadores. Un tipo hibrida con una secuencia que comprende la secuencia mutante. En otras palabras, el cebador es exactamente complementario de una secuencia que contiene C en posición -13910 o G en posición -22018 o de la hebra complementaria del mismo. El otro tipo de cebador es exactamente complementario de una secuencia que tiene T en posición -13910 o A en posición -22018 o de la hebra complementaria del mismo. Puesto que las condiciones de hibridación se elegirían preferiblemente para ser suficientemente rigurosas, la puesta en contacto de, por ejemplo, un cebador exactamente complementario de la secuencia mutante con un alelo de tipo silvestre no daría como resultado

una hibridación eficaz debido a la formación de desapareamiento. Después del lavado, no se detectaría señal debido a la retirada del cebador.

Adicionalmente, la invención se refiere a un hospedador no humano transformado con el vector de la invención como se describe anteriormente en la presente memoria. El hospedador puede portar la secuencia mutante o de tipo silvestre. Tras el cultivo y demás, el hospedador puede ser heterocigótico u homocigótico de uno o ambos SNP. El hospedador de la invención puede portar el vector de la invención integrado transitoria o establemente en el genoma. Los procedimientos para generar el hospedador no humano de la invención son bien conocidos en la materia. Por ejemplo, pueden emplearse protocolos de transfección convencionales descritos en Sambrook *et al.*, *loc. cit.*, para generar bacterias transformadas (tales como *E. coli*) o levaduras transformadas. El hospedador no humano de la invención puede usarse, por ejemplo, para dilucidar el inicio de la hipolactasia de tipo adulto.

En una realización preferida de la invención, el hospedador no humano es una bacteria, una célula de levadura, una célula de insecto, una célula fúngica, una célula de mamífero, una célula de planta, un animal transgénico o una planta transgénica. Mientras que la *E. coli* es una bacteria preferida, las células de levadura preferidas son células de *S. cerevisiae* o *Pichia pastoris*. Son células fúngicas preferidas las células de *Aspergillus* y las células de insecto preferidas incluyen células de *Spodoptera frugiperda*. Son células de mamífero preferidas las líneas celulares de carcinoma de colon que muestran expresión de la enzima LPH, e incluyen células CaCO2.

Un procedimiento para la producción de un animal transgénico no humano, por ejemplo un ratón transgénico, comprende la introducción del polinucleótido o vector direccionador anteriormente mencionado en una célula germinal, una célula embrionaria, citoblasto u óvulo o célula derivada de las mismas. El animal no humano puede usarse de acuerdo con un procedimiento de cribado de la invención descrito en la presente memoria. La producción de embriones transgénicos y el cribado de los mismos puede efectuarse, por ejemplo, como se describe por A. L. Joyner Ed., "Gene Targeting, A Practical Approach" (1993), Oxford University Press. El ADN de las membranas embrionarias de embriones puede analizarse usando, por ejemplo, transferencias Southern con una molécula de ácido nucleico complementaria apropiada; véase *supra*. Se describe en la materia un procedimiento general para preparar animales transgénicos no humanos, véase por ejemplo el documento WO 94/24274. Para preparar organismos transgénicos no humanos (que incluyen animales no humanos modificados de forma dirigida homológamente), se prefieren citoblastos embrionarios (células ES). Las células ES de murino, tales como la línea AB-1 cultivada en capas de alimentación celular SNL76/7 mitóticamente inactivas (McMahon y Bradley, *Cell* 62: 1073-1085 (1990)) esencialmente como se describen por (Robertson, E. J. (1987) en "Teratocarcinomas and Embryonic Stem Cells: A Practical Approach". E. J. Robertson, ed. (Oxford: IRL Press), pág. 71-112) pueden usarse para modificación genética dirigida homóloga. Otras líneas de ES adecuadas incluyen, pero sin limitación, la línea E14 (Hooper *et al.*, *Nature* 326: 292-295 (1987)), la línea D3 (Doetschman *et al.*, *J. Embryo J. Exp. Morph.* 87: 27-45 (1985)), la línea CCE (Robertson *et al.*, *Nature* 323: 445-448 (1986)), la línea AK-7 (Zhuang *et al.*, *Cell* 77: 875-884 (1994)). El éxito de generar una línea de ratón a partir de células ES portadoras de una mutación dirigida específica depende de la pluripotencia de las células ES (concretamente, su capacidad, una vez inyectadas en un embrión en desarrollo hospedador, tal como un blastocito o mórula, de participar en la embriogénesis y contribuir a las células germinales del animal resultante). Los blastocitos que contienen las células ES inyectadas se dejan desarrollar en los úteros de hembras no humanas pseudoembarazadas y nacen como ratones quiméricos. Los ratones transgénicos resultantes que son quiméricos de células que tienen la molécula de ácido nucleico deseada, se retrocruzan y se criba la presencia del transgén o transgenes correctamente dirigidos por PCR o análisis de transferencia Southern en el ADN de biopsia de cola de la progenie, para identificar los ratones transgénicos heterocigóticos de la molécula de ácido nucleico de la invención.

Los animales transgénicos no humanos pueden ser, por ejemplo, ratones, ratas, hámsteres, perros, monos (simios), conejos, cerdos o vacas transgénicos. Preferiblemente, dicho animal transgénico no humano es un ratón. Los animales transgénicos de la invención son, entre otras cosas, útiles para estudiar la expresión/resultado fenotípico de los ácidos nucleicos y vectores de la presente invención. Además, los animales transgénicos de la presente invención son útiles para estudiar la expresión de desarrollo de la enzima LPH, por ejemplo en el intestino de roedor. Se prevé además que los animales transgénicos no humanos de la invención puedan emplearse para ensayar los agentes/composiciones terapéuticos u otras posibles terapias que sean útiles para mejorar la hipolactasia de tipo adulto.

Además, la invención se refiere a un anticuerpo o aptámero o fago que se une específicamente a la molécula de ácido nucleico mutante de la invención, pero no a la correspondiente molécula de ácido nucleico de tipo silvestre. Puede ensayarse la unión del anticuerpo y usarse en cualquier técnica serológica bien conocida en la materia, tal como técnicas de aglutinación en tubos, geles, fase sólida y técnicas de captura con o sin anticuerpos secundarios, o en citometría de flujo con o sin potenciación de inmunofluorescencia (véanse, por ejemplo, las técnicas descritas en Harlow y Lane "Antibodies, A Laboratory Manual", CSH Press, Cold Spring Harbor, EE.UU., 1988 (véase la referencia 53).

En consonancia con la invención, el anticuerpo reconoce específicamente un epítipo que comprende la posición -13910 (en la que el nucleótido es C) o la posición -22018 (en la que el nucleótido es G). No reacciona de forma cruzada, o no esencialmente, con un epítipo que comprende la posición -13910 con una T en esta posición ni con un epítipo que comprende la posición -22018 con una G en esta posición. La especificidad del anticuerpo que puede generarse según protocolos estándares puede ensayarse poniéndolo en contacto con moléculas de ADN portadoras de la secuencia de tipo silvestre y mutante tal como en un ensayo ELISA. Se seleccionarán solo aquellos anticuerpos que produzcan una señal superior al fondo con la secuencia mutante pero no con la secuencia de tipo silvestre.

El anticuerpo de la invención puede ser un anticuerpo monoclonal o un anticuerpo derivado de o comprendido en un antisuero policlonal. El término "anticuerpo", como se usa de acuerdo con la presente invención, comprende adicionalmente fragmentos de dicho anticuerpo tales como fragmentos Fab, F(ab')₂, Fv o scFv; véase, por ejemplo, Harlow y Lane⁵³, *loc. cit.* El anticuerpo o fragmento del mismo puede ser de origen natural o puede producirse (semi)sintéticamente. Dichos productos sintéticos comprenden también material no proteico como material semiproteico que tiene la misma, o esencialmente la misma, especificidad de unión que el anticuerpo de la invención. Dichos productos pueden obtenerse, por ejemplo, mediante peptidomiméticos.

El término "aptámero" es bien conocido en la materia y se define, por ejemplo, en Osborne *et al.*, Curr. Opin. Chem. Biol. 1 (1997), 5-9 (véase la referencia 51) o en Stall y Szoka, Pharm. Res. 12 (1995), 465-483 (véase la referencia 52).

- 10 Además, la invención se refiere a un anticuerpo o aptámero o fago que se une específicamente a la molécula de ácido nucleico de tipo silvestre como se describe anteriormente en la presente memoria pero no a la correspondiente secuencia mutante que contribuye a o es indicativa de hipolactasia de tipo adulto. Las afirmaciones con respecto a la especificidad y demás realizadas para el anticuerpo que es específico de la secuencia mutante se aplican aquí cambiando lo necesario.
- 15 Además, la invención se refiere a una composición farmacéutica que comprende las moléculas de ácido nucleico de tipo silvestre descritas anteriormente en la presente memoria.

La composición farmacéutica de la invención puede usarse en enfoques de terapia génica, particularmente en terapia génica somática. La molécula de ácido nucleico de tipo silvestre designada anteriormente y contenida en la composición farmacéutica de la invención puede combinarse con un portador y/o diluyente farmacéuticamente aceptable. Los ejemplos de portadores farmacéuticos adecuados son bien conocidos en la materia e incluyen soluciones salinas tamponadas con fosfato, agua, emulsiones tales como emulsiones de aceite en agua, diversos tipos de agentes humectantes, soluciones estériles, etc. Las composiciones que comprenden dichos portadores pueden formularse mediante procedimientos convencionales bien conocidos. Estas composiciones farmacéuticas pueden administrarse al sujeto a una dosis adecuada. La administración de las composiciones adecuadas puede efectuarse por diferentes vías, por ejemplo, mediante administración intravenosa, intraperitoneal, subcutánea, intramuscular, tópica, intradérmica, intranasal o intrabronquial. El régimen de dosificación se determinará por el médico a cargo y según factores clínicos. Como es bien conocido en la técnica médica, las dosificaciones de cualquier paciente dependen de muchos factores, incluyendo tamaño del paciente, área superficial corporal, edad, el compuesto particular a administrar, sexo, momento y ruta de administración, salud general y otros fármacos que se estén administrando simultáneamente. Una dosis típica puede estar, por ejemplo, en el intervalo de 0,001 a 1000 µg de ácido nucleico para expresión o para inhibición de la expresión; sin embargo, se conciben dosis inferiores o superiores a este intervalo ejemplar, considerando especialmente los factores anteriormente mencionados. Las dosificaciones variarán, pero es una dosificación preferida para administración intravenosa de ADN de aproximadamente 10⁶ a 10¹² copias de la molécula de ADN. La progresión puede controlarse mediante evaluación periódica. Las composiciones de la invención pueden administrarse por vía local o sistémica. La administración será generalmente parenteral, por ejemplo intravenosa; el ADN puede administrarse también directamente al sitio diana, por ejemplo, mediante suministro biolístico a un sitio diana interno o externo o mediante catéter a un sitio en una arteria. Las preparaciones para administración parenteral incluyen soluciones, suspensiones y emulsiones estériles acuosas o no acuosas. Son ejemplos de disolventes no acuosos propilenglicol, polietilenglicol, aceites vegetales tales como aceite de oliva y ésteres orgánicos inyectables tales como oleato de etilo. Los portadores acuosos incluyen agua, soluciones, emulsiones o suspensiones alcohólicas/acuosas, incluyendo medios salinos y tamponados. Los vehículos parenterales incluyen solución de cloruro de sodio, dextrosa de Ringer, dextrosa y cloruro de sodio, solución de Ringer-lactato o aceites no volátiles. Los vehículos intravenosos incluyen reponedores de fluidos y nutrientes, reponedores de electrolitos (tales como aquellos basados en dextrosa de Ringer) y similares. Pueden estar presentes también conservantes y otros aditivos tales como, por ejemplo, agentes antimicrobianos, antioxidantes, quelantes y gases inertes y similares.

Adicionalmente, la invención se refiere a una composición de diagnóstico que comprende la molécula de ácido nucleico como se describe anteriormente en la presente memoria, el vector como se describe anteriormente en la presente memoria, el cebador o par de cebadores como se describe anteriormente en la presente memoria y/o el anticuerpo, aptámero y/o fago como se describe anteriormente en la presente memoria. La composición de diagnóstico es útil para evaluar el estado genético de una persona con respecto a su predisposición a desarrollar hipolactasia de tipo adulto con respecto al diagnóstico de la afección aguda. Los diversos posibles componentes de la composición de diagnóstico pueden envasarse en uno o más viales, en un disolvente o de otro modo tal como en forma liofilizada. Si se disuelve en un disolvente, la composición de diagnóstico se enfría preferiblemente al menos a +8 a +4°C. Puede preferirse la congelación en otros casos.

55 La invención se refiere también a un procedimiento para ensayar la presencia de o predisposición a hipolactasia de tipo adulto o un rasgo asociado, que comprende ensayar en una muestra obtenida de un paciente previsible o de una persona sospechosa de portar dicha predisposición la presencia de la molécula de ácido nucleico como se describe anteriormente en la presente memoria en estado homocigótico o heterocigótico. En diversas realizaciones, puede ensayarse la presencia de la secuencia o secuencias de tipo silvestre o la secuencia o secuencias mutantes.

El procedimiento de la invención es útil para detectar la configuración genética de dicha persona/paciente y extraer las conclusiones apropiadas de si la afección que padece dicho paciente es hipolactasia de tipo adulto. Como alternativa, puede evaluarse si una persona que no padece la afección porta la predisposición a hipolactasia de tipo adulto. Con respecto a la posición -13910 en dirección 5' del gen *LPH*, solo si se encuentra citosina en estado homocigótico se diagnosticaría una afección como hipolactasia de tipo adulto o se manifestaría la correspondiente predisposición. Por otro lado, si se encuentra timidina en estado homocigótico o si el individuo es heterocigótico (C/T), entonces puede concluirse que la afección que padece un paciente no está relacionada con hipolactasia de tipo adulto y, además, que el paciente no porta la predisposición a desarrollar esta afección. Sin embargo, puede concluirse que los hijos de personas que portan el genotipo heterocigótico pueden desarrollar la afección si el cromosoma que porta el residuo C se empareja con el correspondiente cromosoma del otro progenitor.

La situación es similar y se aplican esencialmente las mismas conclusiones para el análisis del SNP en posición -22018. Un residuo G de aparición homocigótica marca la predisposición a o la aparición de hipolactasia de tipo adulto aguda. Un estado G/A heterocigótico se correlaciona con una alta probabilidad de no desarrollar la afección. Los individuos que portan A en estado homocigótico no se esperaría que desarrollaran la afección. De forma similar, los pacientes que padecen una afección se diagnosticarían por no padecer hipolactasia de tipo adulto.

En una realización preferida del procedimiento de la invención, dicho ensayo comprende hibridar la molécula de ácido nucleico complementaria como se describe anteriormente en la presente memoria que es complementaria de la molécula de ácido nucleico que contribuye a o es indicativa de hipolactasia de tipo adulto, o la molécula de ácido nucleico como se describe anteriormente en la presente memoria que es complementaria de la secuencia de tipo silvestre, como sonda en condiciones (altamente) rigurosas con moléculas de ácido nucleico comprendidas en dicha muestra, y detectar dicha hibridación. De nuevo, dependiendo de la sonda de ácido nucleico usada, se detectarían secuencias de tipo silvestre o mutante (concretamente, secuencias que contribuyen a o son indicativas de hipolactasia de tipo adulto). Se entiende que se elegirían condiciones de hibridación tales que una molécula de ácido nucleico complementaria de secuencias de tipo silvestre no hibridaría, o esencialmente no hibridaría, con la secuencia mutante. De forma similar, una molécula de ácido nucleico complementaria de la secuencia mutante no hibridaría, o esencialmente no hibridaría, con la secuencia de tipo silvestre. Para diferenciar entre los resultados obtenidos de genotipos homocigóticos y heterocigóticos en los procedimientos de hibridación de la invención, puede por ejemplo controlarse/detectarse la potencia/intensidad de la señal de detección respectiva después de la hibridación. Para diferenciar entre los alelos homocigótico de tipo silvestre, heterocigótico y/o homocigótico mutante en los procedimientos de hibridación de la invención, se incluirán en el análisis muestras de control interno de los correspondientes genotipos.

En una realización preferida adicional, el procedimiento de la invención comprende adicionalmente digerir el producto de dicha hibridación con una endonucleasa de restricción o someter el producto de dicha hibridación a digestión con una endonucleasa de restricción y analizar el producto de dicha digestión. Esta realización preferida de la invención permite, mediante medios convenientes, la diferenciación entre hibridación eficaz e hibridación no eficaz. Por ejemplo, si la secuencia de ADN adyacente a la posición -13910 o la posición -22018 comprende un sitio de restricción de endonucleasa, el producto hibridado será escindible por una enzima de restricción apropiada tras una hibridación eficaz, mientras que la falta de hibridación no proporcionará un producto bicatenario o no comprenderá el sitio de restricción reconocible y, por consiguiente, no se escindirán. En particular, las enzimas de restricción específicas para la secuencia de la variante de ADN C/T₋₁₃₉₁₀ es CviI I, para la variante de ADN G/A₋₂₂₀₁₈ son HhaI y AclI. Dichas enzimas de restricción que cortan rg/cy se encontraron mediante el uso del programa Webcutter. El análisis del producto de digestión puede efectuarse mediante medios convencionales, tales como mediante electroforesis en gel que puede combinarse opcionalmente con tinción del ácido nucleico, por ejemplo, con bromuro de etidio. Las combinaciones con técnicas adicionales tales como transferencia Southern están también previstas.

La detección de dicha hibridación puede efectuarse, por ejemplo, mediante un anticuerpo anti-ADN bicatenario o empleando un oligonucleótido marcado. Convenientemente, se emplea el procedimiento de la invención junto con técnicas de transferencia tales como transferencia Southern o Northern y técnicas relacionadas. El marcaje puede efectuarse, por ejemplo, mediante protocolos estándares e incluye marcaje con marcadores radiactivos, marcadores fluorescentes, fosforescentes, quimioluminiscentes, enzimáticos, etc. (véase también anteriormente).

De acuerdo con lo anterior, en otra realización preferida del procedimiento de la invención, dicha sonda se marca detectablemente, por ejemplo mediante los procedimientos y con los marcadores descritos anteriormente en la presente memoria.

En aún otra realización preferida del procedimiento de la invención, dicho ensayo comprende determinar la secuencia de ácido nucleico de al menos una porción de la molécula de ácido nucleico como se describe anteriormente en la presente memoria, comprendiendo dicha porción la posición nucleotídica -13910 y/o la posición nucleotídica -22018 del gen *LPH*. La determinación de la molécula de ácido nucleico puede efectuarse de acuerdo con uno de los protocolos convencionales tales como los protocolos de Sanger o Maxam/Gilbert (véase Sambrook *et al.*, *loc. cit.*, para una guía adicional).

En una realización preferida adicional del procedimiento de la invención, se efectúa la determinación de la secuencia de ácido nucleico mediante minisequenciación en fase sólida. La minisequenciación en fase sólida está basada en el análisis cuantitativo del nucleótido de tipo silvestre y mutante en una solución. En primer lugar, se amplifica la región

genómica que contiene la mutación mediante PCR con un cebador biotinilado y otro no biotinilado, en que el cebador biotinilado está unido a una placa recubierta con estreptavidina (SA). Se desnaturaliza el producto de PCR hasta una forma monocatenaria para permitir al cebador de miniselección unirse a esta hebra justo antes del sitio de mutación. Se añaden los nucleótidos mutados y de tipo silvestre marcados con tritio (^3H) o fluorescencia junto con dNTP no marcados a la reacción de miniselección y se secuencian usando polimerasa Taq. El resultado está basado en la cantidad de nucleótidos de tipo silvestre y mutantes en la reacción medidos por contador beta o fluorómetro y expresados como relación de R. Véanse también Syvänen AC, Sajantila A, Lukka M. *Am J. Hum. Genet.* 1993; 52, 46-59 y Suomalainen A y Syvänen AC. *Methods Mol. Biol.* 1996; 65: 73-79. Una realización preferida del procedimiento de la invención comprende adicionalmente, antes de determinar dicha secuencia de ácido nucleico, la amplificación de al menos una porción de dicha molécula de ácido nucleico. Preferiblemente, se efectúa la amplificación mediante reacción en cadena de la polimerasa (PCR). Pueden emplearse también otros procedimientos de amplificación tales como reacción en cadena de la ligasa.

En una realización preferida del procedimiento de la invención, dicho ensayo comprende llevar a cabo una reacción de amplificación en la que al menos uno de los cebadores empleados en dicha reacción de amplificación es el cebador que se describe anteriormente en la presente memoria o pertenece al par de cebadores como se describe anteriormente en la presente memoria, que comprende ensayar el producto de amplificación. En esta realización, y dependiendo de la información que el investigador/facultativo desee obtener, pueden emplearse cebadores que hibriden con las secuencias de tipo silvestre o mutantes.

El procedimiento de la invención dará como resultado la amplificación de solo la secuencia diana si dicha secuencia diana porta una secuencia exactamente complementaria del cebador usado para hibridación. Esto es debido a que el cebador oligonucleotídico preferiblemente no hibridará en condiciones de hibridación (altamente) rigurosas con la secuencia de tipo silvestre/mutante- dependiendo de cuál tipo de cebador se use- (con la consecuencia de que no se obtiene producto de amplificación) sino solo con la secuencia emparejada exactamente. Naturalmente, pueden usarse combinaciones de pares cebadores que hibridan con ambos SNP. En este caso, el análisis de los productos de amplificación esperados (que pueden ser ninguno, uno, dos, tres o cuatro productos de amplificación si el segundo cebador no diferenciador es el mismo para cada locus) proporcionará información sobre el estado genético de ambas posiciones -13910 y -22018.

En una realización preferida del procedimiento de la invención, dicha amplificación se efectúa mediante, o dicha amplificación es, la reacción en cadena de la polimerasa (PCR). La PCR está bien establecida en la materia. Las condiciones típicas para usarse de acuerdo con la presente invención incluyen, por ejemplo, un total de 35 ciclos con un total de 50 μl de volumen, ejemplificadas por una etapa de desnaturalización a 93°C durante 3 minutos; una etapa de asociación a 55°C durante 30 segundos; una etapa de extensión a 72°C durante 75 segundos y una etapa de extensión final a 72°C durante 10 minutos.

La invención se refiere además a un procedimiento para ensayar la presencia de o predisposición a hipolactasia de tipo adulto que comprende ensayar en una muestra obtenida de un ser humano la unión específica al anticuerpo o aptámero o fago como se describe anteriormente en la presente memoria. En este contexto, una tinción más débil por la presencia del antígeno de la invención en comparación con las muestras de control de tipo silvestre homocigóticas (que comprenden dos alelos persistentes) es indicativa del tipo silvestre heterocigótico (un alelo persistente y un alelo hipolactásico), mientras que para el individuo hipolactásico homocigótico, no se espera tinción si se usa el anticuerpo apropiado. Preferiblemente, se efectúa el procedimiento de la invención en presencia de muestras de control correspondientes a las tres combinaciones alélicas posibles como controles internos. El ensayo puede llevarse a cabo con un anticuerpo y demás específico de la secuencia de tipo silvestre o específico de la secuencia mutante. El ensayo de unión puede implicar, de nuevo el empleo de técnicas estándares tales como ELISA; véase, por ejemplo, Harlow y Lane⁵³, *loc. cit.*

En una realización preferida del procedimiento de la invención, dicho anticuerpo o aptámero o fago está marcado detectablemente. Mientras que los aptámeros están preferiblemente marcados radiactivamente con ^3H o ^{32}P o con un marcador fluorescente como se describe anteriormente, el fago o anticuerpo puede estar marcado de la correspondiente manera (con ^{125}I con marcador radiactivo preferido) o estar marcado con un marcador tal como el marcador His, el marcador FLAG o el marcador myc.

En una realización preferida adicional del procedimiento de la invención, el ensayo es un inmunoensayo.

En otra realización preferida del procedimiento de la invención, dicha muestra es sangre, suero, plasma, tejido fetal, saliva, orina, tejido mucoso, moco, tejido vaginal, tejido fetal obtenido de la vagina, piel, pelo, folículo piloso u otro tejido humano.

En una realización preferida adicional del procedimiento de la invención, dicha molécula de ácido nucleico de dicha muestra está fijada a un soporte sólido.

La fijación de la molécula de ácido nucleico a un soporte sólido permitirá un fácil manejo del ensayo de prueba y, además, al menos algunos soportes sólidos tales como chips, obleas de silicio o microplacas permiten el análisis

simultáneo de mayores números de muestras. Idealmente, el soporte sólido permite un ensayo automatizado que emplea, por ejemplo, dispositivos robotizados.

En una realización particularmente preferida del procedimiento de la invención, dicho soporte sólido es un chip, una oblea de silicio, una perla o una placa de microvaloración.

5 Además, la invención se refiere a la molécula de ácido nucleico como se describe anteriormente en la presente memoria para uso en el análisis de la presencia de o predisposición a hipolactasia de tipo adulto. La molécula de ácido nucleico permite simultáneamente el análisis de la ausencia de la afección o la predisposición a la afección, como se ha descrito con detalle anteriormente en la presente memoria.

10 Además, la invención se refiere a un kit que comprende la molécula de ácido nucleico como se describe anteriormente en la presente memoria, el cebador o par de cebadores como se describe anteriormente en la presente memoria, el vector como se describe anteriormente en la presente memoria y/o el anticuerpo, aptámero y/o fago como se describe anteriormente en la presente memoria en uno o más envases.

15 Se describe también en la presente memoria el uso de la molécula de ácido nucleico como se describe anteriormente en la presente memoria o del vector como se describe anteriormente en la presente memoria en terapia génica. Los enfoques de terapia génica se han discutido anteriormente en la presente memoria en relación con el vector de la invención, y se aplican igualmente aquí. Se describen también en la presente memoria fragmentos de las moléculas de ácido nucleico como se definen anteriormente en la presente memoria y como se exhiben, en particular, en las SEQ ID NO: 3 a 4 para emplear en enfoques de terapia génica. Dichos fragmentos comprenden el nucleótido en posición -13910 como se define en (c) anteriormente en la presente memoria (y también mostrado en la SEQ ID NO: 3) o la posición -22018 como se define en (d) anteriormente en la presente memoria (y también mostrado en la SEQ ID NO 4).
20 Preferiblemente, dichos fragmentos comprenden al menos 200, al menos 250, al menos 300, al menos 400 y lo más preferiblemente al menos 500 nucleótidos.

Se describe también en la presente memoria que dicha terapia génica trata o previene la hipolactasia de tipo adulto.

Las figuras muestran:

25 **Fig. 1:** Las familias finlandesas con hipolactasia de tipo adulto estudiadas. Los símbolos oscuros indican individuos hipolactásicos, el asterisco (*) indica que no estaba disponible muestra, la interrogación (?) indica estado de afección desconocido, ↑ indica los individuos usados para secuenciación en la identificación de SNP (Tabla 2).

30 **Fig. 2:** Mapa físico del locus de hipolactasia de tipo adulto. Los clones de BAC se muestran encima de la línea horizontal. Los tres genes *LPH*, *MCM6* y *DARS* se muestran mediante flechas negras gruesas con la punta hacia el extremo 3' del gen por encima de los recuadros negros. Se muestra la posición de 10 marcadores polimórficos de microsatélites usados para la cartografía fina del locus. La barra invertida en la línea horizontal denota un hueco en la secuencia de la secuencia de cóntigo. Se confirmó la posición del marcador D2S2169 rellenando el hueco con PAC 106020 aislado de la colección PAC como se describe anteriormente⁴⁰. Se muestra la organización del gen *MCM6*, incluyendo la posición de las variantes asociadas al fenotipo de lactasa persistente en los intrones 9 y 13, localizados a 13,9 kb y 22 kb en dirección 5' del primer ATG de *LPH*.
35

40 **Fig. 3:** Análisis haplotípico extendido de los cromosomas persistentes derivados de familias finlandesas con hipolactasia de tipo adulto usando 7 marcadores de microsatélite estrechamente ligados. Los haplotipos que representan el cromosoma persistente fundacional ancestral están sombreados. Se muestran solo los haplotipos de cromosomas no persistentes que estaban presentes también en los cromosomas persistentes. Basándose en las recombinaciones ancestrales, el locus de hipolactasia de tipo adulto pudo limitarse a un intervalo de 47 kb entre los marcadores LPH1 y AC3.

Fig. 4: La secuencia comprendida en la secuencia del intrón 13 del gen *MCM6* (3220 pb) que comprende el SNP en posición -13910 en que la T, que es específica de persistencia de lactasa, está sustituida por una C. Dicha posición está indicada por el uso de una minúscula. Esta secuencia designa la SEQ ID NO: 1.

45 **Fig. 5:** La secuencia comprendida en la secuencia del intrón 9 del gen *MCM6* (1295 pb) que comprende el SNP en posición -22018 en que la A, que es específica de la secuencia de tipo persistencia de lactasa, está sustituida por una G. Dicha posición está indicada por el uso de una minúscula. Esta secuencia se designa como SEQ ID NO: 2.

50 **Fig. 6:** La secuencia del intrón 13 del gen *MCM6* (3220 pb) de tipo persistencia de lactasa que comprende en la posición -13910 una T. Dicha posición está indicada por el uso de una minúscula. Esta secuencia se designa como SEQ ID NO: 3.

Fig. 7: La secuencia del intrón 9 del gen *MCM6* (1295 pb) de tipo persistencia de lactasa que comprende en la posición -22018 una A. Dicha posición está indicada por el uso de una minúscula. Esta secuencia se designa como SEQ ID NO: 4.

Fig. 8: La secuencia del intrón 13 del gen *MCM6* (3220 pb) que comprende el SNP en la posición -13910 en que la T, que es específica de la secuencia de tipo persistencia de lactasa, está sustituida por una C. Dicha posición está indicada por el uso de una minúscula. Esta secuencia se designa como SEQ ID NO: 5.

Fig. 9: La secuencia del intrón 9 del gen *MCM6* (1295 pb) que comprende el SNP en la posición -22018 en que la A, que es específica de la secuencia de tipo persistencia de lactasa, está sustituida por una G. Dicha posición está indicada por el uso de una minúscula. Esta secuencia se designa como SEQ ID NO: 6.

Los ejemplos ilustran la invención.

Ejemplo 1: Análisis de ligamiento y desequilibrio de ligamiento

Se analizaron 7 marcadores polimórficos de microsatélites entre D2S114 y D2S2385 que flanquean el gen *LPH* en 2q21 en 9 familias finlandesas extensas con hipolactasia (Fig. 1). Se encontró evidencia significativa de ligamiento con los marcadores D2S314, D2S442, D2S2196 y D2S1334, con una puntuación de LOD máxima de 7,67 a $\theta = 0$, obtenida con el marcador D2S2196 (Tabla 1). Se detectaron eventos de recombinación obligatorios con el marcador D2S114 (familia B, IV3), que define los límites centroméricos para el locus de persistencia/no persistencia de lactasa, y con el marcador D2S2385 (familia B, IV17) (Fig. 1, Tabla 1), que define los límites teloméricos del locus. Para la cartografía fina de la región crítica, se analizaron 9 marcadores polimórficos adicionales (Tabla 1). Se controló el desequilibrio de ligamento (DL) en la región condicionado al ligamiento detectado, tratando las frecuencias alélicas y la fracción de recombinación como parámetros de ruido¹⁶⁻¹⁷. 6 de los 9 marcadores (LPH13, LPH2, LPH1, AC3, AC4 y AC10), que se extienden en un intervalo de más de ~200 kb, mostraron una evidencia altamente significativa de DL ($p < 10^{-4}$), mientras que los marcadores en 3' del gen *LPH* no mostraron evidencia de DL (Tabla 1). Dos marcadores, LPH2 y AC3, expusieron el desequilibrio de ligamiento más significativo en los alelos de persistencia de lactasa ($p < 10^{-7}$).

El material familiar consistía en 9 árboles genealógicos finlandeses extensos estudiados originalmente por Sahi⁵. Se ensayó en todo el material familiar la hipolactasia de tipo adulto en los años 70. El material familiar de este estudio se aumentó recogiendo el ADN de miembros de la familia de las nuevas generaciones. El material familiar de este estudio consistía en 194 individuos en total (Fig. 1). Se confirmó el estado fenotípico de todos los miembros de la familia mediante ensayos de tolerancia a la lactosa con etanol (ETLE)⁴⁻⁵ en todos menos en 49 individuos. Se ha excluido la enteropatía por gluten de todos los pacientes afectados mediante la medida de la IgG anti-transglutaminasa sérica de tejido⁴⁵. Se extrajo ADN de muestras de sangre tomadas de todos los miembros de la familia participantes de acuerdo con protocolos estándares⁴⁶ después de obtener el consentimiento informado. Como estudio de casos y controles, se secuenciaron 196 muestras de ADN aleatorias aisladas de especímenes de biopsia de yeyuno en los cuales se habían medido las actividades disacaridasa en el Hospital universitario de Helsinki⁴⁷. Se aisló el ADN de biopsias intestinales según el protocolo estándar⁴⁸. Estas series comprendían 137 muestras con persistencia de lactasa y 59 sin persistencia. Además, se analizó el ADN de 9 italianos, amablemente proporcionado por M. Rossi, Universidad de Nápoles, 9 muestras de ADN de alemanes, amablemente proporcionadas por M. Lentze, Universidad de Bonn y 22 especímenes de muestra de biopsia intestinal surcoreanos, amablemente proporcionados por J.K. Seo, Universidad nacional de Seúl. (En la tabla: 23 coreanos, 9 italianos y 7 alemanes (uno de los casos de Alemania era originario de Corea del Sur)). El diagnóstico se basó en la medida de las actividades disacaridasa. Finalmente, para determinar la frecuencia de la variante C/T₋₁₃₉₁₀ en la población finlandesa, se analizó el ADN de 938 donantes anónimos de sangre finlandeses de pequeñas parroquias de Finlandia oriental y occidental y el ADN de 109 progenitores pertenecientes a las familias del CEPH¹⁹. Además, se analizó el ADN genómico de un babuino (*Papio hamadryas ursinus*) aislado de biopsia de hígado usando protocolos estándares⁴⁸ aprobados por los Comités éticos del Hospital de la Universidad de Helsinki y el Servicio de transfusión de sangre de la Cruz Roja finlandesa.

Ejemplo 2: Análisis haplotípico extendido

En la primera etapa, se analizaron 10 marcadores polimórficos de microsatélites que flanquean el gen *LPH* en 2q21 como se describe en otro lugar^{40,55}. Brevemente, se analizaron los 10 marcadores altamente polimórficos de microsatélites en 2q en la vecindad del gen de lactasa del Génethon Resource Center⁵⁵, con distancias genéticas como las siguientes: cen -D2S114 -1 cM -D2S1334 -0 cM -D2S2196 -0 cM -D2S442 -2 cM -D2S314 -2 cM -D2S2385 -1 cM -D2S2288 -1 cM -D2S8397 -1 cM -D2S150 -1 cM -D2S132. El orden de los marcadores se ha obtenido en su mayoría del mapa físico de cóntigos de YAC del cromosoma 2 (Chumakov *et al.*, 1995⁵⁶) suplementado con el mapa de Génethon. Se efectuó PCR en un volumen total de 15 μ l que contenía 12 ng del ADN molde, 5 pmol de cebadores, 0,2 mM de cada nucleótido, TrisHCl 20 mM (pH 8,8), $(\text{NH}_4)_2\text{SO}_4$ 15 mM, MgCl_2 1,5 mM, 0,1% de Tween 20, 0,01% de gelatina y 0,25 U de polimerasa Taq (Dynazyme, Finnzymes). Uno de los cebadores estaba radiomarcado en el extremo 3' con ³²P- γ ATP. Se efectuaron las reacciones en una placa de microvaloración multipocillo durante 35 ciclos con desnaturalización a 94°C durante 30 s, asociación a diversas temperaturas dependiendo de los cebadores durante 30 s y extensión a 72°C durante 30 s; se fijó la desnaturalización a 3 min y la extensión final a 5 min. Se separaron los fragmentos amplificados en gel de poliacrilamida al 6% y se efectuó una autorradiografía.

En la segunda etapa, se identificaron 9 marcadores de microsatélites adicionales en el cóntigo construido en el gen *LPH* a partir de la secuencia genómica publicada de los BAC (NH034L23, NH0318L13, NH0218L22 y RP11-329/1) usando el programa Repeat Masker (<http://ftp.genome.washington.edu/cgi-bin/RepeatMasker>). Se sintetizaron los cebadores que

flanquean las repeticiones. Las condiciones de PCR fueron como se describen en otro lugar⁴⁰. Se separaron los fragmentos amplificados en gel de poliacrilamida al 6% y se efectuó una autorradiografía.

Se calcularon las puntuaciones de LOD apareadas mediante el uso de la opción MLINK del paquete de programas LINKAGE⁴⁹. Se supuso una herencia autosómica recesiva para la hipolactasia de tipo adulto con penetrancia completa, sin diferencia de sexos en las fracciones de recombinación y una frecuencia alélica de la enfermedad de 0,4. Solo se incluyeron en el estudio los individuos de más de 20 años de edad, ya que la afección se manifiesta a esa edad en la población finlandesa⁵⁻⁶. El estado de afección de individuos no confirmados por ETLE se consideró desconocido. Se estimaron las frecuencias alélicas y heterocigosidades de los marcadores a partir del material familiar usando el programa Downfreq con fines de análisis de ligamiento paramétrico⁴⁹. Adicionalmente, se efectuaron análisis de ligamiento de seudomarcadores y desequilibrio de ligamiento, suponiendo un modo de herencia autosómico recesivo¹⁶. Se efectuó un ensayo de DL condicionado al ligamiento detectado, tratando las frecuencias alélicas y la fracción de recombinación como parámetros de ruido^{16,49}. Los valores de P de estos análisis se muestran en la Tabla 1. Se construyeron manualmente los haplotipos para los marcadores de microsatélites en este orden: LPH1-LPH2-LPH13-AC7-AC3-AC4-AC5 (Fig. 3). Estaban disponibles para análisis haplotípico un total de 54 cromosomas no persistentes y 33 cromosomas persistentes en el material familiar.

Se confirmó el orden de los marcadores estrechamente relacionados ensamblando los cuatro clones de BAC NH0034L23, NH0218L22, NH0318L13 y 329/10 en la región crítica en un segmento de secuencia ininterrumpida. Este cóntigo se extendía del marcador AC8 al exón 10 del gen de aspartil-ARNt sintetasa (DARS) y cubría un total de 222,5 kb (Fig. 2). Basándose en este mapa físico de la región ligada, se construyeron haplotipos extendidos con 7 marcadores que cubrían un intervalo de 150 kb (cen-LPH13-LPH2-LPH1-AC7-AC3-AC4-AC5-tel) (Fig. 3). Estaba presente un haplotipo mayoritario en 20 alelos persistentes (60%), frente a 3 de los alelos no persistentes (5%), mientras que se observó una amplia diversidad de haplotipos en los alelos no persistentes. El 40% restante de los haplotipos en los alelos persistentes difería del haplotipo ancestral de manera consistente con una degradación del haplotipo por eventos de recombinación históricos. Basándose en el análisis haplotípico conservado, el locus de la persistencia de lactasa pudo limitarse a un intervalo de 47 kb entre los marcadores LPH1 y AC3 (Fig.3)

Ejemplo 3: Análisis de secuencia del locus de hipolactasia de tipo adulto

Se amplificó la región de 47 kb entre los marcadores LPH1 y AC3 en fragmentos de PCR superpuestos a partir del ADN genómico de varios miembros de las 9 familias con hipolactasia y se secuenció. La región contiene el gen de mantenimiento del minicromosoma (*MCM6*)¹⁸, que cubre 36 kb de la región crítica de 47 kb (Fig. 2). No se detectaron variaciones en la región de codificación del gen *MCM6*, pero se identificaron un total de 52 variantes: 43 SNP y 9 polimorfismos de delección/inserción en la región crítica de 47 kb (Tabla 2). Solo dos de las variantes (C/T₋₁₃₉₁₀, G/A₋₂₂₀₁₈) estaban asociadas al rasgo de persistencia/no persistencia de lactasa en las familias finlandesas (Tablas 2 y 3). La primera variante asociada, C/T₋₁₃₉₁₀, reside en el intrón 13 del gen *MCM6* gene en la posición -13910 pb desde el primer codón ATG del gen *LPH*. La segunda variante asociada, G/A₋₂₂₀₁₈, está localizada en el intrón 9 del gen *MCM6* en la posición -22018 desde el primer codón ATG del gen *LPH* (Fig. 2). Estas dos variantes, separadas 8 kb entre sí, se cosegregaban completamente con hipolactasia de tipo adulto en 9 familias finlandesas extensas. Todos los miembros hipolactásicos (no persistentes) de la familia eran homocigóticos tanto de C₋₁₃₉₁₀ como de G₋₂₂₀₁₈ (Tabla 3). De forma interesante, ambas variantes residen en elementos repetidos, C/T₋₁₃₉₁₀ en un elemento derivado de L2 y G/A₋₂₂₀₁₈ en un elemento derivado de Alu.

Experimentalmente, se usaron 3 individuos sin persistencia, 2 con persistencia homocigótica y 2 con persistencia heterocigótica que comparten un haplotipo similar en la región crítica del material familiar para secuenciación en la primera etapa (Fig. 1). Usando la secuencia genómica preliminar publicada, se ensamblaron en un cóntigo los BAC NH0034L23, NH0218L22, NH0318L23 y RP-329/10, que cubrían la región crítica de hipolactasia de tipo adulto, usando el software Sequencher 4 (Gene Codes Corporation). Se diseñaron cebadores oligonucleotídicos que se extienden por la región crítica entre los marcadores LPH1 y AC3 (se describe a continuación en la presente memoria una lista de cebadores oligonucleotídicos). Se llevaron a cabo amplificaciones por PCR en un volumen de 50 µl con ADN genómico (100 ng), cebadores (20 ng cada uno), dNTP (200 µM), 0,5 U de polimerasa Taq (Dynazyme, Finnzymes) en un tampón estándar. Se amplificaron la mayoría de PCR usando las condiciones de ciclo de PCR siguientes: una ronda inicial de desnaturalización a 94°C durante 3 min, y entonces 35 ciclos a 94°C durante 30 s, 55°C durante 30 s y 72°C durante 1,25 min y una extensión final a 72°C durante 10 min, excepto porque en los casos en que el tamaño de los productos de PCR era mayor de 1 kb se usó el kit de extensión Dynazyme (las condiciones se describen a continuación en la presente memoria). Se secuenciaron cíclicamente los productos de PCR (15-40 ng) usando la química de terminación BigDye (PE Biosystems). Se analizaron los datos usando ABI Sequencing Analysis 3.3 (PE Biosystems) y Sequencher4.1 (Gene Codes).

Detección de variantes de lactasa mediante secuenciación:

Se llevaron a cabo amplificaciones por PCR en un volumen de 50 µl con ADN genómico (100 ng), cebadores (20 ng cada uno), dNTP (200 µM), 0,5 U de polimerasa Taq (Dynazyme, Finnzymes) en un tampón estándar. Se amplificaron ambas PCR usando las condiciones de ciclo de PCR siguientes: una ronda inicial de desnaturalización a 94°C durante 3 min, y entonces 35 ciclos a 94°C durante 30 s, 55°C durante 30 s y 72°C durante 1,25 min y una extensión final a 72°C durante 10 min. Se purificó la PCR mediante reacción enzimática. Se secuenciaron cíclicamente los productos de PCR

purificados (15-40 ng) usando la química de terminación BigDye (PE Biosystems). Se analizaron los datos usando ABI Sequencing Analysis 3.3 (PE Biosystems) y Sequencher 4.1 (Gene Codes).

Cribado de las variantes de lactasa mediante minisecuencia en fase sólida:

Se amplificó el fragmento de ADN que se extiende por la variante C/T₋₁₃₉₁₀ usando un cebador biotinilado (5'-Bio-CCTCGTTAATACCCACTGACCTA-3') y un cebador no biotinilado (5'-GTCACCTTTGATATGATGAGAGCA-3'). Para G/A₋₂₂₀₁₈, se usaron un cebador biotinilado (5'-Bio-TGCTCAGGACATGCTGATCAA-3') y otro no biotinilado (5'-CTACCCTATCAGTAAAGGCCTA-3') en las condiciones descritas anteriormente. Se capturaron 10 µg del producto de PCR en un pocillo de microvaloración recubierto con estreptavidina (Lab Systems, Finlandia), se lavaron los pocillos y se desnaturizó el ADN unido como se describe por Syvänen *et al.*, (*Am. J. Hum. Genet.* (1993), 52, 46-59) y Syvänen y Landegren (*Hum. Mutat.* (1994), 3, 172-9). Se añadieron a cada pocillo 50 µl de la mezcla de reacción de minisecuencia que contenían 10 pmoles de los cebadores de minisecuencia de C/T₋₁₃₉₁₀ (5'-GGCAATACAGATAAGATAATGTAG-3'), G/A₋₂₂₀₁₈ (5'-AAAAACAGCATTCTCAGCTGGGC-3'), y 0,1 µl de cualquiera de H-dCTP y H-dGTP correspondientes al alelo de no persistencia de lactasa (115 Ci/mmol; Amersham, RU) o H-dTTP y H-sATP correspondientes al alelo de persistencia de lactasa, y 0,05 unidades de ADN polimerasa (Dynazyme II, Finnzymes) en su tampón. Se incubaron las placas de microvaloración durante 20 min a 50°C y se lavaron los pocillos. Se eluyó la detección y se midió la radiactividad eluida en un contador de centelleo líquido (Rackbeta 1209, Wallac, Finlandia). Se llevaron a cabo dos reacciones de minisecuencia paralelas para cada producto de PCR.

Cebadores de PCR y cebador de detección de la variante C/T₋₁₃₉₁₀:

Cebador de PCR de codificación:	GTCACCTTTGATATGATGAGAGCA	Tm 58	SEQ ID NO: 8
Cebador de detección:	GGCAATACAGATAAGATAATGTAG	Tm 58	SEQ ID NO: 10
Cebador inverso Bio:	Bio-CCTCGTTAATACCCACTGACCTA	Tm 62	SEQ ID NO: 9
o	Bio-TAGGTCAGTGGGTATTAACGAGGT		SEQ ID NO: 7

Cebadores de PCR y cebador de detección para la variante G/A₋₂₂₀₁₈:

Cebador de PCR de codificación:	CTACCCTATCAGTAAAGGCCTA	Tm 58	SEQ ID NO: 12
Cebador de detección:	AAAAACAGCATTCTCAGCTGGGC	Tm 62	SEQ ID NO: 14
Cebador inverso Bio:	Bio-TGCTCAGGACATGCTGATCAA	Tm 62	SEQ ID NO: 13
o	Bio-TTGATCAGCATGTCCTGAGCA		SEQ ID NO: 11

20 Ejemplo 4: Control de las variantes de ADN en una muestra de estudio de casos y controles

Se analizó la frecuencia de las variantes C/T₋₁₃₉₁₀ y G/A₋₂₂₀₁₈ en muestras de ADN aisladas de un total de 196 especímenes de muestra de biopsia intestinal en que se había analizado la actividad disacaridasa como ensayo de diagnóstico para hipolactasia. Un total de 59 muestras mostraron deficiencia de lactasa primaria. 6 de los 59 casos (Tabla 3) eran GA heterocigóticos de la variante G/A₋₂₂₀₁₈, siendo los 53 restantes homocigóticos del alelo G. Las 59 muestras eran homocigóticas del alelo C de la variante C/T₋₁₃₉₁₀. Entre los 137 casos que muestran persistencia de lactasa, se encontró que 74 eran homocigóticos de los alelos T y A, siendo 63 heterocigóticos de CT y GA y no siendo ninguno homocigótico de los alelos C y G en C/T₋₁₃₉₁₀ y G/A₋₂₂₀₁₈, respectivamente (Tabla 3).

Para analizar estas variantes en otras poblaciones, se secuenciaron muestras de ADN aisladas de especímenes de biopsia intestinal de 40 casos no finlandeses con deficiencia de disacaridasa establecida: 23 casos originarios de Corea del Sur, 9 de Italia y 8 de Alemania. Un caso italiano era heterocigótico GA para G/A₋₂₂₀₁₈, mientras que todos los 39 casos restantes eran homocigóticos CC y GG para C/T₋₁₃₉₁₀ y G/A₋₂₂₀₁₈ respectivamente (Tabla 3). Un estudio extenso dio lugar a los datos proporcionados en la Tabla 7, que representa datos de la asociación completa de la variante C/T₋₁₃₉₁₀ con la hipolactasia verificada bioquímicamente (no persistencia de lactasa) en 400 individuos de 6 poblaciones diferentes. La variante G/A₋₂₂₀₁₈ se asociaba a no persistencia de lactasa en 400 de los 401 casos.

35 Ejemplo 5: Epidemiología molecular de la variante de persistencia de lactasa C/T₋₁₃₉₁₀

Para controlar la prevalencia de la variante asociada a hipolactasia en la población finlandesa, se usó un procedimiento de minisecuencia en fase sólida^{19,20} para cribar las muestras de ADN de 938 donantes de sangre anónimos finlandeses originarios de la región occidental de asentamiento temprano o de la región oriental de asentamiento tardío de Finlandia (Tabla 4). Experimentalmente, se amplificó el fragmento de ADN que se extiende por la variante C/T₋₁₃₉₁₀ usando un cebador biotinilado (5'-CCTCGTTAATACCCCTGACCTA-3') y otro no biotinilado (5'-GTCACCTTTGATATGATGAGAGCA-3'). Para G/A₋₂₂₀₁₈, se usó un cebador biotinilado (5'-AGTCTGTGGCATGTGTCTTCATG-3') y otro no biotinilado (5'-TGCTCAGGACATGCTGATCAACT-3') en las condiciones descritas anteriormente. Se capturaron 10 µl del producto de PCR en un pocillo de microvaloración recubierto con estreptavidina (Lab System, Finlandia). Se lavaron los pocillos y se desnaturizó el ADN unido como se

describe anteriormente^{19,20}. Se añadieron a cada pocillo 50 µl de la mezcla de reacción de minisecueñación que contienen 10 pmoles de los cebadores de minisecueñación para G/A-₂₂₀₀₅ (5'-GACAAAGGTGTGAGCCACCG-3'), G/A-₁₃₉₁₅ (5'-GGCAATACAGATAAGATAATGTAG-3') y 0,1 µl de cualquier H-dCTP correspondiente al alelo de no persistencia de lactasa (115 Ci/mmol; Amersham, RU) o H-dTTP correspondiente al alelo de persistencia de lactasa y 0,05 unidades de ADN polimerasa (Dynazyme II, Finnzymes) en su tampón. Se incubaron las placas de microvaloración durante 20 min a 50°C y se lavaron los pocillos. Se eluyó el cebador de detección y se midió la radiactividad eluida en un contador de centelleo líquido (Rackbeta 1209, Wallac, Finlandia). Se llevaron a cabo dos reacciones de minisecueñación paralelas para cada producto de PCR. La prevalencia global del presunto genotipo de hipolactasia CC-₁₃₉₁₀ (170 casos) era de 18,1%, con mayor prevalencia (16,8% frente a 18,9%) en la muestra occidental que en la oriental (Tabla 4). Estos valores están bien de acuerdo con el estudio epidemiológico que reseña una prevalencia del 17% entre los finlandeses que hablan finés, con un gradiente creciente de oeste a este². Se genotipó también el mismo conjunto de muestras para el polimorfismo G/A-₂₂₀₁₈ y se controló el DL entre estos dos SNP usando el estadístico D'²¹. Se encontró un DL casi completo (D' = 0,98, $p = 7,62 \times 10^{-11}$, Tabla 5).

Es conocido que la prevalencia de la hipolactasia en diferentes poblaciones varía en gran medida desde menos de 5% a casi 100%^{3,6}. Para determinar si estos cambios en la prevalencia de hipolactasia se correlacionaban con la distribución del genotipo CC-₁₃₉₁₀, se analizó el ADN de los progenitores de familias del CEPH²². Las familias del CEPH se han recogido principalmente en Francia, con una prevalencia reseñada de hipolactasia de aproximadamente 37%²³ y en Utah, siendo las poblaciones de Utah originarias del norte de Europa, con una prevalencia de hipolactasia de menos de 5%²⁴. El genotipado de los progenitores en familias del CEPH reveló que un 41,2% (7 de 17 muestras) de las familias francesas tienen un genotipo CC, mientras que solo un 7,6% (7 de 92 muestras) de las familias de Utah tienen un genotipo CC (Tabla 4). De nuevo, a pesar del pequeño número de muestras analizadas, estos números concuerdan con los valores obtenidos en los estudios epidemiológicos de hipolactasia en estas poblaciones^{23,24}. La Tabla 8 demuestra que la prevalencia observada de las variantes está bien de acuerdo con las frecuencias de intolerancia a la lactosa de la población descrita.

Ejemplo 6: Genealogía de la variante de persistencia de lactasa C/T-₁₃₉₁₀

El análisis haplotípico de las familias finlandesas sugería que la mayoría, sino todos, los alelos de persistencia de lactasa en Finlandia descendían de un antecesor común. Se usó el desequilibrio de ligamiento para estimar el momento de introducción del alelo de persistencia en la población finlandesa²⁵. Suponiendo un tiempo por generación de 20 años, esta estimación indicaría que la mutación fundacional se introdujo en la población finlandesa hace unos 9000-11400 años (Tabla 6). Esto está bien de acuerdo con las señales más tempranas de asentamiento en el territorio finlandés hace unos 8000-9000 años²⁶, y coincidiría razonablemente bien con el inicio de la industria láctea en los años 8000-10000 a de C²⁷. De forma más importante, la presencia de la misma variante de ADN en los alelos de persistencia en diferentes poblaciones sugeriría que esta variante es aún más antigua y que la mutación ocurrió antes de la diferenciación de las poblaciones analizadas.

Para llegar a comprender mejor el origen filogenético del alelo de lactasa, se secuenciaron el intrón 9 y parte del intrón 13 del gen *MCM6* de un babuino (*Papio hamadryas*). El genotipo GG y CC estaba presente en el ADN de babuinos tanto en G/A-₂₂₀₁₈ como en C/T-₁₃₉₁₀. Esto podría sugerir que los alelos G y C reflejan, respectivamente, la aparición del alelo ancestral, que presenta el tipo de no persistencia, y una mutación que ha transformado este alelo creando el alelo de persistencia. Esta suposición está apoyada por la identificación de DL y el haplotipo compartido en los alelos de persistencia frente a la alta diversidad de los alelos encontrados en los alelos de no persistencia.

Ejemplo 7: DL apareado de las variantes C/T y G/A

Se estimó el DL entre C/T-₁₃₉₁₀ y G/A-₂₂₀₁₈ usando el estadístico D'²¹. Se estimaron las frecuencias haplotípicas mediante la probabilidad máxima usando el programa EH⁵⁰. La D' se calcula como $\max(D/D_{\max}, D/D_{\min})$, en que la medida del desequilibrio $D = h_{pq} - pq$, en que h_{pq} es la frecuencia del haplotipo con el alelo raro en cada locus, p y q son la frecuencia de los alelos raros en los loci 1 y 2 y $D_{\max} = \min p(1-p), q(1-q)$ si $D > 0$, y $D_{\min} = -\min pq, (1-p)(1-q)$ si $D < 0$. Se determinó la significación de la desviación de D' de 0 usando el estadístico D²

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

que se distribuye como χ^2 con 1 df²¹.

Números de acceso a genes. Para los BAC NH0218L22, N0034L34, NH0318L13 y RP11-329/10 son AC012551, AC011893, AC011999 y AC016516 respectivamente. Los números de acceso para polimorfismos humanos son GenBank AF395607-AF395615.

Referencias

1. Flatz, G. y Rotthauwe, H. "The human lactase polymorphism: physiology and genetics of lactose absorption and malabsorption". *Prog. Med. Genet.* 2, 205-249 (1977).

2. Sahi, T., Isokoski, M., Jussila, J. y 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).
- 5 4. Sahi, T., Isokoski, M., Jussila, J., Launiala, K. y Pyorala, 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).
- 10 6. Sahi, T. "Genetics and epidemiology of adult-type hypolactasia". Scand. J. Gastroenterol. Suppl. 202, 7-20 (1994).
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).
- 15 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).
10. Harvey, C.B., Pratt, W.S., Islam, I., Whitehouse, D.B. y Swallow, D.M. "DNA polymorphisms in the lactase gene: linkage disequilibrium across the 70 kb region". Eur. J. Hum. Genet. 3, 27-41 (1995).
11. Escher, J.C *et al.* "Molecular basis of lactase levels in adult humans". J. Clin. Invest. 89, 480-483 (1992).
- 20 12. Lloyd, M *et al.* "Regulation of intestinal lactase in adult hypolactasia". J. Clin. Invest. 89, 524-529 (1992).
13. Fajardo, O., Naim, H.Y. y Lacey, S.W. "The polymorphic expression of lactase in adults is regulated at the messenger RNA level". Gastroenterology 106, 1233.
14. Luigi, M. *et al.* "Mosaic regulation of lactase in human adult-type", Gastroenterology 112, 1506-1514 (1997).
- 25 15. Rossi, 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. y 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).
- 30 17. Terwilliger, J.D. y 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).
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).
- 35 20. Syvänen, A-C. y Landegren, U. "Detection of point mutations by solid-phase methods". Hum. Mutat. 3, 172-179 (1994).
21. Thompson, E. A., Deeb, S., Walker, D. y 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).
- 40 22. Dausset, J. *et al.* "Centre d'etude du polymorphisme humain (CEPH): Collaborative genetic mapping of human genome". Genomics 6, 575-577 (1990).
23. Cuddenec, Y., Delbrück, H. y 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. McLellan, T., Jorde, L.B. y Skolnick, M.H. "Genetic distance between the Utah Mormons and related populations". Am. J. Hum. Genet. 36, 836-857 (1984).
- 45 25. 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. Núñez, 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).
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).
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).
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. y Moran, J.V. "The impact of L1 retrotransposons on the human genome". Nature Genet. 19, 19-24 (1998).
34. Moran, J.V., DeBerardinis, R.J. y Kazazian, H.H. "Exon shuffling by L1 retrotransposition". Science 283, 1530-1534 (1999).
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. y Moss, S.E. "A conserved nuclear element with a role in mammalian gene regulation". Hum. Mol. Genet. vol. 8, 9, 1723-1728 (1999).
37. Boeke, J.D. "LINEs and Alus - the polyA connection". Nature Genet. 16, 6-7 (1997).
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).
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).
41. Simoons, F. J. "The geographic hypothesis and lactose malabsorption. A weighing of the evidence". Am. J. Dig. Dis. 23, 963-980 (1978).
42. Flatz, G. y 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).
45. Sulkanen, S. *et al.* "Tissue transglutaminase autoantibody enzyme-linked immunosorbent assay in detecting celiac disease". Gastroenterology 115 (6), 1322-1328 (1998).
46. Sambrook, J., Fritsch, E.F. y Maniatis, T. "Molecular cloning: a laboratory manual" (2^a ed). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989).
47. Messer, M. y Dahlqvist, A. "A one-step ultramicro method for the assay of intestinal disaccharidases". Anal. Biochem. 14 (3), 376-92 (1966).
48. Cottingham, Jr. R.W., Idury, R.M. y Schaffer, A.A. "Faster sequential genetic linkage computations". Am. J. Hum. Genet. 53, 252-263 (1993).
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).

51. Osborne *et al.*, *Curr. Opin. Chem. Biol.* 1 (1997); 5-9
52. Stall y Szoka, *Pharm. Res.* 12 (1995), 465-483
53. Harlow y 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.
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*. 14 de marzo de 1996; 380 (6570): 152-4.
56. Chumakov IM, Rigault P, Le Gall 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*. 28 de septiembre de 1995; 377(supl. 6547): 175-297
- 10 **Tabla 1. Análisis de ligamiento y desequilibrio de ligamiento en familias con hipolactasia de tipo adulto (los marcadores de cartografía fina se muestran en negrita)**

Marcador	puntuación de LOD a Θ					valor de P ^a
	0,0	0,1	0,2	0,3	0,4	
D2S114	$-\infty$	2,44	1,92	1,13	0,41	0,87195
P6112	2,76	2,20	1,45	0,75	0,22	0,66207
D2S1334	3,15	2,45	1,61	0,84	0,25	0,91039
AC8	2,26	1,99	1,36	0,71	0,21	0,53670
LPH13	3,67	2,94	1,96	1,03	0,31	4×10^{-6}
LPH2	4,09	3,07	2,00	1,00	0,26	$5,7 \times 10^{-7}$
LPH1	5,91	4,52	2,96	1,53	0,46	5×10^{-6}
AC7	3,63	2,60	1,66	0,83	0,23	0,03471
AC3	6,63	4,88	3,16	1,61	0,44	$3,2 \times 10^{-8}$
AC4	3,07	2,22	1,42	0,71	0,19	4×10^{-5}
AC5	5,33	4,10	2,72	1,39	0,39	0,02166
AC10	6,60	4,99	3,25	1,65	0,46	1×10^{-5}
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	0,26	0,46457

a: valores de p producidos usando el ensayo de desequilibrio de ligamiento dado el ligamiento ^{16,49}

Tabla 2. Variaciones identificadas en el locus de hipolactasia de tipo adulto en familias finlandesas

Posición ^a	Variante	Persistencia de lactasa		Persistencia de lactasa		No persistencia de lactasa		
		(homocigótico)		(heterocigótico)				
		BIV4	AIV3	BIV8	CIV3	BIV9	DIV4	EIII2 ^b
-694	A→G	AA	AA	AG	AA	GG	N ^c	AA
-1640/50	T ₁₃ →T ₁₂	T _{13/13}	T _{13/13}	T _{13/13}	T _{13/13}	T _{13/13}	T _{12/12}	T _{12/12}
-2131	C→T	CC	CC	CT	CC	TT	CT*	TT
-3058/72	T ₁₅ →T ₁₆	T _{15/15}	T _{15/15}	T _{15/15}	T _{15/15}	T _{15/15}	T _{16/16}	T _{16/16}
-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 ₁₁ →T ₁₂	T _{11/11}	T _{11/11}	T _{11/12}	T _{11/11}	T _{12/12}	T _{11/11}	T _{12/12}
-16888	A→G	AA	AA	GA	AA	GG	GA	GG
-17300	C→T	CC	CC	CC	CC	CC	CT	TT

-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 ₇ →A ₆	A _{7/7}	A _{7/7}	A _{7/7}	A _{7/7}	A _{7/7}	A ₇ /A ₆	A _{7/7}
-21731	A→C	AA	AA	AA	AA	AA	CC	AA
-21736/43	A ₉ →A ₈	A _{9/9}	A _{9/9}	A ₉ /A ₈	A _{9/9}	A _{8/8}	A _{8/8}	A _{8/8}
-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	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
-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	Δ I	ΔΔ	II	ΔI	N
-36671/77	T7→T6	T _{7/7}	T _{7/7}	T _{7/6}	T _{7/7}	T _{6/6}	T _{7/7}	T _{7/7}
-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: El número es desde el codón de iniciación de la traducción (ATG) del gen de LPH usando la secuencia génica compilada de los BAC NH034L23, NH0218L22, NH0318L13 y RP11-329110, b: individuos secuenciados de familias finlandesas estudiados y mostrados por una flecha en la fig. 1, c: no determinado

Tabla 3. Distribución de los genotipos C/T₋₁₃₉₁₀ y G/A₋₂₂₀₁₈ en alelos persistente/no persistente de lactasa

		C/T ₋₁₃₉₁₀			G/A ₋₂₂₀₁₈			Total
Genotipo		CC	CT	TT	GG	GA	AA	
Miembros de familia	No persistencia de lactasa	45	0	0	45	0	0	45
	Persistencia de lactasa	0	32	13	0	32	13	45
Muestras de casos y controles								
Finlandeses	No persistencia de lactasa	59	0	0	53	6	0	59
	Persistencia de lactasa	0	63	74	0	63	74	137
No finlandeses*	No persistencia de lactasa	40	0	0	39	1	0	40
	Persistencia de lactasa	0	5	0	0	5	0	5
Total	No persistencia de lactasa						0	144

*Las muestras no finlandesas consisten en 23 individuos surcoreanos, 9 italianos y 7 alemanes

Tabla 4. Prevalencia de la variante C/T₋₁₃₉₁₀ en las muestras de población

Muestras de ADN	Genotipo			Total	Frecuencia alélica (%)		% de genotipo (CC)
	CC	CT	TT		C	T	
I. Población finlandesa:							
1. Regiones orientales	108	287	176	571	0,440	0,560	18,9%
2. Regiones occidentales	62	159	146	367	0,385	0,615	16,8%
Total	170	446	322	938	0,418	0,582	18,1%
II. Progenitores del CEPH:							
1. Familias de Utah	7	33	52	92	0,255	0,745	7,6%
2. Familias francesas	7	9	1	17	0,676	0,324	41,2%

Un total de 938 muestras de ADN de donantes de sangre anónimos finlandeses de pequeñas parroquias de regiones orientales y occidentales de Finlandia y 109 muestras de ADN de progenitores del CEPH. La prevalencia de la hipolactasia en las poblaciones se refleja por las frecuencias genotípicas de los alelos CC

Tabla 5. DL entre las variantes C/T₋₁₃₉₁₀ y G/A₋₂₂₀₀₁₈ en muestras aleatorias finlandesas

	Genotipo en C/T ₋₁₃₉₁₀			Total	D'	χ^2 (1 df)	Valor de P
	CC	CT	TT				
Genotipo en G/A ₋₂₂₀₁₈							
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 x 10 ⁻¹¹

Se calculó el DL usando el estadístico D'¹⁸, el valor p es la significación de D' desde 0 como se describe en los procedimientos¹⁸.

Tabla 6. Estimación de la introducción de la variante C/T₋₁₃₉₁₀ en la población finlandesa usando el programa DISLAMB

Marcador		AC3			LPH2		
Alelo	Persistencia de lactasa	No persistencia de lactasa			Persistencia de lactasa	No persistencia de lactasa	
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	
λ^a	0,838			0,999			
θ^b	0,00031 (0,000038-0,00099)			0,0000 (0,00000-0,00052)			
n^c	570			450			

a: λ es la proporción de aumento de un cierto alelo en cromosomas enfermos (alelo de persistencia de lactasa) respecto a su frecuencia de población (0,60). b: θ es la fracción de recombinación, reflejada por la distancia de la mutación desde el marcador más cercano, suponiendo 1 cM= 1 Mb. c: n es el número de generación desde la introducción de la mutación fundacional en una población aplicando la fórmula $\lambda = \infty (1-\theta)^n$. d: alelo hipotético usado en los cálculos cuando θ es cero y ∞ es 1.

Tabla 7 Prevalencia de las variantes de intolerancia a lactosa en muestras verificadas bioquímicamente

Población		Número	C/T ₁₃₉₁₀			G/A ₂₂₀₁₈		
			CC	CT	TT	GG	GA	AA
1. Finlandesa	Persistencia de lactasa	182	0	95	87	0	95	97
	No persistencia de lactasa	116	116	0	0	110	6	0
2. Italiana	Persistencia de lactasa	7	0	7	0	0	7	0
	No persistencia de lactasa	23	23	0	0	22	1	0
3. Alemana	Persistencia de lactasa	0	0	0	0	0	0	0
	No persistencia de lactasa	8	8	0	0	8	0	0
4. Somalí	Persistencia de lactasa	0	0	0	0	0	0	0
	No persistencia de lactasa	42	42	0	0	42	0	0
5. Surcoreana	Persistencia de lactasa	0	0	0	0	0	0	0
	No persistencia de lactasa	23	23	0	0	23	0	0
Total		401	212	102	87	205	109	87

Tabla 8. Prevalencia de las variantes de intolerancia a la lactosa en diversas muestras de población

Población	Genotipo							% de prevalencia del alelo de persistencia de lactasa
	Número	C/T13910			G/A22018			
		CC	CT	TT	GG	GA	AA	
Surcoreanos	23	23	0	0	23	0	0	0*
Francia	17	7	9	1	6	10	1	59*
Vascos	85	7	44	34	13	35	37	92*
Italianos meridionales	100	89	11	0	88	12	0	11*
Somalíes	79	74	5	0	78	1	0	6
Utah	92	7	33	52	7	30	55	92*
Afroamericanos	96	76	15	5	78	12	5	21*
Marroquíes	90	62	25	3	65	22	3	31*
Saharauis (africanos)	57	29	26	2	28	26	3	49*
Lapones	30	20	10	0	21	9	0	33*
Tíbet	23	23	0	0	0	3	0	0
Finlandeses orientales	571	108	287	176	107	288	176	81*

Finlandeses occidentales	367	62	159	146	58	161	148	83*
Tribus finougrias								
Xan	20	19	1	0	19	1	0	5
Xm	20	19	1	0	19	1	0	5
Mansis	22	20	2	0	20	2	0	9
Lkomis	10	7	3	0	7	3	0	30
Erzas	30	17	10	3	19	9	2	43
Moksas	30	13	17	0	14	16	0	57*
Udmurtos	30	12	16	2	11	15	4	60*
Tribus paquistaníes								
Kalashas	30	30	0	0	28	2	0	0
Burushos	30	29	1	0	27	3	0	3
Hazaras	14	13	1	0	11	3	0	7
Cachemiros	20	15	5	0	14	6	0	25
Makranis baluches	29	19	10	0	19	8	1	34
Brahuis	30	17	10	3	16	11	3	43
Makranis (negroides)	29	16	10	3	16	10	3	45
Pastunes	29	12	16	1	13	14	2	59*
Indios	29	11	13	5	10	12	5	62*
Total	2032							

*La prevalencia del alelo de persistencia de lactasa está muy bien correlacionada con las prevalencias reseñadas para el alelo de persistencia de lactasa (Simoons Fj. "The geographic hypothesis and lactose malabsorption", Am. J. Dig. Dis. 1978, 23 (11): 963-80)

LISTADO DE SECUENCIAS

<110> National Public Health Institute

PELTONEN, Leena

ENATTAH, Nabil

5

JÄRVEL Ä, Irma

SAHI, Timo

SAVILAHTI, Erkki

TERWILLIGER, Joseph

<120> Identificación de una variante de ADN asociada a hipolactasia de tipo adulto

10

<130> F 2034 PCT

<150> EP 01 11 9377.8

<151> 10-08-2001

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

5 <150> US 60/315.955
 <151> 31-08-2001

<160> 14

10 <170> PatentIn versión 3.1

<210> 1
 <211> 180
 <212> ADN

15 <213> *Homo sapiens*

<400> 1
 accttttcatt caggaaaaat gtacttagac cctacaatgt actagtaggc ctctgcgctg 60
 gcaatacaga taagataatg tagcccctgg cctcaaagga actctcctcc ttaggttgca 120
 tttgtataat gtttgatfff tagattgttc tttgagccct gcattccacg aggataggtc 180

<210> 2
 20 <211> 180
 <212> ADN
 <213> *Homo sapiens*

<400> 2
 taagaacatt ttacactctt cagtataaag aagtcagaat acccctaccc tatcagtaaa 60
 ggcctataag ttaccattaa aaagatgtcc ttaaaaaagc cattctcagc tgggcgcggt 120
 ggctcacacc tttgtcccag tactttggga agccgaggtg ggtggatcac ctgagggtcag 180

25 <210> 3
 <211> 3213
 <212> ADN
 <213> *Homo sapiens*

30 <400> 3

atcagagtca	ctttgatatg	atgagagcag	agataaacag	at ttgttgca	tg tttttaat	60
ctttggatg	ggacatacta	gaattcactg	caaatacatt	tttatgtaac	tg ttgaatgc	120
tcatacgacc	atggaattct	tccctttaaa	gagcttggtg	agcatttgag	tg tagttggt	180
agacggagac	gatcacgtca	tagtttatag	agtgcataaa	gacgtaagtt	accatttaat	240
acctttcatt	caggaaaaat	gtacttagac	cctacaatgt	actagtaggc	ctctgcgctg	300
gcaatacaga	taagataatg	tagtccttgg	cctcaaagga	actctcctcc	ttaggttgca	360
tttgtataat	gtttgatttt	tagattgttc	tttgagccct	gcattccacg	aggataggtc	420
agtgggtatt	aacgaggtaa	aaggggagta	gtacgaaagg	gcattcaagc	gtcccatctt	480
cgcttcaacc	aaagcagccc	tgogttttcc	tagttttatt	aataggtttg	atgtaaggtc	540
gtctttgaaa	aggggggttg	gctttttttt	acagtgtgac	tgagggtataa	tttataaaaa	600
gggaaatgta	tggcatggtg	agttttttca	catacatcct	tg tgaatacc	cagctcaaga	660
tccaaaacat	ttccataatt	tcagaaagtt	ccaaaccctt	gcctcttttc	agtcttagcc	720
ctcttcccct	gaagtaacca	ctgttccgac	ttcaatcact	acttttatcc	cacagggtta	780
ttttttggct	ttttccact	aaattttcaa	attctttgat	atggtaacttt	actattgacg	840
aagtactttc	acactagggt	atttaatat	ctttgattca	cccaatat	agggaacacc	900
tgtaggggac	aaaaaatgaa	tgagagcccc	tgccttccat	tgctgcta	ctgggtggaa	960
cgagacatgt	atttaattaa	gcatgtaaaa	aatagagtgg	gtgatgaaat	aatctatata	1020
ctaaatcccc	atgacacaca	gtttacctat	gtaacaaacc	tgcatgtgta	cccccgaa	1080
taaaatataa	gttggaat	aaaaaaaaac	gagagggaga	atagagcatc	acaaccagag	1140
tgctgagatg	aattacttta	ttaccaaaaga	aggaggagga	ctcagggagg	tgccgacgtt	1200
taaaccagct	cactgaaggg	tgtgcagaat	ttggataggc	aagataccct	gggacaaggt	1260
cattctaaaa	ccatgctaac	at ttgtactt	tttttttcat	tg t gatagtt	cctgaaatga	1320
gttgcataaa	actggtacat	gtcttagggc	agtctcta	tgatttttat	tttgttctat	1380
ttttaaaaat	tagtcttcaa	atagcagatt	cacatgat	taaaatat	gcacataaat	1440
tatatacaca	aatatat	ctgaatgaaa	tttagtat	gcatat	aagagctatt	1500
tctgtctcat	atgttcataa	tcttcatcca	ttaaaaaac	ttttgttagg	cctttctcac	1560
tctaagatta	taaaaaat	tccattatt	tacctagcta	gttttctagt	tg ttccaaaa	1620
ccatttattg	aacaatccat	ctttttgaca	ctgggttggc	atgcctta	tatatattct	1680
tgtgtgtgtt	aggatctcct	tttggacttt	ccattctgtt	cattgagct	tatcagctcc	1740
tcttacattg	gtaccatgat	gttttaatt	atggggcttt	gtagtttaaa	tgtagggcta	1800

gttcacgcgc attgttctct atcagctggt aggaacttag aaatcagctt gctctgtttt 1860
aaagaaaaac ctggtatttt tttatcagta taacattcta tttatattaa cttgaagaat 1920
tgaaaacatc tatgattttt cctattcagt aacgtatcac ttagaatagg ttaggttgta 1980
ctactataaa atctcagctg cataaaaacaa tttttttttg cttgtgctac acatccatta 2040
ggatcatcaag ggactcacct tgtcaagtta ctacagagatt caggctgata taaagggttg 2100
atcttgacat acgctttcat gatgacagaa agcagggaag agaagggtgt gagccatgtg 2160
ctttctcccc cttctatcca gaaatgacac atactcacat ttcattcgcc agagaaatta 2220
acatggcccc tcctaagttc aaatggatag agaaatgcct tcctaccagg tgcccagaat 2280
tagaagagca aacatttgtg aacagttctg agtaccacaa ataccgttat ctttccactt 2340
aagtcttctg tttcactcag tagtgcttta aacttttctt catatgtttt tcagtgtttc 2400
ttgttgaatt tcttgatatt ttatcatggt tgttcgtact gggagtagcc tttttttcca 2460
tttcattttc tggctgggtt cattgctggt tgtttttttg ttttgtttg ttttgagat 2520
ggagtctcac tctgtcgccc aggctggagt gcagtgtcac aatctcggt cactgcaacc 2580
tctgcctccc aggttcaagc gattcttctt tctcagcctc ctgagtagct gggattacag 2640
gcatgtgcca ccatgcccag ctaatttttt atatttttag tagagatggg gtttctccat 2700
gttggtcagg ctggtctcaa actcccaatc tcagggtgat cgctgcctc tgccttccaa 2760
agtgtggga ttatagacat gagccaccgt gcctggccta gttcttatgg gatgtatatg 2820
tctttggatt catatgatat gtatatatgt ttatatctt acaagtacat acctaggagt 2880
ggaattgttg ggtcataggt taatgcatgt ttttctgcca aacagttgtg tcaatttctg 2940
ttttcacccg tgtgaatgag agttgttcta ctttcttgac aacacttgat attgtcagtc 3000
attttagcca ttctggtgaa tttatagtgc tatttctgtg tgtgtaagag agagaatgag 3060
agaggggtgt tgtgagaaaa ccaaagcaac actgtgagag tgtgtgtgtt tgtgagaaaa 3120
ccaaaataca tactactgtg atttcattgg gagaaaatct gtttggtata tcaaaaaaag 3180
tagcttaatt acttcatcat tattgggtta ggt 3213

<210> 4

<211> 1296

<212> ADN

5 <213> *Homo sapiens*

<400> 4

taagaacatt ttacactctt cagtataaag aagtcagaat accctaccc tatcagtaaa 60
ggcctataag ttaccattaa aaagatgtcc ttaaaaaacag cattctcagc tgggcacggg 120

ggctcacacc tttgtcccag tactttggga agccgaggtg ggtggatcac ctgaggtcag 180
 gagttcgaga ccagcctggc caacatggcg aaaacccatt ttctctacta aaaatacaaa 240
 aattagccgg gcatgggtggc ggggtgcttgt ggtcccagct actcaagagg ctgaggtggg 300
 aggatcactg agcccaggag gtggaggctg cattgagcca agattgtgcc actgcactcc 360
 agcctgggtg acagagcgag actctgtctc aaaaaaacca aaacaaaaaa aaccagcat 420
 tcttttagtaa ataattcata gttttcttca tctagaattt aaaattgtga tagttgatca 480
 gcatgtcctg agcacgtgtg tttgctgtta ctagtttaga tcggtagatg tgtatataag 540
 ttataggtat aaaatcaatc ctgagttgac acaaggtttt gatgttgagt acaagtacag 600
 taagtgtata tttttagtta tgctcttagt ttttaagtcaa ttgtgtgggt ctttctagct 660
 ttaggatctg ttgaattatc ttcttagaa aagggagtta agaattctca cttacctatc 720
 ttctacttgt ttggagaata gaagagtccc tgtggtagca gactttgtga gtttacttgt 780
 aattttccat ctgaaagact gttcttggtt ttctgtatga agtcttgctc tgtcgcccag 840
 gctggagtgc agtggtgcaa ccttggtcca ctgcaacctc tgcctcccg gttcaagcaa 900
 ttctcctgcc tcagcctccc gagtatctgg gattacaggt gcacaccacc acacctggct 960
 aatttttgta ttttcagtag agacgggggt tcaccatgtt ggccaggctg gtctcgaact 1020
 cttgacctca tgatcagccc acctcagcct tccaaagtgc tgggattaca ggtgtgagcc 1080
 cccacactcg gcggttggtt ttttttaaga gacaggggtc cactotgtca cctaacctgg 1140
 agtacagtgg caatcatggc tcaactgtaac ctcaaagtgc cggccttagt gaagcgttct 1200
 tcctgccttg gcctcccaa gtgctgggat tacaagtgtg agccatgcat ccagcttgaa 1260
 agacagcttc ttaggcttga tttgtttggt tacagg 1296

<210> 5

<211> 3213

<212> ADN

5 <213> *Homo sapiens*

<400> 5

atcagagtca ctttgatatg atgagagcag agataaacag atttgttgca tgtttttaat 60
 ctttggtatg ggacatacta gaattcactg caaatacatt tttatgtaac tgttgaatgc 120
 tcatacgacc atggaattct tccctttaaa gagcttggtg agcatttgag tgtagttggt 180
 agacggagac gatcacgtca tagtttatag agtgcataaa gacgtaagtt accatttaat 240
 acctttcatt caggaaaaat gtacttagac cctacaatgt actagtaggc ctctgcgctg 300
 gcaatacaga taagataatg tagcccctgg cctcaaagga actctcctcc ttaggttgca 360

tttgtataat gtttgatttt tagattgttc tttgagccct gcattccacg aggataggtc 420
 agtgggtatt aacgaggtaa aaggggagta gtacgaaagg gcattcaagc gtcccatctt 480
 cgcttcaacc aaagcagccc tgcgttttcc tagttttatt aataggtttg atgtaaggtc 540
 gtctttgaaa aggggggttg gctttttttt acagtgtgac tgagggtataa tttataaaaa 600
 gggaaatgta tggcatgggtg agttttttca catacatcct tgtgaatacc cagctcaaga 660
 tccaaaacat ttccataatt tcagaaagtt ccaaaccctt gcctcttttc agtcttagcc 720
 ctcttcccct gaagtaacca ctgttccgac ttcaatcact acttttatcc cacagggtta 780
 ttttttggct tttttccact aaattttcaa attctttgat atggtaactt actattgacg 840
 aagtactttc aactagggtt atttaatat ctttgattca cccaatattt aggggaacacc 900
 tgtagggggac aaaaaatgaa tgagagcccc tgccttccat tgctgctaatt ctggtgggaa 960
 cgagacatgt atttaattaa gcatgtaaaa aatagagtgg gtgatgaaat aatctatata 1020
 ctaaattccc atgacacaca gtttacctat gtaacaaacc tgcattgtga ccccggaacc 1080
 taaaatataa gttggaaatt aaaaaaaaaac gagaggggaga atagagcatc acaaccagag 1140
 tgctgagatg aattacttta ttaccaaaga agggaggagga ctgaggaggg tgccgacgtt 1200
 taaaccaggt cactgaaggg tgtgcagaat ttggataggc aagataccct gggacaaggt 1260
 cattctaaaa ccatgctaac atttgtaactt tttttttcat tgtgatagtt cctgaaatga 1320
 gttgcataaa actggtacat gtcttagggc agtctctaatt tgatttttat tttgttctat 1380
 ttttaaaaat tagtcttcaa atagcagatt cacatgatat taaaatatat gcacataaat 1440
 tatatacaca aatatatttt ctgaatgaaa tttagtatct gcatatattt aagagctatt 1500
 tctgtctcat atgttcataa tcttcatcca ttaaaaaaac ttttgttagg cctttctcac 1560
 tctaagatta taaaaaattc tccattatt tacctagcta gttttctagt tgttccaaaa 1620
 ccatttattg aacaatccat ctttttgaca ctggtttggc atgccttaat tatatattct 1680
 tgtgtgtgtt aggatctcct tttggacttt ccattctgtt cattgagtct tatcagctcc 1740
 tcttacattg gtaccatgat gttttaatct atggggcttt gtagtttaaa tgtagggcta 1800
 gttccagcgc attgttctct atcagctgtt aggaacttag aaatcagctt gctctgtttt 1860
 aaagaaaaac ctggtatttt tttatcagta taacattcta tttatattaa cttgaagaat 1920
 tgaaaacatc tatgattttt cctattcagt aacgtatcac ttagaatagg ttaggttgta 1980
 ctactataaa atctcagctg cataaaacaa tttttttttg cttgtgctac acatccatta 2040
 ggtcatcaag ggactcacct tgtcaagtta ctgagagatt caggctgata taaagggttg 2100

atcttgacat acgctttcat gatgacagaa agcaggggaag agaaggtggt gagccatgtg 2160
ctttctcccc cttctatcca gaaatgacac atactcacat ttcattcgcc agagaaatta 2220
acatggcccc tctaagttc aaatggatag agaaatgcct tctaccagg tgcccagaat 2280
tagaagagca aacatttgtg aacagttctg agtaccacaa ataccgttat ctttccactt 2340
aagtcttctg tttcactcag tagtgcttta aacttttctt catatgtttt tcagtgtttc 2400
ttgttgaatt tcttgatatt ttatcatgtt tgttcgtact gggagtagcc tttttttcca 2460
tttcattttc tggctgggtt cattgctggt tgtttttttg ttttgttttg tttttgagat 2520
ggagtctcac tctgtcgccc aggctggagt gcagtgtcac aatctcggct cactgcaacc 2580
tctgcctccc aggttcaagc gattcttctt tctcagcctc ctgagtagct gggattacag 2640
gcatgtgcca ccatgccag ctaatttttt atatttttag tagagatggg gtttctccat 2700
gttggtcagg ctggtctcaa actcccaatc tcaggtgatc cgctgcctc tgccttccaa 2760
agtgtgga ttatagacat gagccaccgt gcctggccta gttcttatgg gatgtatatg 2820
tctttggatt catatgatat gtatatatgt ttatatctt acaagtacat acctaggagt 2880
ggaattgttg ggtcataggt taatgcatgt ttttctgcca aacagttgtg tcaatttctg 2940
ttttcaccgc tgtgaatgag agttgttcta cttcttgac aacacttgat attgtcagtc 3000
atttttagcca ttctggtgaa tttatagtgc tatttctgtg tgtgtaagag agagaatgag 3060
agaggggtgtt tgtgagaaaa ccaaagcaac actgtgagag tgtgtgtgtt tgtgagaaaa 3120
ccaaaataca tactactgtg atttcattgg gagaaaatct gtttggtata tcaaaaaaag 3180
tagcttaatt acttcatcat tattgggtta ggt 3213

<210> 6

<211> 1296

<212> ADN

5 <213> *Homo sapiens*

<400> 6

taagaacatt ttacactctt cagtataaag aagtcagaat accctaccc tatcagtaaa 60
ggcctataag ttaccattaa aaagatgtcc ttaaaaacag cattctcagc tgggcgcggg 120
ggctcacacc tttgtcccag tactttggga agccgagggt ggtggatcac ctgaggtcag 180
gagttcgaga ccagcctggc caacatggcg aaaaccatt ttctctacta aaaatacaaa 240
aattagccgg gcatgggtggc ggggtgctgt ggtcccagct actcaagagg ctgagggtgg 300
aggatcactg agcccaggag gtggaggctg cattgagcca agattgtgcc actgcactcc 360
agcctgggtg acagagcgag actctgtctc aaaaaaacca aaacaaaaaa aaccagcat 420

tcttttagtaa ataattcata gttttcttca tctagaattt aaaattgtga tagttgatca 480
 gcatgtcctg agcacgtgtg ttgctgttta ctagtttaga tcggtagatg tgtatataag 540
 ttataggtat aaaatcaatc ctgagttgac acaagggttt gatgttgagt acaagtacag 600
 taagtgtata ttttttagtta tgctcttagt ttttaagtcaa ttgtgtggtt ctttctagct 660
 ttaggatctg ttgaattatc ttcccttagaa aaggaggtta agaattctca cttacctatc 720
 ttctacttgt ttggagaata gaagagtccc tgtggtagca gactttgtga gtttaacttgt 780
 aattttccat ctgaaagact gttcttggtt ttctgtgatga agtcttgctc tgtcgcccag 840
 gctggagtgc agtggtgcaa ccttggtcctca ctgcaacctc tgccctcccg gttcaagcaa 900
 ttctcctgcc tcagcctccc gagtatctgg gattacaggt gcacaccacc acacctggct 960
 aatttttgta ttttcagtag agacgggggtt tcaccatggt ggccaggctg gtctcgaact 1020
 cttgacctca tgatcagccc acctcagcct tccaaagtgc tgggattaca ggtgtgagcc 1080
 cccacactcg gccgttggtg ttttttaaga gacaggggtct cactctgtca cctaacctgg 1140
 agtacagtgg caatcatggc tcactgtaac ctcaaagtgc cggccttagt gaagcggtct 1200
 tcctgccttg gcctcccaa gtgctgggat tacaagtgtg agccatgcat ccagcttgaa 1260
 agacagcttc ttaggcttga tttgtttggt tacagg 1296

<210> 7

<211> 24

<212> ADN

5 <213> Secuencia artificial

<220>

<223> Cebador

10 <400> 7

taggtcagtg ggtattaacg aggt 24

<210> 8

<211> 23

<212> ADN

15 <213> Secuencia artificial

<220>

<223> Cebador

20 <400> 8

	gtcactttga tatgatgaga gca	23
	<210> 9	
	<211> 23	
	<212> ADN	
5	<213> Secuencia artificial	
	<220>	
	<223> Cebador	
10	<400> 9	
	cctcgttaat acccactgac cta	23
	<210> 10	
	<211> 24	
	<212> ADN	
15	<213> Secuencia artificial	
	<220>	
	<223> Cebador	
20	<400> 10	
	ggcaatacag ataagataat gtag	24
	<210> 11	
	<211> 21	
	<212> ADN	
25	<213> Secuencia artificial	
	<220>	
	<223> Cebador	
30	<400> 11	
	ttgatcagca tgtcctgagc a	21
	<210> 12	
	<211> 22	
	<212> ADN	
35	<213> Secuencia artificial	

	<220>		
	<223> Cebador		
	<400> 12		
5	ctaccctatc agtaaaggcc ta	22	
	<210> 13		
	<211> 21		
	<212> ADN		
	<213> Secuencia artificial		
10	<220>		
	<223> Cebador		
	<400> 13		
15	tgctcaggac atgctgatca a	21	
	<210> 14		
	<211> 23		
	<212> ADN		
	<213> Secuencia artificial		
20	<220>		
	<223> Cebador		
	<400> 14		
25	aaaaacagca ttctcagctg ggc	23	

REIVINDICACIONES

1. Una molécula de ácido nucleico que comprende una porción 5' de un gen de lactasa-florizina hidrolasa (LPH) intestinal que contribuye a o es indicativa de hipolactasia de tipo adulto, en la que dicha molécula de ácido nucleico se selecciona del grupo consistente en
 - 5 (a) una molécula de ácido nucleico que tiene o comprende la secuencia de ácido nucleico de SEQ ID NO: 1, la secuencia de SEQ ID NO: 1 se exhibe también en la Fig. 4 y comprendida en la secuencia como se exhibe en la Fig. 8; y
 - (b) una molécula de ácido nucleico que tiene o comprende la secuencia de ácido nucleico de SEQ ID NO: 2, la secuencia de SEQ ID NO: 2 se exhibe también en la Fig. 5 y comprendida en la secuencia como se exhibe en la Fig. 9;
 en la que dicha molécula de ácido nucleico se extiende, como máximo, 30.000 nucleótidos más allá del extremo 5' y/o 3' de la molécula de ácido nucleico de SEQ ID NO: 1 o 2, respectivamente.
2. Una molécula de ácido nucleico consistente en una porción 5' de un gen de lactasa-florizina hidrolasa (LPH) intestinal que contribuye a o es indicativa de hipolactasia de tipo adulto, en la que dicha molécula de ácido nucleico se selecciona del grupo consistente en
 - 15 (a) una molécula de ácido nucleico de al menos 20 nucleótidos, cuya hebra complementaria hibrida en condiciones rigurosas con una molécula de ácido nucleico que tiene o comprende la secuencia de ácido nucleico de SEQ ID NO: 1, en la que dicho polinucleótido tiene en la posición correspondiente a la posición -13910 en 5' del primer ATG del gen *LPH* un residuo de citosina; y
 - (b) una molécula de ácido nucleico de al menos 20 nucleótidos, cuya hebra complementaria hibrida en condiciones rigurosas con una molécula de ácido nucleico que tiene o comprende la secuencia de ácido nucleico de SEQ ID NO: 2, en la que dicho polinucleótido tiene en la posición correspondiente a la posición -22018 en 5' del primer ATG del gen *LPH* un residuo de guanina;
 en la que dicha molécula de ácido nucleico se extiende, como máximo, 30.000 nucleótidos más allá del extremo 5' y/o 3' de la molécula de ácido nucleico de SEQ ID NO: 1 o 2, respectivamente.
3. Una molécula de ácido nucleico que comprende una porción 5' de un gen de lactasa-florizina hidrolasa (LPH) intestinal en la que dicha molécula de ácido nucleico se selecciona del grupo consistente en
 - 25 (a) una molécula de ácido nucleico que tiene o comprende la secuencia de ácido nucleico de SEQ ID NO: 3; la secuencia de SEQ ID NO: 3 se exhibe también en la Fig. 6;
 - (b) una molécula de ácido nucleico que tiene o comprende la secuencia de ácido nucleico de SEQ ID NO: 4; la secuencia de SEQ ID NO: 4 se exhibe también en la Fig. 7.
4. Una molécula de ácido nucleico consistente en una porción 5' de un gen de lactasa-florizina hidrolasa (LPH) intestinal, en la que dicha molécula de ácido nucleico se selecciona del grupo consistente en:
 - 30 (a) una molécula de ácido nucleico cuya hebra complementaria hibrida en condiciones rigurosas con una molécula de ácido nucleico que tiene o comprende la secuencia de ácido nucleico de SEQ ID NO: 3, en la que dicha molécula de ácido nucleico tiene en la posición correspondiente a la posición 13910 del primer ATG del gen *LPH* un residuo de timidina; y
 - (b) una molécula de ácido nucleico cuya hebra complementaria hibrida en condiciones rigurosas con una molécula de ácido nucleico que tiene o comprende la secuencia de ácido nucleico de SEQ ID NO: 4, en la que dicha molécula de ácido nucleico tiene en la posición correspondiente a la posición 22018 del primer ATG del gen *LPH* un residuo de adenosina.
5. La molécula de ácido nucleico según una cualquiera de las reivindicaciones 1 a 4 que es ADN genómico.
6. La molécula de ácido nucleico de la reivindicación 5, en la que dicho ADN genómico es parte de un gen.
7. Un fragmento de la molécula de ácido nucleico de una cualquiera de las reivindicaciones 1 a 6 que tiene al menos 14 nucleótidos, en el que dicho fragmento comprende la posición nucleotídica -13910 o la posición nucleotídica -22018 del gen *LPH*.
8. Una molécula de ácido nucleico que es complementaria de la molécula de ácido nucleico de una cualquiera de las reivindicaciones 1 y 2 y 5 a 7
9. Una molécula de ácido nucleico que es complementaria de la molécula de ácido nucleico de una cualquiera de las reivindicaciones 3 a 7
10. Un vector que comprende la molécula de ácido nucleico de una cualquiera de las reivindicaciones 1 o 2 y 5 a 7.

11. Un vector que comprende la molécula de ácido nucleico de una cualquiera de las reivindicaciones 3 a 7.
12. Un cebador o par de cebadores, en el que el cebador o par de cebadores hibrida en condiciones rigurosas con la molécula de ácido nucleico de una cualquiera de las reivindicaciones 1 o 2 y 5 a 7 que comprende la posición nucleotídica -13910 del gen *LPH* o con la hebra complementaria de la misma, en el que el cebador es exactamente complementario de una secuencia que contiene C en la posición -13910 o con la hebra complementaria de la misma.
13. Un cebador o par de cebadores, en el que el cebador o par de cebadores hibrida en condiciones rigurosas con la molécula de ácido nucleico de una cualquiera de las reivindicaciones 1 o 2 y 5 a 7 que comprende la posición nucleotídica -22018 del gen *LPH* o con la hebra complementaria de la misma, en el que el cebador es exactamente complementario de una secuencia que contiene G en la posición -22018 o con la hebra complementaria de la misma.
14. Un cebador o par de cebadores, en el que el cebador o par de cebadores hibrida en condiciones rigurosas con la molécula de ácido nucleico de una cualquiera de las reivindicaciones 3 a 7 que comprende la posición nucleotídica -13910 del gen *LPH* o con la hebra complementaria del mismo, en el que el cebador es exactamente complementario de una secuencia que contiene T en la posición -13910 o con la hebra complementaria de la misma.
15. Un cebador o par de cebadores, en el que el cebador o par de cebadores hibrida en condiciones rigurosas con la molécula de ácido nucleico de una cualquiera de las reivindicaciones 3 a 7 que comprende la posición nucleotídica -22018 del gen *LPH* o con la hebra complementaria de la misma, en el que el cebador es exactamente complementario de una secuencia que contiene A en la posición -22018 o con la hebra complementaria de la misma.
16. Un hospedador no humano transformado con el vector de la reivindicación 10.
17. Un hospedador no humano transformado con el vector de la reivindicación 11.
18. El hospedador no humano de la reivindicación 16 o 17 que es una bacteria, una célula de levadura, una célula de insecto, una célula fúngica, una célula de mamífero, una célula de planta, un animal transgénico o una planta transgénica.
19. Un anticuerpo o aptámero o fago que se une específicamente con la molécula de ácido nucleico mutante de una cualquiera de las reivindicaciones 1 o 2 y 5 a 8, pero no con la correspondiente molécula de ácido nucleico de tipo silvestre, en el que la molécula de ácido nucleico de tipo silvestre tiene en la posición correspondiente a la posición -13910 del gen *LPH* una timidina y/o en la posición correspondiente a la posición -22018 una adenosina, y una molécula de ácido nucleico mutante tiene en la posición -13910 una citosina y/o en la posición correspondiente a la posición -22018 una guanina.
20. Un anticuerpo o aptámero o fago que se une específicamente con la molécula de ácido nucleico de tipo silvestre de una cualquiera de las reivindicaciones 3 a 7 y 9, pero no con la correspondiente secuencia mutante que contribuye a o es indicativa de hipolactasia de tipo adulto, en el que la molécula de ácido nucleico de tipo silvestre tiene en la posición correspondiente a la posición -13910 del gen *LPH* una timidina y/o en la posición correspondiente a la posición -22018 una adenosina, y una molécula de ácido nucleico mutante tiene en la posición correspondiente a la posición -13910 una citosina y/o en la posición correspondiente a la posición -22018 una guanina.
21. Una composición farmacéutica que comprende la molécula de ácido nucleico de tipo silvestre de la reivindicación 3 a 6 o el vector de la reivindicación 11, en la que la molécula de ácido nucleico de tipo silvestre tiene en la posición correspondiente a la posición -13910 del gen *LPH* una timidina y/o en la posición correspondiente a la posición -22018 una adenosina.
22. Una composición de diagnóstico que comprende una molécula de ácido nucleico de una cualquiera de las reivindicaciones 1 a 9, el vector de la reivindicación 10 u 11, el cebador o par de cebadores de la reivindicación 12 a 15 y/o el anticuerpo, aptámero y/o fago de la reivindicación 19 o 20.
23. Un procedimiento para ensayar la presencia de o predisposición a hipolactasia de tipo adulto que comprende ensayar en una muestra obtenida de un paciente previsible o de una persona sospechosa de portar dicha predisposición la presencia de la molécula de ácido nucleico de una cualquiera de las reivindicaciones 1 o 2 y 5 a 8 en estado homocigótico o heterocigótico.
24. Un procedimiento para ensayar la presencia de o predisposición a hipolactasia de tipo adulto que comprende ensayar en una muestra obtenida de un paciente previsible o de una persona sospechosa de portar dicha predisposición la presencia de la molécula de ácido nucleico de una cualquiera de las reivindicaciones 3 a 7 y 9 en estado homocigótico o heterocigótico.
25. El procedimiento de la reivindicación 23 o 24, en el que dicho ensayo comprende hibridar la molécula de ácido nucleico complementaria de la reivindicación 8 que es complementaria de la molécula de ácido nucleico que contribuye a o es indicativa de hipolactasia de tipo adulto, o la molécula de ácido nucleico de la reivindicación 9 que es complementaria de la secuencia de tipo silvestre en forma de sonda en condiciones rigurosas con moléculas de ácido nucleico comprendidas en dicha muestra y detectar dicha hibridación, en el que la molécula de ácido nucleico de tipo

silvestre tiene en la posición correspondiente a la posición -13910 del gen *LPH* una timidina y/o en la posición correspondiente a la posición -22018 una adenosina, y una molécula de ácido nucleico mutante tiene en la posición correspondiente a la posición -13910 una citosina y/o en la posición correspondiente a la posición -22018 una guanina.

- 5 26. El procedimiento de una cualquiera de las reivindicaciones 23 a 24, que comprende adicionalmente digerir el producto de dicha hibridación con una endonucleasa de restricción o someter el producto de dicha hibridación a digestión con una endonucleasa de restricción y analizar el producto de dicha digestión.
27. El procedimiento de la reivindicación 25, en el que dicha sonda está marcada detectablemente.
28. El procedimiento de la reivindicación 23 o 24, en el que dicho ensayo comprende determinar la secuencia de ácido nucleico de al menos una porción de la molécula de ácido nucleico de una cualquiera de las reivindicaciones 1 a 10 9, comprendiendo dicha porción la posición nucleotídica -13910 y/o la posición nucleotídica -22018 del gen *LPH*.
29. El procedimiento de la reivindicación 28, en el que se efectúa la determinación de la secuencia de ácido nucleico mediante minisequenciación en fase sólida.
30. El procedimiento de la reivindicación 28, que comprende adicionalmente, antes de determinar dicha secuencia de ácido nucleico, la amplificación de al menos dicha porción de dicha molécula de ácido nucleico.
- 15 31. El procedimiento de la reivindicación 23 o 24, en el que dicho ensayo comprende llevar a cabo una reacción de amplificación en la que al menos uno de los cebadores empleados en dicha reacción de amplificación sea el cebador de la reivindicación 12 o 13 o pertenezca al par de cebadores de las reivindicaciones 12 o 13, que comprende ensayar un producto de amplificación.
- 20 32. El procedimiento de la reivindicación 23 o 24, en el que dicho ensayo comprende llevar a cabo una reacción de amplificación en la que al menos uno de los cebadores empleados en dicha reacción de amplificación sea el cebador de la reivindicación 14 o 15 o pertenezca al par de cebadores de las reivindicaciones 14 o 15, que comprende ensayar un producto de amplificación.
33. El procedimiento de una cualquiera de las reivindicaciones 30 a 32, en el que se efectúa dicha amplificación mediante, o dicha amplificación es, la reacción en cadena de la polimerasa (PCR).
- 25 34. Un procedimiento para ensayar la presencia de o predisposición a hipolactasia de tipo adulto que comprende ensayar en una muestra obtenida a partir de un ser humano la unión específica con el anticuerpo o aptámero o fago de la reivindicación 19.
- 30 35. Un procedimiento para ensayar la presencia de o predisposición a hipolactasia de tipo adulto que comprende ensayar en una muestra obtenida a partir de un ser humano la unión específica con el anticuerpo o aptámero o fago de la reivindicación 20.
36. El procedimiento de la reivindicación 34 o 35, en el que dicho anticuerpo o aptámero o fago está marcado detectablemente.
37. El procedimiento de una cualquiera de las reivindicaciones 34 a 36, en el que el ensayo es un inmunoensayo.
- 35 38. El procedimiento de una cualquiera de las reivindicaciones 23 a 37, en el que dicha muestra es sangre, suero, plasma, tejido fetal, saliva, orina, tejido mucoso, moco, tejido vaginal, tejido fetal obtenido a partir de vagina, piel, pelo, folículo piloso u otro tejido humano.
39. El procedimiento de una cualquiera de las reivindicaciones 23 a 38, en el que dicha molécula de ácido nucleico de dicha muestra se fija a un soporte sólido.
- 40 40. El procedimiento de la reivindicación 39, en el que dicho soporte sólido es un chip, una oblea de silicio, una perla o una placa de microvaloración.
41. La molécula de ácido nucleico de una cualquiera de las reivindicaciones 1 a 9 para uso en el análisis de la presencia de o predisposición a hipolactasia de tipo adulto.
- 45 42. Kit que comprende la molécula de ácido nucleico de una cualquiera de las reivindicaciones 1 a 9, el cebador o par de cebadores de las reivindicaciones 12 a 15, el vector de la reivindicación 10 u 11 y/o el anticuerpo, aptámero y/o fago de la reivindicación 19 o 20 en uno o más envases.

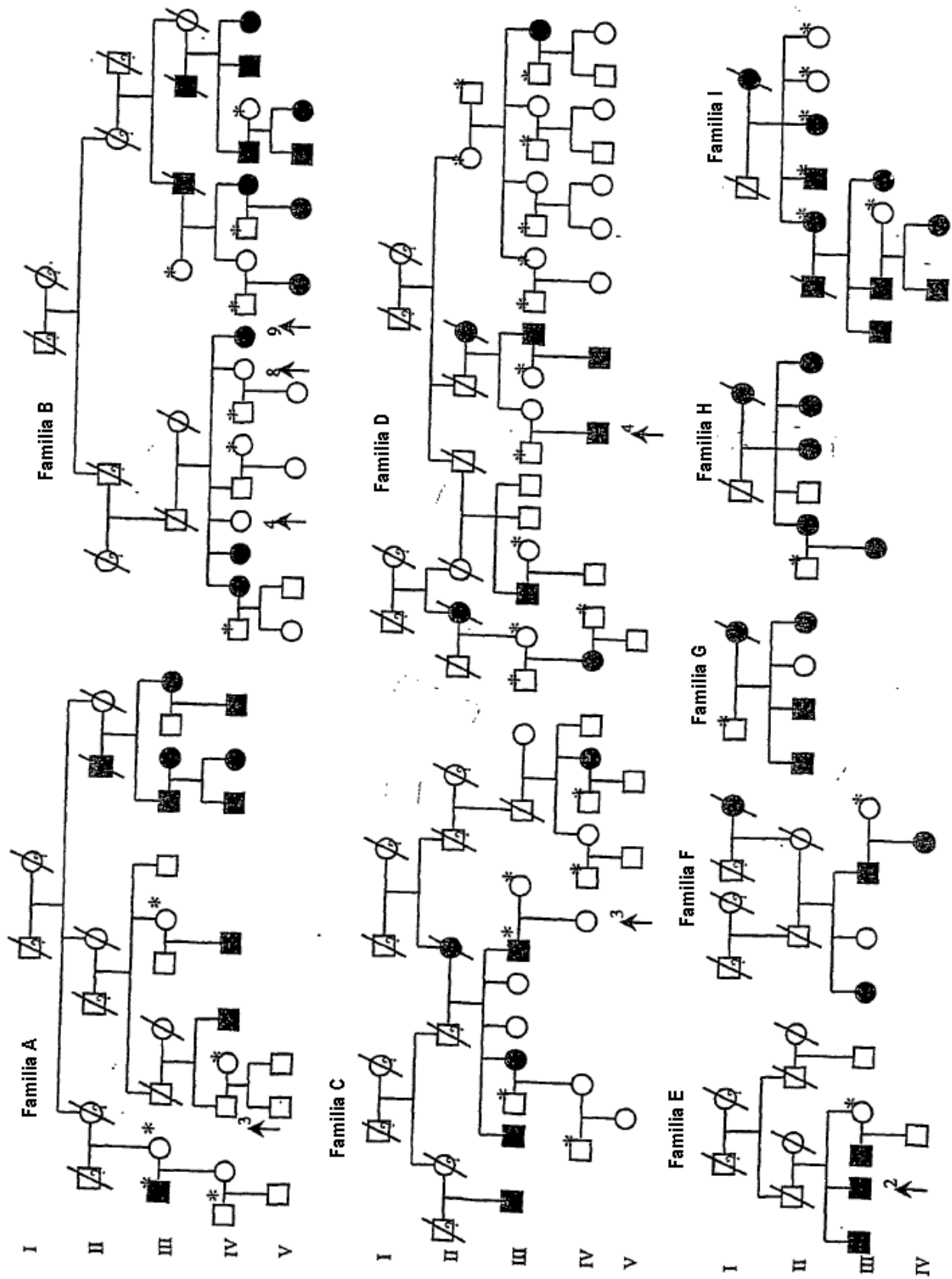


Figura 1

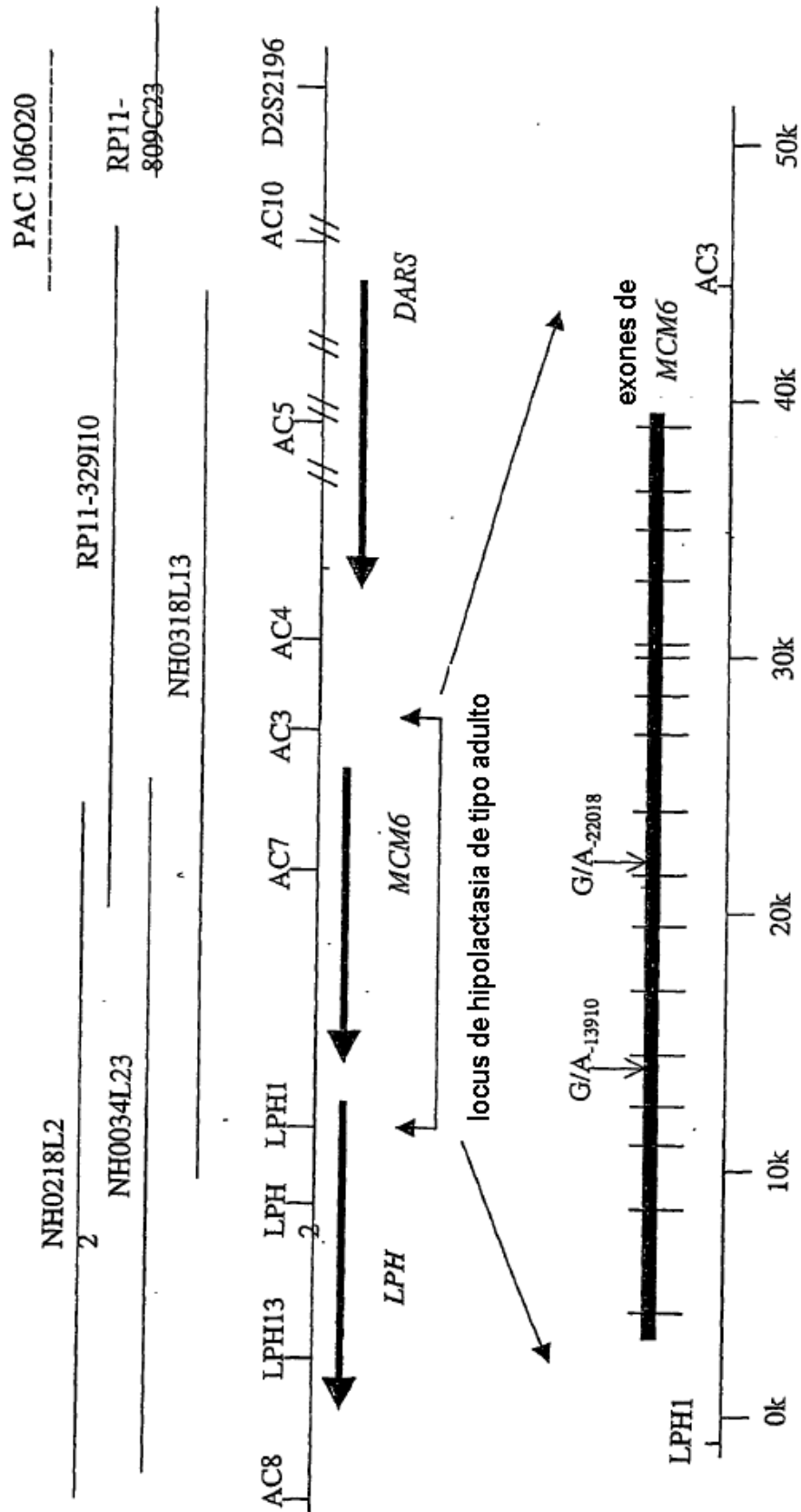


Figura 2

marcador	haplotipo										Total
LPH13	6	6	6	6	6	6	6	6	6	6	6
LPH2	4	4	4	4	4	4	4	4	4	4	4
LPH1	3	2	2	2	1	1	4	4	4	4	4
AC7	5	5	5	5	5	5	5	5	5	5	5
AC3	2	2	2	4	2	2	2	2	2	2	2
AC4	5	5	5	5	5	5	5	5	5	5	5
AC5	6	6	5	6	6	3	6	5	5	3	3
alelos de lactasa persistente	20	4	1	2	2	1	1	1	1	1	33
alelos de lactasa no persistente	3	2	0	0	0	0	2	0	0	2	54

Figura 3

SEQ ID NO:1

ACCTTTCATTCAGGAAAAATGTACTTAGACCCTACAATGTACTAGTAGGCCTCTGCGCTG
GCAATACAGATAAGATAATGTAGcCCCTGGCCTCAAAGGAACTCTCCTCCTTAGGTTGCA
TTTGTATAATGTTTGATTTTGTAGATTGTTCTTTGAGCCCTGCATTCCACGAGGATAGGTC

Figura 4

SEQ ID NO:2

TAAGAACATTTTACACTCTTCAGTATAAAGAAGTCAGAATACCCCTACCCTATCAGTAAA
GGCCTATAAGTTACCATTAAAAAGATGTCCTTAAAAACAGCATTCTCAGCTGGGCgCGGT
GGCTCACACCTTTGTCCCAGTACTTTGGGAAGCCGAGGTGGGTGGATCACCTGAGGTCAG

Figura 5

SEQ ID NO:3

ATCAGAGTCACTTTGATATGATGAGAGCAGAGATAAACAGATTTGTTGCATGTTTTTAAT
 CTTTGGTATGGGACATACTAGAATTCAGTGCAAATACATTTTTATGTAAGTGTGAATGC
 TCATACCACCATGGAATTCTTCCCTTTAAAGAGCTTGGTAAGCATTTGAGTGTAGTTGTT
 AGACGGAGACGATCACGTCATAGTTTATAGAGTGCATAAAGACGTAAGTTACCATTTAAT
 ACCTTTTCATTTCAGGAAAAATGTACTTAGACCCTACAATGTACTAGTAGGCCCTCTGCGCTG
 GCAATACAGATAAGATAATGTAGTCCCTGGCCTCAAAGGAACCTCTCCTCCTTAGGTTGCA
 TTTGTATAATGTTTGATTTTTAGATTGTTCTTTGAGCCCTGCATTCACGAGGATAGGTC
 AGTGGGTATTAACGAGGTAAAAGGGGAGTAGTACGAAAGGGCATTCAGCGTCCCCTCTT
 CGCTTCAACCAAAGCAGCCCTGCGTTTTCTTAGTTTTTATTAATAGGTTTGATGTAAGGTC
 GTCTTTGAAAAGGGGGTTTTGGCTTTTTTTTACAGTGTGACTGAGGTATAATTTATAAAAA
 GGGAAATGTATGGCATGGTGAGTTTTTTCACATACATCCTTGTGAATACCEAGCTCAAGA
 TCCAAAACATTTCCATAATTTAGAAAGTTCCAAACCCCTGCCTCTTTTCAGTCTTAGCC
 CTCTTCCCCTGAAGTAACCACTGTTCCGACTTCAATCACTACTTTTTATCCCACAGGTTAA
 TTTTTTGGCTTTTTTCCACTAAATTTTCAAATTCCTTGATATGGTACTTTACTATTGACG
 AAGTACTTTCACACTAGGTTATTTAATATCTTTGATTCACCCAATATTTAGGGAACACC
 TGTAGGGGACAAAAAATGAATGAGAGCCCCCTGCCTTCCATTGCTGCTAATCTGGTGGGAA
 CGAGACATGTATTTAATTAAGCATGTAAAAAATAGAGTGGGTGATGAAATAATCTATATA
 CTAAATCCCCATGACACACAGTTTACCTATGTAACAAACCTGCATGTGTACCCCCGAACC
 TAAATATAAGTTGGAAATTAACAAAAAACGAGAGGGAGAATAGAGCATCACAAACAGAG
 TGCTGAGATGAATTACTTTATTACCAAAGAAGGAGGAGGACTCAGGGAGGTGCCGACGTT
 TAAACCCAGTCACTGAAGGGTGTGCAGAATTTGGATAGGCAAGATACCCTGGGACAAGGT
 CATTCATAAACCATGCTAACATTTGTACTTTTTTTTTTCATTGTGATAGTTCTTGAAATGA
 GTTGCATAAAACTGGTACATGTCTTAGGGCAGTCTCTAATTGATTTTTATTTTGTCTAT
 TTTTAAAAATTAGTCTTCAAATAGCAGATTCACATGATATTAATAATATATGCACATAAAT
 TATATACACAAATATATTTTCTGAATGAAATTTAGTATCTGCATATATTTAAGAGCTATT
 TCTGTCTCATATGTTTCATAATCTTCATCCATTAAAAAACTTTTGTTAGGCCTTTCTCAC
 TCTAAGATTATAAAAAATTCTCCCATTTATTACCTAGCTAGTTTTCTAGTTGTTCCAAAA
 CCATTTATTGAACAATCCATCTTTTTTGACACTGGTTTGGCATGCCTTAATTATATATTCT
 TGTGTGTGTTAGGATCTCCTTTTGGACTTTCATTCTGTTTATTGAGTCTTATCAGCTCC
 TCTTACATTGGTACCATGATGTTTTAATCTATGGGGCTTTGTAGTTTAAATGTAGGGCTA
 GTTCCAGCGCATTGTTCTCTATCAGCTGTTAGGAACTTAGAAATCAGCTTGCTCTGTTTT
 AAAGAAAAACCTGGTATTTTTTTTATCAGTATAACATTCTATTTATATTAAGTGAAGAAT
 TGAAAAACATCTATGATTTTTTCTTATTCAGTAACGTATCACTTAGAATAGGTTAGGTTGTA
 CTACTATAAAATCTCAGCTGCATAAAACAATTTTTTTTTTGCTTGTGCTACACATCCATTA
 GGTCATCAAGGGACTCACCTTGTCAAGTTACTCAGAGATTCAGGCTGATATAAAGGTTTG
 ATCTTGACATACGCTTTCATGATGACAGAAAGCAGGGAAGAGAAGGTGGTGAGCCATGTG
 CTTTCTCCCCCTTCTATCCAGAAATGACACATACTCACATTTCAATTCGCCAGAGAAATTA
 ACATGGCCCCCTCCTAAGTTCAAATGGATAGAGAAATGCCTTCTACCAGGTGCCCAGAAT
 TAGAAGAGCAAACATTTGTGAACAGTCTGAGTACCACAAATACCGTTATCTTTCCACTT
 AAGTCTTCTGTTTCACTCAGTAGTGCTTTAAACTTTTCTTCATATGTTTTTCAGTGTTTC
 TTGTTGAATTTCTTGATATTTTATCATGTTTGTTCGTACTGGGAGTAGCCTTTTTTTCCA

Figura 6

TTTCATTTTCTGGCTGGTTTCATTGCTGGTTGTTTTTTGTTTTGTTTTGTTTTGAGAT
 GGAGTCTCACTCTGTCGCCCAGGCTGGAGTGCAGTGTCAACAATCTCGGCTCACTGCAACC
 TCTGCCTCCCAAGTTCAAGCGATTCTTCTTTCTCAGCCTCCTGAGTAGCTGGGATTACAG
 GCATGTGCCACCATGCCCAGCTAATTTTTTATATTTTATAGTAGAGATGGGGTTTCTCCAT
 GTTGGTCAGGCTGGTCTCAAACCTCCCAATCTCAGGTGATCCGCCTGCCTCTGCCTTCCAA
 AGTGCTGGGATTATAGACATGAGCCACCGTGCCTGGCCTAGTTCTTATGGGATGTATATG
 TCTTTGGATTATATGATATGTATATATGTTTATATTTCTACAAGTACATACCTAGGAGT
 GGAATTGTTGGGTCATAGGTTAATGCATGTTTTTCTGCCAAACAGTTGTGTCAATTTCTG
 TTTTCACCGCTGTGAATGAGAGTTGTTCTACCTTCTTGACAACACTTGATATTGTCAGTC
 ATTTTAGCCATTCTGGTGAATTTATAGTGCTATTTCTGTGTGTGTAAGAGAGAGAATGAG
 AGAGGGTGTTTGTGAGAAAACCAAAGCAACACTGTGAGAGTGTGTGTGTTTGTGAGAAAA
 CCAAAATACATACTACTGTGATTTTCATTGGGAGAAAAATCTGTTTGGTATATCAAAAAAAG
 TAGCTTAATTACTTCATCATTATTGGTTTAGGT

Figura 6 cont.

SEQ ID NO:4

TAAGAACATTTTACACTCTTCAGTATAAAGAAGTCAGAATACCCCTACCCTATCAGTAAA
 GGCTATAAGTTACCATTAAAAAGATGTCCTTAAAAACAGCATTCCTCAGCTGGGCaCGGT
 GGCTCACACCTTTGTCCCAGTACTTTGGGAAGCCGAGGTGGGTGGATCACCTGAGGTCAG
 GAGTTCGAGACCAGCCTGGCCAACATGGCGAAAACCCATTTTCTCTACTAAAAATACAAA
 AATTAGCCGGGCATGGTGGCGGGTGCTTGTGGTCCCAGCTACTCAAGAGGCTGAGGTGGG
 AGGATCACTGAGCCCAGGAGGTGGAGGCTGCATTGAGCCAAGATTGTGCCACTGCACTCC
 AGCCTGGGTGACAGAGCGAGACTCTGTCTCAAAAAAACCAAAACAAAAAAAACCCAGCAT
 TCTTTAGTAAATAATTCATAGTTTTCTTCATCTAGAATTTAAAATTGTGATAGTTGATCA
 GCATGTCCTGAGCACGTGTGTTTGCTGTTACTAGTTTAGATCGGTAGATGTGTATATAAG
 TTATAGGTATAAAATCAATCCTGAGTTGACACAAGGTTTTGATGTTGAGTACAAGTACAG
 TAAGTGTATATTTTTAGTTATGCTCTTAGTTTTAAGTCAATTGTGTGGTTCTTTCTAGCT
 TTAGGATCTGTTGAATTATCTTCCTTAGAAAAGGGAGTTAAGAATCTTCACTTACCTATC
 TTCTACTTGTTTGGAGAATAGAAGAGTCCCTGTGGTAGCAGACTTTGTGAGTTTACTTGT
 AATTTTCCATCTGAAAGACTGTTCTTGTTTTTCGTGATGAAGTCTTGCTCTGTCGCCCAG
 GCTGGAGTGCAAGTGGTGCAACCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGCAA
 TTCTCCTGCCTCAGCCTCCCGAGTATCTGGGATTACAGGTGCACACCACCACACCTGGCT
 AATTTTTGTATTTTCAGTAGAGACGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGAACT
 CTTGACCTCATGATCAGCCCACCTCAGCCTTCCAAAGTGCTGGGATTACAGGTGTGAGCC
 CCCACACTCGGCCGTTGTTGTTTTTTAAGAGACAGGGTCTCACTCTGTACCTAACCTGG
 AGTACAGTGGCAATCATGGCTCACTGTAACCTCAAATGCCCGGCCTTAGTGAAGCGTTCT
 TCCTGCCTTGGCCTCCCAAAGTGCTGGGATTACAAGTGTGAGCCATGCATCCAGCTTGAA
 AGACAGCTTCTTAGGCTTGATTTGTTTGGTTACAGG

Figura 7

SEQ ID NO:5

ATCAGAGTCACTTTGATATGATGAGAGCAGAGATAAACAGATTTGTTGCATGTTTTTAAT
 CTTTGGTATGGGACATACTAGAATTCAGTCAAATACATTTTTATGTAAGTGTGAATGC
 TCATACGACCATGGAATTCTTCCCTTTAAAGAGCTTGGTAAGCATTTGAGTGTAGTTGTT
 AGACGGAGACGATCACGTCATAGTTTATAGAGTGCATAAAGACGTAAGTTACCATTTAAT
 ACCTTTCATTCAAGGAAAAATGTACTTAGACCCTACAATGTACTAGTAGGCCTCTGCGCTG
 GCAATACAGATAAGATAATGTAGCCCCTGGCCTCAAAGGAACTCTCCTCCTTAGGTTGCA
 TTTGTATAATGTTTGATTTTTAGATTGTTCTTTGAGCCCTGCATTCCACGAGGATAGGTC
 AGTGGGTATTAACGAGGTAAAAGGGGAGTAGTACGAAAGGGCATTCAGCGTCCCATCTT
 CGCTTCAACCAAGCAGCCCTGCGTTTTCTAGTTTATTAATAGGTTTGATGTAAGGTC
 GTCTTTGAAAAGGGGGTTTGGCTTTTTTTTACAGTGTGACTGAGGTATAATTTATAAAAA
 GGGAAATGTATGGCATGGTGAGTTTTTTCACATACATCCTTGTGAATACCCAGCTCAAGA
 TCCAAAACATTTCCATAATTTCAAGAAAGTTCCAAACCCCTGCCTCTTTTCAGTCTTAGCC
 CTCTTCCCCTGAAGTAACCACTGTTCCGACTTCAATCACTACTTTTATCCACAGGTAA
 TTTTTTGGCTTTTTTCCACTAAATTTTCAAATCTTTGATATGGTACTTTACTATTGACG
 AAGTACTTTCACACTAGGTTATTTAATATTCTTTGATTCAACCAATATTTAGGGAACACC
 TGTAGGGGACAAAAAATGAATGAGAGCCCCCTGCCTTCCATTGCTGCTAATCTGGTGGGAA
 CGAGACATGTATTTAATTAAGCATGTAAAAAATAGAGTGGGTGATGAAATAATCTATATA
 CTAAATCCCCATGACACACAGTTTACCTATGTAACAAACCTGCATGTGTACCCCCGAACC
 TAAATATAAGTTGGAAATTAACAAAAAACGAGAGGGAGAATAGAGCATCACAACCAGAG
 TGCTGAGATGAATTACTTTATTACCAAAGAAGGAGGAGGACTCAGGGAGGTGCCGACGTT
 TAAACCCAGTCACTGAAGGGTGTGCAGAATTTGGATAGGCAAGATACCCTGGGACAAGGT
 CATCTAAAACCATGCTAACATTTGTACTTTTTTTTTTCATTGTGATAGTTCCTGAAATGA
 GTTGCAATAAACTGGTACATGTCTTAGGGCAGTCTCTAATTGATTTTTATTTTGTCTAT
 TTTTAAAAATTAGTCTTCAAATAGCAGATTCACATGATATTAAAAATATATGCACATAAAT
 TATATACACAAATATATTTTCTGAATGAAATTTAGTATCTGCATATATTTAAGAGCTATT
 TCTGTCTCATATGTTTCATAATCTTCATCCATTAAAAAACTTTTGTAGGCCTTCTCAC
 TCTAAGATTATAAAAAATTCTCCCATTATTTACCTAGCTAGTTTTCTAGTTGTTCCAAAA
 CCATTTATTTGAACAATCCATCTTTTTTGACACTGGTTTGGCATGCCTTAATTATATATTCT
 TGTGTGTGTTAGGATCTCCTTTTGGACTTTCCATTCTGTTCATTGAGTCTTATCAGCTCC
 TCTTACATTGGTACCATGATGTTTTAATCTATGGGGCTTTGTAGTTTAAATGTAGGGCTA
 GTTCCAGCGCATTGTTCTCTATCAGCTGTTAGGAACTTAGAAATCAGCTTGCTCTGTTTT
 AAAGAAAAACCTGGTATTTTTTTTATCAGTATAACATTCTATTTATATTAACCTGAAGAAT
 TGAAAACATCTATGATTTTTTCTATTCAGTAACGTATCACTTAGAATAGGTTAGGTTGTA
 CTACTATAAAATCTCAGCTGCATAAAACAATTTTTTTTTTGCTTGTGCTACACATCCATTA
 GGTCATCAAGGGACTCACCTTGTCAAGTTACTCAGAGATTCAGGCTGATATAAAGGTTTG
 ATCTTGACATACGCTTTCATGATGACAGAAAGCAGGGAAGAGAAGGTGGTGAGCCATGTG
 CTTTCTCCCCCTTCTATCCAGAAATGACACATACTCACATTTTCATTGCGCCAGAGAAATTA
 ACATGGCCCCCTCCTAAGTTCAAATGGATAGAGAAATGCCCTTCCCTACCAGGTGCCCAGAAT
 TAGAAGAGCAAACATTTGTGAACAGTTCTGAGTACCACAAATACCGTTATCTTTCCACTT
 AAGTCTTCTGTTTCACTCAGTAGTGCTTTAACTTTTCTTCATATGTTTTTCAGTGTTTC
 TTGTTGAATTTCTTGATATTTTATCATGTTTGTTCGTACTGGGAGTAGCCTTTTTTTTCCA

Figura 8

TTTCATTTTCTGGCTGGTTTCATTGCTGGTTGTTTTTTGTTTTGTTTTGTTTTGAGAT
 GGAGTCTCACTCTGTGCGCCAGGCTGGAGTGCAGTGTACAATCTCGGCTCACTGCAACC
 TCTGCCTCCCAGGTTCAAGCGATTCTTCTTTCTCAGCCTCCTGAGTAGCTGGGATTACAG
 GCATGTGCCACCATGCCCAGCTAATTTTTTATATTTTATAGTAGAGATGGGGTTTCTCCAT
 GTTGGTCAGGCTGGTCTCAAACCTCCCAATCTCAGGTGATCCGCCTGCCTCTGCCTTCCAA
 AGTGCTGGGATTATAGACATGAGCCACCGTGCCTGGCCTAGTTCTTATGGGATGTATATG
 TCTTTGGATTTCATATGATATGTATATATGTTTATATTTCTACAAGTACATACCTAGGAGT
 GGAATTGTTGGGTCATAGGTTAATGCATGTTTTTCTGCCAAACAGTTGTGTCAATTTCTG
 TTTTCACCGCTGTGAATGAGAGTTGTTCTACCTTCTTGACAACACTTGATATTGTCAGTC
 ATTTTAGCCATTCTGGTGAATTTATAGTGCTATTTCTGTGTGTGTAAGAGAGAGAATGAG
 AGAGGGTGTTTGTGAGAAAACCAAAGCAACACTGTGAGAGTGTGTGTGTTTGTGAGAAAA
 CCAAATACATACTACTGTGATTTTCATTGGGAGAAAATCTGTTTGGTATATCAAAAAAAG
 TAGCTTAATTACTTCATCATTATTGGTTTAGGT

Figura 8 cont.

SEQ ID NO:6

TAAGAACATTTTACACTCTTCAGTATAAAGAAGTCAGAATACCCCTACCCTATCAGTAAA
 GGCTTATAAGTTACCATTAATAAAGATGTCCTTAATAAACAGCATTCTCAGCTGGGC_gCGGT
 GGCTCACACCTTTGTCCCAGTACTTTGGGAAGCCGAGGTGGGTGGATCACCTGAGGTCAG
 GAGTTCGAGACCAGCCTGGCCAACATGGCGAAAACCCATTTTCTCTACTAAAAATACAAA
 AATTAGCCGGGCATGGTGGCGGGTGCTTGTGGTCCCAGCTACTCAAGAGGCTGAGGTGGG
 AGGATCACTGAGCCCAGGAGGTGGAGGCTGCATTGAGCCAAGATTGTGCCACTGCACTCC
 AGCCTGGGTGACAGAGCGAGACTCTGTCTCAAAAAAACCAAAACAAAAAAAACCCAGCAT
 TCTTTAGTAAATAATTCATAGTTTCTTCATCTAGAATTTAAAATTGTGATAGTTGATCA
 GCATGTCCTGAGCACGTGTGTTTGCTGTTACTAGTTTAGATCGGTAGATGTGTATATAAG
 TTATAGGTATAAAATCAATCCTGAGTTGACACAAGGTTTGTGATGTTGAGTACAAGTACAG
 TAAGTGTATATTTTTAGTTATGCTCTTAGTTTAAAGTCAATTGTGTGGTCTTTCTAGCT
 TTAGGATCTGTTGAATTATCTTCCTTAGAAAAGGGAGTTAAGAATCTTCACTTACCTATC
 TTCTACTTGTTTGGAGAATAGAAGAGTCCCTGTGGTAGCAGACTTTGTGAGTTTACTTGT
 AATTTTCCATCTGAAAGACTGTTCTTGTTTTTCGTGATGAAGTCTTGCTCTGTCGCCAG
 GCTGGAGTGCAGTGGTGCAACCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGCAA
 TTCTCCTGCCTCAGCCTCCCGAGTATCTGGGATTACAGGTGCACACCACCACACCTGGCT
 AATTTTTGTATTTTCAGTAGAGACGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGAACT
 CTTGACCTCATGATCAGCCCACCTCAGCCTTCCAAAGTGCTGGGATTACAGGTGTGAGCC
 CCCACACTCGGCCGTTGTTGTTTTTTAAGAGACAGGGTCTCACTCTGTCACCTAACCTGG
 AGTACAGTGGCAATCATGGCTCACTGTAACCTCAAATGCCCGGCCTTAGTGAAGCGTTCT
 TCCTGCCTTGGCCTCCCAAAGTGCTGGGATTACAAGTGTGAGCCATGCATCCAGCTTGAA
 AGACAGCTTCTTAGGCTTGATTTGTTTGGTTACAGG

Figura 9