(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 783 698**

(21) Número de solicitud: 201930240

(51) Int. CI.:

C12Q 1/6837 (2008.01)

(12) PATENTE DE INVENCIÓN CON EXAMEN B2

(22) Fecha de presentación:

15.03.2019

(43) Fecha de publicación de la solicitud:

17.09.2020

Fecha de modificación de las reivindicaciones:

20.09.2023

Fecha de concesión:

13.11.2023

(45) Fecha de publicación de la concesión:

20.11.2023

(73) Titular/es:

UNIVERSIDAD DEL PAÍS VASCO / EUSKAL HERRIKO UNIBERTSITATEA (100.0%)
Barrio Sarriena, s/n
48940 Leioa (Bizkaia) ES

(72) Inventor/es:

BILBAO CATALÁ, José Ramón y
FERNÁNDEZ JIMÉNEZ, Nora

(74) Agente/Representante:

ARIAS SANZ, Juan

Observaciones:

La lista de secuencias está accesible al público en la página web de la OEPM para su descarga en formato electrónico.

(54) Título: **MÉTODO DE DIAGNÓSTICO DE ENFERMEDAD CELIACA BASADO EN EL NIVEL DE EXPRESIÓN DEL GEN UBE2L3**

(57) Resumen:

Método de diagnóstico de enfermedad celiaca basado en el nivel de expresión del gen UBE2L3.

La presente invención se refiere a métodos para el diagnóstico de la enfermedad celiaca basados en la cuantificación del nivel de expresión del gen UBE2L3 así como de la expresión relativa de la isoforma 2 del gen UBE2L3 con respecto a la expresión total del dicho gen. Asimismo, la invención se refiere a kits con los medios necesarios para realizar el diagnóstico según los métodos de la invención.

Aviso: Se puede realizar consulta prevista por el art. 41 LP 24/2015.
Dentro de los seis meses siguientes a la publicación de la concesión en el Boletín Oficial de la Propiedad Industrial cualquier persona podrá oponerse a la concesión. La oposición deberá dirigirse a la OEPM en escrito motivado y previo pago de la tasa correspondiente (art. 43 LP 24/2015).

DESCRIPCIÓN

MÉTODO DE DIAGNÓSTICO DE ENFERMEDAD CELIACA BASADO EN EL NIVEL DE EXPRESIÓN DEL GEN UBE2L3

CAMPO DE LA INVENCIÓN

La presente invención se refiere a métodos y kits para el diagnóstico de la enfermedad celíaca. Por lo tanto la presente invención se engloba dentro del campo técnico de la medicina, específicamente en el campo técnico del diagnóstico clínico y más específicamente en el diagnóstico de pacientes que padecen celiaquía.

ANTECEDENTES DE LA INVENCIÓN

La celiaquía o enfermedad celíaca es una enfermedad autoinmune que aparece en personas genéticamente predispuestas y que se caracteriza por una inflamación crónica de la parte proximal del intestino delgado, causada por la exposición a la gliadina. La gliadina es uno de los componentes principales del gluten, por lo que el único tratamiento a día de hoy para la celiaquía es la estricta ausencia de gluten en la dieta.

Los síntomas de la enfermedad celíaca incluyen diarrea crónica, retraso del crecimiento y/o desarrollo infantil, fatiga, erupciones en la piel, pérdida de peso, cambios en el carácter (irritabilidad, apatía, introversión, tristeza), vómitos, distensión abdominal, pérdida de apetito, fatiga, etc.

La prevalencia de la enfermedad celíaca en europeos y sus descendientes es del 1%, siendo más frecuente en las mujeres con una proporción 2:1. Habitualmente la celiaquía se manifiesta durante la etapa infantil, aunque se ha demostrado que también puede manifestarse a lo largo de la etapa adulta.

Los pacientes tratados con dietas libres de gluten muestran menores actividades séricas de antígenos frente al gluten. Sin embargo, esa reducida actividad al gluten, tras la introducción de este tipo de dietas, puede ser consecuencia del catabolismo de los anticuerpos ya formados, así como de una disminución de su síntesis por una menor estimulación antigénica al gluten, y no necesariamente por una mejora en la integridad de la mucosa intestinal.

La sensibilidad al gluten auto-diagnosticada y la dieta sin gluten autoimpuesta son fenómenos en auge. No obstante el 10% de la población general y hasta un 12 o 13%

en países como Italia o Reino Unido, reporta una sensibilidad al gluten no diagnosticada, a pesar de que la prevalencia de la EC es en principio mucho menor. En este contexto, una herramienta diagnóstica menos invasiva y más eficiente que una provocación con gluten y su consiguiente biopsia constituiría un gran avance clínico.

La estrategia actual de diagnóstico para la enfermedad celíaca se basa fundamentalmente en demostrar una enteropatía dependiente del gluten, para lo cual aún se utilizan preferentemente métodos invasivos (biopsias), aunque también se analizan marcadores séricos como los anticuerpos frente a la gliadina. Sin embargo, aún no hay métodos analíticos fiables para el seguimiento de la efectividad del tratamiento propuesto, ya que el análisis histológico de biopsias duodenales no es factible como método de monitorización rutinario para analizar la remisión y eficacia del tratamiento.

Se ha propuesto la determinación de los niveles de anticuerpos frente a otros antígenos alimentarios, así como las medidas de la permeabilidad intestinal como métodos de diagnóstico alternativos de celiaquía. Estos métodos tendrían como ventaja la ausencia de requerimiento de muestras de biopsias.

Los métodos analíticos actuales de la enfermedad celíaca basados en el análisis histológico de biopsias duodenales tampoco permiten evaluar la adherencia a una dieta libre de gluten por parte de los celíacos. Se ha sugerido que el análisis de los cambios cuantitativos y cualitativos en anticuerpos relacionados con la enfermedad celíaca puede ser de utilidad para monitorizar la adherencia al tratamiento con dieta libre de gluten, y de hecho, algunos de los anticuerpos analizados para ello son los anticuerpos frente a la transglutaminasa (anti-tTG) y a la gliadina desaminada (AGA). Sin embargo, estos análisis no reflejan necesariamente el grado real de integridad intestinal, ya que una disminución en esos anticuerpos puede ser consecuencia de un aumento de su catabolismo o una disminución de su síntesis por menor estimulación antigénica al gluten en las dietas libres de gluten.

Se ha descrito que la detección de tetrámeros HLA-DQ-gluten en sangre podría emplearse como método diagnóstico de la EC en ausencia de gluten de la dieta, con sensibilidad del 90% y especificidad del 93% (Sarna VK *et al*., Gastroenterology. 2018;154:886-96)

Por lo tanto, a la vista del estado de la técnica, sigue existiendo la necesidad de desarrollar métodos que permitan tanto el diagnóstico de la enfermedad celíaca cuya

fiabilidad sea mayor o al menos aporten más información que la de los métodos descritos hasta la fecha

COMPENDIO DE LA INVENCIÓN

En un primer aspecto la invención se refiere a un método *in vitro* para el diagnóstico de la enfermedad celiaca en un sujeto sospechoso de padecer enfermedad celiaca que comprende cuantificar en una muestra de dicho sujeto el contenido en:

(i) el mRNA que codifica la isoforma 2 del gen UBE2L3 y/o

(ii) el mRNA total del gen UBE2L3

en donde:

i) un nivel reducido del mRNA que codifica la isoforma 2 del gen UBE2L3 con respecto a un valor de referencia,

ii) un nivel elevado del mRNA total del gen UBE2L3 con respecto a un valor de referencia y/o

iii) un contenido relativo del mRNA que codifica la isoforma 2 del gen UBE2L3 con respecto al total de mRNA del gen UBE2L3 elevado con respecto a un valor de referencia es indicativo de que el sujeto sufre enfermedad celiaca.

En un segundo aspecto, la invención se refiere a un kit para la puesta en práctica de un método tal y como se define en las reivindicaciones 1 a 10 que comprende:

(i) Una sonda que hibrida específicamente con el exón 4 del gen UBE2L3,

(ii) Una sonda que hibrida específicamente con el exón 5 del gen UBE2L3 y, opcionalmente

(iii) Una sonda que hibrida de forma específica con un gen de expresión constitutiva

en donde los componentes (i) y (ii) constituyen al menos un 1% del total de las sondas presentes en el kit.

En un tercer aspecto, la invención se refiere a un método para el diagnóstico y tratamiento de un sujeto que comprende:

(i) diagnosticar en dicho sujeto la presencia de enfermedad celiaca mediante el método *in vitro* de diagnóstico de la invención;

(ii) administrar al sujeto diagnosticado con enfermedad celiaca en la etapa (i) un tratamiento adecuado para dicha enfermedad.

BREVE DESCRIPCIÓN DE LAS FIGURAS

Figura 1: Localización genómica de las isoformas del gen UBE2L3 en el cromosoma 22 (flecha señalando la variante 2; números 4 y 5 señalando el cuarto y quinto exón, respectivamente). En el esquema se muestran el SNP priorizado mediante aleatorización mendeliana (AM) (rs5754217) y las sondas del microarray de expresión de Illumina realizado en pacientes de EC en dieta sin gluten (ILMN_1796830 e ILMN_1677877).

Figura 2: **A**. Expresión en unidades de intensidad reportada por las dos sondas de Illumina localizadas en el gen UBE2L3 en controles (círculos) y celiacos en dieta sin gluten (cuadros). **B**: Expresión relativa entre la variante 2 y el resto de isoformas del gen UBE2L3 calculada a partir de las unidades de intensidad reportadas por las sondas de Illumina de forma independiente.

Figura 3: curva ROC con el score de expresión relativa de UBE2L3 como clasificador binario (EC en dieta sin gluten / control).

DESCRIPCIÓN DETALLADA DE LA INVENCIÓN

Como se muestra en los ejemplos de la solicitud, los autores de la presente invención han demostrado que una disminución en la expresión de la isoforma 2 del gen UBE2L3 con respecto a un grupo control es indicativo de la enfermedad celiaca, con una gran especificidad y sensibilidad (AUC=0,997). Asimismo, un aumento en la expresión relativa de la isoforma 2 del gen UBE2L3 en relación a la expresión de todas las isoformas del gen, con respecto a un grupo control, es también indicador de la enfermedad celiaca. En este caso la especificidad del método de diagnóstico es del 100% (AUC = 1). El altísimo potencial predictivo del *score* de expresión relativa de *UBE2L3* lo sitúa en un lugar privilegiado de cara al desarrollo de un test diagnóstico en sangre periférica que podría eliminar las provocaciones con gluten del diagnóstico de la EC en ausencia de gluten de la dieta.

Métodos de diagnóstico

Así, en base a los hallazgos realizados, en un primer aspecto la invención se relaciona con un método *in vitro* para el diagnóstico de la enfermedad celiaca en un sujeto sospechoso de padecer enfermedad celiaca que comprende cuantificar en una muestra de dicho sujeto el contenido en:

(i) el mRNA que codifica la isoforma 2 del gen UBE2L3 y/o

(ii) el mRNA total del gen UBE2L3

en donde:

- un nivel reducido del mRNA que codifica la isoforma 2 del gen UBE2L3 con respecto a un valor de referencia,
- un nivel elevado del mRNA total del gen UBE2L3 con respecto a un valor de referencia y/o
- un contenido relativo del mRNA que codifica la isoforma 2 del gen UBE2L3 con respecto al total de mRNA del gen UBE2L3 elevado con respecto a un valor de referencia

es indicativo de que el sujeto sufre enfermedad celiaca.

El término "diagnóstico" como se usa aquí se refiere tanto al proceso de tratar de determinar y/o identificar una posible enfermedad en un sujeto, es decir, el procedimiento de diagnóstico, y la opinión alcanzada por este proceso, es decir, la opinión de diagnóstico. Como tal, también puede ser considerado como un intento de clasificación de la condición de un individuo en categorías separadas y distintas que permiten decisiones médicas sobre el tratamiento y el pronóstico a realizar.

Como entenderá el experto en la materia, tal diagnóstico puede no ser correcto para el 100% del sujeto a diagnosticar, aunque es preferible que lo sea. Sin embargo, el término requiere que una parte estadísticamente significativa de los sujetos pueda ser identificada como que padece una enfermedad o que tiene una predisposición a la misma. El experto en la materia puede determinar si un resultado es estadísticamente significativo usando diferentes herramientas de evaluación estadística bien conocidas, por ejemplo, mediante la determinación de los intervalos de confianza, la determinación del valor p, la prueba de Student, el Mann-Whitney, etc. (ver Dowdy y Wearden, 1983). Los intervalos de confianza preferidos son al menos 50%, al menos 60%, al menos 70%, al menos 80%, al menos 90%, o al menos 95% o al menos 99%. Los valores de p son preferiblemente 0.05, 0.025, 0.001 o menos.

La expresión "método de diagnóstico" a la que se hace referencia de acuerdo con la presente invención significa que el método puede consistir esencialmente en los pasos mencionados anteriormente o puede incluir pasos adicionales.

Como se usa en este documento, la expresión "diagnóstico *in vitro*" se refiere a un método de diagnóstico que no se aplica a un cuerpo humano o animal.

El término “celiaquía” o “enfermedad celíaca” en la presente invención hace referencia a una patología autoinmune caracterizada por una inflamación crónica de la parte próxima del intestino delgado o yeyuno, causada por la exposición a la gliadina, que es uno de los componentes del gluten. La enfermedad celíaca supone una intolerancia permanente al gluten cuya característica es una reacción inflamatoria de base inmune en la mucosa intestinal que dificulta la absorción de los nutrientes.

El gluten es una proteína presente en cereales tales como trigo (*Triticum spp*), cebada (*Hordeum vulgare*), centeno (*Secale cereale*), triticale (Triticosecale, cereal procedente del cruce entre trigo y centeno), kamut (*Triticum turgidum*), espelta (*Triticum spelta*) y posiblemente avena (Avena spp), pero ausente en arroz (Oriza sativa) y maíz (*Zea mays*). El gluten se compone de gliadina y glutenina, siendo la gliadina una glucoproteína de tipo prolamina. En el sujeto que padece la celiaquía o celíaco, la ingestión de gliadina provoca que la enzima transglutaminasa tisular modifique dicha proteína y el sistema inmune origine una reacción cruzada contra el intestino delgado, causando una reacción inflamatoria que causa atrofia de las vellosidades que recubren el intestino e interferencia en la absorción de nutrientes.

El desarrollo de la enfermedad celíaca está determinado tanto por factores ambientales (alimentación) como genéticos. Así, la celiaquía implica también una predisposición genética ya que la mayor parte de celíacos presentan el antígeno de leucocito humano (HLA) de tipo DQ2 (HLA-DQ2) y DQ8 (HLA-DQ8) (May-Ling J. et al. 2010 Immunogenetics 62:641-651).

El término “sujeto”, tal como se usa en la invención, se refiere a todos los animales clasificados como mamíferos e incluye, pero no está restringido a, animales domésticos y de granja, primates y humanos, por ejemplo, seres humanos, primates no humanos, vacas, caballos, cerdos, ovejas, cabras, perros, gatos o roedores. Preferiblemente, el sujeto es un ser humano hombre o mujer de cualquier edad o raza.

La expresión "que comprende" pretende incluir, pero sin limitación, lo que sigue a la expresión "que comprende". Por lo tanto, el uso de la expresión "que comprende" indica que los elementos citados son necesarios u obligatorios, pero que son opcionales otros elementos y existe la posibilidad de que estén presentes o no.

"Cuantificar", en relación a una secuencia de ácido nucleico, se refiere al uso de cualquier método para estudiar la cantidad de una secuencia de ácido nucleico particular, incluyendo, sin limitación, métodos para determinar el número de copias de

una secuencia de ácido nucleico o para determinar el cambio en la cantidad de copias de la secuencia de ácido nucleico a lo largo del tiempo, o determinar la concentración relativa de una secuencia cuando se compara con otra secuencia.

Prácticamente cualquier método convencional puede ser utilizado dentro del marco de la invención para detectar y cuantificar los niveles de ARNm o de su ADNc correspondiente. A modo ilustrativo, no limitativo, los niveles de ARNm codificados por dichos genes pueden ser cuantificados mediante el empleo de métodos convencionales, por ejemplo, métodos que comprenden la amplificación del ARNm y la cuantificación del producto de la amplificación de dicho ARNm, tales como electroforesis y tinción, o alternativamente, mediante Northern blot y empleo de sondas apropiadas y empleo de sondas específicas del ARNm de los genes de interés o de su ADNc correspondiente, mapeo con la nucleasa S1, RT-PCR, hibridación, microarrays, etc. Análogamente, los niveles del ADNc correspondiente a dichos ARNm también puede ser cuantificado mediante el empleo de técnicas convencionales; en este caso, el método de la invención incluye una etapa de síntesis del correspondiente ADNc mediante transcripción inversa (RT) del ARNm correspondiente seguida de amplificación y cuantificación del producto de la amplificación de dicho ADNc. Métodos convencionales de cuantificar los niveles de expresión pueden encontrarse, por ejemplo, en Sambrook y cois., 2001 “Molecular cloning: a Laboratory Manual”, 3 rd ed., Cold Spring Harbor Laboratory Press, N.Y., Vol.1-3.

En una realización preferida, la cuantificación del producto de la expresión de los genes se realiza determinando el nivel de mRNA derivado de su transcripción, donde el análisis del nivel de mRNA se puede realizar, a título ilustrativo y sin que limite el alcance de la invención, mediante amplificación por reacción en cadena de la polimerasa (PCR), retrotranscripción en combinación con la reacción en cadena de la ligasa (RT-LCR), retrotranscripción en combinación con la reacción en cadena de la polimerasa (RT-PCR), retrotranscripción en combinación con la reacción en cadena de la polimerasa cuantitativa (qRT-PCR), o cualquier otro método de amplificación de ácidos nucleicos; microarrays de DNA elaborados con oligonucleótidos depositados por cualquier mecanismo; microarrays de DNA elaborados con oligonucleótidos sintetizados *in situ* mediante fotolitografía o por cualquier otro mecanismo; hibridación *in situ* utilizando sondas específicas marcadas con cualquier método de marcaje; mediante geles de electroforesis; mediante transferencia a membrana e hibridación con una sonda específica; mediante RMN o cualquier otra técnica de

diagnóstico por imagen utilizando nanopartículas paramagnéticas o cualquier otro tipo de nanopartículas detectables funcionalizadas con anticuerpos o por cualquier otro medio.

Adicionalmente, el método de la invención puede incluir la realización de una etapa de extracción con el fin de obtener el ARN total, lo que puede realizarse mediante técnicas convencionales (Chomczynski y cois., Anal. Biochem., 1987, 162:156; Chomczynski P., Biotechniques, 1993, 15:532).

El término “muestra”, tal como se emplea en la presente invención, hace referencia a una muestra, obtenida partir del sujeto bajo estudio (salvo que se indique lo contrario), tal como una muestra de sangre o de suero obtenida a partir de dicho sujeto. En una realización particular, dicha muestra de sangre comprende sangre periférica. El término “sangre periférica” se relaciona con el volumen de sangre circulante distante del corazón, esto es, la sangre que circula por el organismo de un sujeto. La muestra de sangre puede obtenerse por métodos convencionales conocidos por el técnico en la materia. El término “suero”, tal como se utiliza en la presente invención, se refiere al componente de la sangre resultante tras la coagulación de ésta y eliminación del coágulo resultante. Métodos de obtención de muestras de sangre a partir de un sujeto están ampliamente recogidos en el estado de la técnica, así como métodos de obtención de suero a partir de muestras de sangre.

El término "RNA mensajero" o "mRNA" se refiere a un RNA que está sin intrones y que puede o no traducirse en un polipéptido.

El término “cDNA” se refiere a una secuencia de nucleótidos, complementaria de una secuencia de mRNA.

El término “isoforma” o "variante" se refiere a todos los tipos de ARN transcritos de un gen dado que, cuando se procesan colectivamente mediante empalme alternativo o procesamiento alternativo, codifican isoformas de proteínas plurales. El término "empalme alternativo" o “procesamiento alternativo”, se refiere a todos los tipos de procesamiento de ARN que conducen a la expresión de isoformas de proteínas plurales de un solo gen. Algunos genes, generalmente genes eucariotas, se transcriben primero como precursores largos de ARNm que luego se acortan mediante una serie de pasos de procesamiento para producir la molécula de ARNm madura. El empalme alternativo del pre-ARNm es un proceso postranscripcional que permite la producción de distintos ARNm a partir de un solo gen con el potencial de expandir la

estructura y diversidad de proteínas. El empalme alternativo también puede introducir o eliminar elementos reguladores para afectar la traducción, la localización o la estabilidad del ARNm. Más del 70% de los genes humanos pueden experimentar un empalme alternativo con muchos genes capaces de producir docenas e incluso cientos de isoformas diferentes. Uno de estos pasos es el empalme de ARN, en el que las secuencias de intrones se eliminan del precursor de ARNm. Una célula puede unir el transcrito primario de diferentes maneras, creando diferentes "variantes de empalme" y, por lo tanto, creando diferentes cadenas polipeptídicas del mismo gen o de la misma molécula de ARNm. Las variantes de empalme pueden incluir, por ejemplo, inserciones de exones, extensiones de exones, truncamientos de exones, deleciones de exones, alternativas en la región no traducida 5 'y alternativas en la región no traducida 3'. Las variantes de empalme son impredecibles, ya que la célula puede unir la transcripción primaria de diferentes maneras, creando diferentes cadenas polipeptídicas del mismo gen o de la misma molécula de ARNm.

EL término UBE2L3 hace referencia al gen UBE2L3, que se encuentra en el cromosoma 22 y codifica para la enzima E3 L3 que conjuga ubiquitina (Ubiquitin Conjugating Enzyme E2 L3). Al gen UBE2L3 se le conoce también por UBCH7, L-UBC, UbcM4, E2-F1 o UBCE7. El procesamiento alternativo del gen UBE2L3 da lugar a 5 isoformas o variantes de mRNA. Dichas isoformas se pueden definir por su correspondiente número de acceso en el NCBI y por el número y secuencias específicas de exones, según se indica en la siguiente tabla:

ISOFORMA	**POSICIÓN EN LA SECUENCIA**				
	Exón 1	**Exón 2**	**Exón 3**	**Exón 4**	**Exón 5**
1 NM_003347.3 24/02/2019	1-225 SEQ ID NO: 1	226-321 SEQ ID NO: 2	322-508 SEQ ID NO: 3	----	509-3028 SEQ ID NO: 5
2 NR_028436 24/02/2019	1-225 SEQ ID NO: 1	226-321 SEQ ID NO: 2	322-508 SEQ ID NO: 3	509-617 SEQ ID NO: 4	618-3137 SEQ ID NO: 5
3 NM_001256356.1 24/02/2019	1-225 SEQ ID NO: 1	----	226-412 SEQ ID NO: 3	----	413-2932 SEQ ID NO: 5
4 NM_001256355.1 24/02/2019	1-204 SEQ ID NO: 6	205-300 SEQ ID NO: 2	301-487 SEQ ID NO: 3	----	488-3007 SEQ ID NO: 5
5 NR_046082.1 24/02/2019	1-648 SEQ ID NO: 7	----	745-931 SEQ ID NO: 3	----	932-3451 SEQ ID NO: 5

Tabla 1. Isoformas del gen UBE2L3, secuencias, número de acceso y exones correspondientes.

Así, el término isoforma 2 del gen UBE2L3, según se utiliza en el contexto de la presente invención, se refiere a la secuencia identificada con el número de acceso en

el NCBI NR_028436 del 24 de febrero de 2019. Dicha secuencia corresponde a una isoforma no codificante del gen y comprende los exones 1 a 5 definidos por las secuencias SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4 y SEQ ID NO: 5, según se especifica en la tabla 1.

El mRNA total, se refiere a la suma de todas las isoformas del gen UBE2L3, es decir, a las isoformas definidas por los números de acceso del NCBI NM_003347.3, NR_028436, NM_001256356.1, NM_001256355.1 y NR_046082.1 del 24 de febrero de 2019.

Como ya se ha indicado anteriormente, niveles de mRNA pueden medirse y cuantificarse por diversos métodos bien conocidos por aquellos con destreza en la técnica, incluyendo el uso de equipos y reactivos comercialmente disponibles. La cuantificación de los niveles de expresión de del gen UBE2L3 puede realizarse a partir del ARN resultante de la transcripción de dichos genes (ARNm) o, alternativamente, a partir del ADN complementario (ADNc) de dichos genes. Por tanto, en una realización particular, la cuantificación de los niveles de expresión de del gen UBE2L3 comprende la cuantificación de las isoformas de dicho gen, o un fragmento de dichas isoformas, el ADN complementario (ADNc) de dichas isoformas, o un fragmento de dicho ADNc, o sus mezclas.

La determinación de los niveles de expresión del gen UBE2L3, ya sea mediante la cuantificación del mRNA o del cDNA de la isoforma 2 o mediante la cuantificación del mRNA o cDNA total, necesita ser comparada con los valores de referencia.

El término “valor de referencia” o “punto de corte” o “cantidad de referencia”, tal y como se utiliza en la presente descripción, se refiere a cualquier valor o rango de valores derivado de la cuantificación del producto de la expresión del gen UBE2L en una muestra biológica, o en una colección de muestras biológicas, procedentes de individuos que no padecen la enfermedad celiaca. La cantidad de referencia se ha de medir de la misma forma, y ser obtenida en el mismo tipo de muestra biológica aislada, que la cantidad medida en el sujeto analizado.

El término “comparación”, tal y como se utiliza en la descripción, se refiere pero no se limita, a la comparación de la cantidad del producto de expresión del gen UBE2L3 o de cualquiera de sus isoformas con un valor de referencia. La comparación descrita en el método de la presente invención puede ser realizada manualmente o asistida por ordenador.

En la presente invención el valor de referencia corresponde al valor medio de los niveles de expresión del gen UBE2L3 o de cualquiera de sus isoformas, medidos en una muestra de sujetos no celiacos. Una vez que se ha establecido este valor medio, se puede comparar el nivel de este marcador expresado en muestras de sujetos celiacos con este valor medio, y de esta manera ser asignado al nivel de expresión "reducido" o “elevado”.

Debido a la variabilidad entre sujetos (por ejemplo, aspectos referidos a la edad, raza, etc.) es muy difícil (si no prácticamente imposible) establecer valores de referencia absolutos de expresión del gen UBE2L3 o de cualquiera de sus isoformas. De esta manera, en una forma de realización particular, los valores de referencia para expresión "elevado" o "reducido" de la expresión del gen UBE2L3 o de cualquiera de sus isoformas se determinan calculando los percentiles por medios convencionales que implica ensayar un grupo de muestras aisladas de sujetos normales (es decir, gente sin la enfermedad celiaca) para los niveles de expresión de UBE2L3 o de cualquiera de sus isoformas. Los niveles "reducidos" de UBE2L3 o de cualquiera de sus isoformas se pueden entonces asignar, preferiblemente, a muestras en donde los niveles de expresión de UBE2L3 o de cualquiera de sus isoformas son iguales a o inferiores al percentil 50 en la población normal, incluyendo, por ejemplo, niveles de expresión iguales a o inferiores del percentil 60 en la población normal, iguales a o inferiores al percentil 70 en la población normal, iguales a o inferiores al percentil 80 en la población normal, iguales a o inferiores al percentil 90 en la población normal, e iguales a o inferiores al percentil 95 en la población normal. Los niveles "elevados" deUBE2L3 o de cualquiera de sus isoformas se pueden entonces asignar, preferiblemente, a muestras en donde los niveles de expresión de UBE2L3 o de cualquiera de sus isoformas son iguales a o superan el percentil 50 en la población normal, incluyendo, por ejemplo, niveles de expresión iguales a o en exceso al percentil 60 en la población normal, iguales a o en exceso al percentil 70 en la población normal, iguales a o en exceso al percentil 80 en la población normal, iguales a o en exceso al percentil 90 en la población normal, e iguales a o en exceso al percentil 95 en la población normal.

Así, en según el método de la invención, se puede diagnosticar la enfermedad celiaca en base a los siguientes indicadores:

- un nivel reducido del mRNA de la isoforma 2 del gen UBE2L3 con respecto a un valor de referencia;

- un nivel elevado del mRNA total del gen UBE2L3 con respecto a un valor de referencia; y
- un nivel elevado del contenido relativo de mRNA de la isoforma 2 del gen UBE2L3 con respecto al total del mRNA del gen UBE2L3 con respecto a un valor de referencia.

“Expresión relativa", como se utiliza en el contexto del método de la invención, es la cantidad de mRNA expresado de la isoforma 2 del gen UBE2L3 con respecto a la cantidad del mRNA total del gen UBE2L3, expresadas ambas cantidades en las mismas unidades en función del método utilizado para la cuantificación de dicho mRNA según las técnicas conocidas en el estado de la técnica, o, como ya se ha indicado anteriormente, del cDNA correspondiente. El valor de referencia en este caso se establece mediante el cálculo de la expresión relativa de la isoforma 2 del gen UBE2L3 con respecto al mRNA total de un grupo control que no padece enfermedad celiaca.

En una realización particular, el valor de referencia se selecciona de:

(i) el contenido en mRNA que codifica la isoforma 2 del gen UBE2L3 en un sujeto que no sufre enfermedad celiaca o el valor medio del contenido en mRNA que codifica la isoforma 2 del gen UBE2L3 en una población de sujetos que no sufren enfermedad celiaca,

(ii) el contenido en mRNA total del gen UBE2L3 en un sujeto que no sufre enfermedad celiaca o el valor medio del contenido en mRNA total del gen UBE2L3 en una población de sujetos que no sufren enfermedad celiaca y/o

(iii) el contenido relativo del mRNA que codifica la isoforma 2 del gen UBE2L3 con respecto al total de mRNA del gen UBE2L3 en un sujeto que no sufre enfermedad celiaca o el valor medio del contenido relativo del mRNA que codifica la isoforma 2 del gen UBE2L3 con respecto al total de mRNA del gen UBE2L3 en una población de sujetos que no sufren enfermedad celiaca.

En otra realización particular, el contenido relativo del mRNA que codifica la isoforma 2 del gen UBE2L3 con respecto al total de mRNA del gen UBE2L3 se determina mediante la fórmula (I):

$$SCORE = 2^{(\text{cantidad mRNA isoforma 2 UBE2L3} - \text{cantidad mRNA total UBE2L3})}$$

En una realización preferida los niveles de mRNA de la isoforma 2 del gen UBE2L3 o del mRNA total del gen UBE2L3 se normalizan frente al nivel de un gen de expresión constitutiva.

Para normalizar los valores de expresión el ARNm entre las diferentes muestras, es posible comparar los niveles de expresión del ARNm de interés en las muestras a ensayar con la expresión de un ARN control correspondiente a un gen de expresión constitutiva. Un "ARN control" como se usa aquí, se refiere a un ARN cuyos niveles de expresión no cambian o solo cambian en cantidades limitadas en células de sujetos celiacos en comparación con sujetos no celiacos. Preferiblemente, el ARN control es ARNm derivado de genes de mantenimiento y que codifica proteínas que se expresan de forma constitutiva y que llevan a cabo funciones celulares esenciales. Los ejemplos de genes de mantenimiento para su uso en la presente invención incluyen β-2-microglobulina, ubiquitina, proteína ribosómica de 18-S, ciclofilina, GAPDH y actina. En un forma de realización preferida, el ARN control es el ARNm de GAPDH. En una forma de realización la cuantificación de la expresión génica relativa se calcula según el método comparativo Ct usando GAPDH como control endógeno y controles de ARN comerciales como calibradores. Los resultados finales, se determinan según la fórmula 2-(ΔCt de la muestra-ΔCt del calibrador), donde los valores ΔCT del calibrador y la muestra se determinan sustrayendo el valor CT del gen diana del valor del gen de GAPDH.

En una realización particular, la muestra comprende células monocucleares de sangre periférica (CMSF).

Las células mononucleares de sangre periférica se caracterizan por poseer un único núcleo redondo. Estas células consisten en linfocitos (células T, células B, células NK) y monocitos. En los seres humanos, los linfocitos constituyen la mayoría de la población de CMSF, seguidos de los monocitos, y solo un pequeño porcentaje de células dendríticas.

Métodos de extracción de CMSF son bien conocidos en el estado de la técnica y forman parte de las técnicas habituales de laboratorios de investigación y clínicos. Por ejemplo, estas células se pueden extraer mediante centrifugación de gradientes de densidad (J. W. Mannhalter y otros, Clin. Immunol. Immunopathol. 1986 38, 390 - 397).

En una realización particular el mRNA que codifica la isoforma 2 del gen UBE2L3 se determina mediante el empleo de una sonda que hibrida de forma específica con una región del exón 4 del gen UBE2L3 y/o el mRNA total del gen UBE2L3 se determina mediante el empleo de una sonda que hibrida de forma específica con una región del exón 5 del gen UBE2L3.

Se define “sonda” como una secuencia de bases nitrogenadas de entre adenina (A), citosina (C), guanina (G), timina (T) o uracilo (U), dispuestas en un polímero lineal que es total o parcialmente complementaria (apareamientos del tipo Watson y Crick: A-T/U o G-C) a otra presente en la muestra a analizar. Tales sondas pueden ser de ácidos nucleicos (ADN, ARN), ácido nucleico peptídico (PNA), u otros ácidos nucleicos sintéticos capaces de hibridar con otros ácidos nucleicos mediante apareamientos del tipo de Watson y Crick.

Dichas sondas pueden ser o un fragmento de ADN amplificado por PCR, o un oligómero de bases nitrogenadas (sintético u obtenido por digestión enzimática u otro medio químico o físico de un polímero existente).

En una realización particular de la invención las sondas son oligonucleótidos sintéticos que llevan modificaciones en uno o en los dos extremos para facilitar la inmovilización a soportes sólidos.

Las sondas se diseñan y se construyen de tal forma que su secuencia de bases sea única para la variedad que representa, y de tal forma que las discrepancias (nucleótidos no complementarios) entre dos o más variantes suelen estar en las zonas centrales de los oligonucleótidos.

La longitud de los oligonucleótidos puede variar entre 5 o más bases, siendo preferentemente entre 11 y 30 nucleótidos la longitud más apropiada para discriminar entre dos secuencias muy similares pero no idénticas de ácidos nucleicos.

El término "hibridación específica" se refiere a la formación de híbridos entre una sonda de polinucleótido y un polinucleótido diana específico (por ejemplo el mRNA del gen UBE2L3 o cualquiera de sus isoformas o exones) en el que la sonda preferentemente hibrida con el polinucleótido diana específico y no hibrida sustancialmente con polinucleótidos que consisten en secuencias que no son sustancialmente idénticas al polinucleótido diana. Sin embargo, los expertos en la materia reconocerá que la longitud mínima de un polinucleótido requerido para una hibridación específica con un polinucleótido diana dependerá de varios factores, por

ejemplo: el contenido G / C, el posicionamiento de bases no coincidentes (si las hay) , el grado de singularidad de la secuencia en comparación con la población de polinucleótidos diana, y la naturaleza química del polinucleótido (por ejemplo, la cadena principal de metilfosfonato o fosforotiolato), entre otros.

En una realización particular las sondas presentan una identidad de al menos un 80%, más preferiblemente de al menos un 90%, aún más preferiblemente de al menos un 95%, aún mucho más preferiblemente de al menos un 98%, y particularmente de un 100%, con un fragmento de las secuencias complementarias a las secuencias definidas por los números de acceso del NCBI descritos en la tabla 1, así como de los exones definidos por las secuencias SEQ ID NO: 1-7

Los términos "complementario" y "sustancialmente complementario" se refieren a la hibridación o emparejamiento de bases entre dos nucleótidos o moléculas de ácido nucleico, como por ejemplo, entre las dos hebras de una molécula de DNA de doble cadena o entre un cebador de oligonucleótido y un sitio de unión a cebador en un ácido nucleico de cadena sencilla a secuenciar o amplificar. Los nucleótidos complementarios son, generalmente, A y T (o A y U), y C y G. Dos moléculas de RNA o de DNA de cadena sencilla se dice que son sustancialmente complementarias cuando los nucleótidos de una hebra, alineadas óptimamente y comparadas con inserciones o deleciones de nucleótidos apropiadas, se emparejan con al menos aproximadamente el 80% de los nucleótidos de la otra cadena, normalmente al menos aproximadamente del 90% al 95%, y más preferiblemente del 98 al 100%. La complementariedad sustancial existe cuando una hebra de RNA o DNA hibrida en condiciones de hibridación selectivas con su complementario. Normalmente, la hibridación selectiva se producirá cuando existe al menos un 65% de complementariedad sobre un tramo de al menos 14 a 25 nucleótidos preferiblemente al menos aproximadamente un 75%, y más preferiblemente al menos un 90% de complementariedad. Véase, por ejemplo, M. Kanehisa (1984) Nucleic Acids Res.12:203.

Como se muestra en la tabla 1 anteriormente mencionada, la isoforma 2 del gen UBE2L3 es la única de las isoformas del gen que contiene el exón 4. El exón 4, tal y como se utiliza en la presente invención se refiere a la secuencia comprendida entre las posiciones 509 y 617 de la secuencia de nucleótidos definida por el número de acceso en el NCBI NR_028436 del 24 de febrero de 2019 y que en la presente invención se ha nombrado como SEQ ID NO: 4.

aa acttcagcgt tcccaattat ggcttctctc agatccagcc ttgaagttct ttgacctcct caattcacaa cctgtagctg actttagcca cccacaagta cagaaaa (SEQ ID NO: 4)

Dicho exón 4 solamente se encuentra presente en la isoforma 2 del gen UBE2L3 de la secuencia definida por el número de acceso en el NCBI NR_028436 del 24 de febrero de 2019.

Tal como se utiliza en la presente invención, exón 5 está definido por la secuencia SEQ ID NO: 5 y corresponde al último exón de todas las isoformas del gen UBE2L3, tal y como se muestra en la tabla 1 de la presente descripción.

En otra realización particular, el mRNA total del gen UBE2L3 se determina mediante el empleo de una sonda que hibrida de forma específica con una región del exón 3 del gen UBE2L3. En este caso, el exón 3 viene definido por la secuencia SEQ ID NO: 3 según se especifica en la tabla 1. En otra realización, el mRNA total del gen UBE2L3 se determina mediante cualquier sonda que se una de manera específica a una región común a todas las isoformas del gen UBE2L3.

En una realización preferida, la cuantificación del mRNA que codifica la isoforma 2 del gen UBE2L3 se lleva a cabo mediante la cuantificación del cDNA cuya secuencia viene definida mediante el número de acceso del NCBI NR_028436.2 de la versión del 4 de diciembre de 2018.

En otra realización particular, la cuantificación del mRNA total del gen UBE2L3 se lleva a cabo mediante la cuantificación de la suma de los niveles de los mRNA de las isoformas 1 a 5 del gen UBE2L3 con números de acceso del NCBI NM_003347.3, NR_028436, NM_001256356.1, NM_001256355.1 y NR_046082.1 del 24 de febrero de 2019.

En otra realización preferida, la cuantificación del ARNm se realiza mediante un microarray de expresión.

Un “microarray” es una tecnología múltiplex que típicamente utiliza una serie ordenada de miles de sondas de ácido nucleico para hibridar, por ejemplo con una muestra de cDNA o cRNA bajo condiciones de hibridación astringentes. La hibridación de la sonda con el ácido nucleico que se quiere identificar genera una señal que se puede detectar y/o cuantificar. A modo de ejemplo, las sondas pueden estar marcadas con un fluoróforo, plata, o quimioluminescencia. La señal emitida al hibridar permite por lo

tanto determinar la abundancia relativa de secuencias del ácido nucleico que se quiere detectar. Las sondas se conectan a un soporte sólido por un enlace covalente con una matriz química (mediante epoxi-silano, amino-silano, lisina, poliacrilamida u otras). El término "soporte sólido" tal como se emplea en la presente invención se refiere a una gran variedad de materiales, por ejemplo, pero sin limitarse, intercambio de iones o resina adsorción, vidrio, plástico, látex, nylon, gel, ésteres de celulosa, esferas paramagnéticas, silicio, perlas microscópicas o la combinación de algunos de ellos. Diversos microarrays están comercialmente disponibles incluyendo aquellas fabricadas por ejemplo por Affymetrix, Inc. y Illumina, Inc.

Las "condiciones de hibridación astringentes" incluyen normalmente concentraciones de sal de al menos 1 M, más generalmente menos de alrededor de 500 mM, y preferiblemente menos de aproximadamente 200 mM. Las temperaturas de hibridación pueden ser tan bajas como 5 °C, pero son normalmente mayores de 22 ºC, más normalmente mayores e aproximadamente 30 °C, y preferiblemente por encima de aproximadamente 37 °C. Los fragmentos más largos pueden requerir temperaturas más altas de hibridación para la hibridación específica. Como otros factores pueden afectar a la astringencia de la hibridación, incluso la composición de las bases y longitud las cadenas complementarias, la presencia de disolventes orgánicos y el grado de desemparejamiento de bases, la combinación de parámetros es más importante que la medición absoluta de cualquier parámetro solo.

En el contexto del método de la presente invención, la detección del contenido de la isoforma 2 del gen UBE2L3 se puede llevar a cabo empleando cualquier sonda capaz de hibridar con una región del mRNA o cDNA del exón 4 (SEQ ID NO: 4). En otra forma preferida de realización, la detección del contenido total en ARNm del gen UBE2L3 se lleva a cabo empleando una sonda capaz de hibridar con una región del mRNA o cDNA del exón 5 (SEQ ID NO: 5), del exón 3 (SEQ ID NO: 3) o de cualquier región común a las 5 isoformas del gen UBE2L3. En una realización concreta, la sonda utilizada para el reconocimiento y cuantificación del exón 4 es ILMN_1796830 de *Illunina*. En otra realización particular, se utiliza la sonda ILMN_1677877 de *Illumina* para el reconocimiento y cuantificación del exón 5 del gen UBE2L3.

En otra forma preferida de realización, el método diagnóistico se aplica sobre un sujeto sospechoso de padecer enfermedad celiaca que sigue una dieta sustancialmente libre de gluten.

Una “dieta sin gluten (DSG)” se refiere a la eliminación de forma estricta de la alimentación todos los productos que contengan o se cocinen con trigo, centeno, cebada y avena, o cualquiera de sus variedades e híbridos (espelta, escanda, kamut, triticale...), y productos derivados, evitando contaminaciones inadvertidas y todo tipo de transgresiones dietéticas.

“Sustancialmente libre de gluten” se refiere en general a productos alimenticios y/o cualquier componente de los mismos que no contienen gluten y/o que contienen una cantidad de gluten aceptable para que una repartición gubernamental aplicable, un ente regulador de alimentos, un grupo de industrias, o similar, los rotule como “libres de gluten”. La Administración de Alimentos y Medicamentos de Estados Unidos de América (FDA) reconoce como productos alimenticios “libres de gluten” a aquellos que no tienen: (1) un ingrediente que es cualquier tipo de trigo, centeno, cebada o híbridos de estos granos; (2) un ingrediente derivado de estos granos y que no se ha procesado para eliminar el gluten; y (3) un ingrediente derivado de estos granos que se han procesado para eliminar el gluten, si deriva en que el producto alimenticio contenga 20 ppm o más de gluten. Otros países tales como, por ejemplo, Nueva Zelanda y Australia, permiten el rótulo de alimento “libre de gluten” en productos alimenticios que tienen menos de 3 ppm de gluten. Un producto alimenticio que es sustancialmente libre de gluten puede tener un contenido de gluten menor o igual a 20 partes por millón (ppm), que incluyen 15 ppm, 10 ppm, 5 ppm, 3 ppm, 1 ppm, 0,5 ppm, 0,1 ppm, 0,05 ppm, 0 ppm, o cualquier valor o rango entre dos cualquiera de estos valores (que incluyen los puntos extremo). Cualquier de los productos alimenticios, composiciones sólidas, composiciones particuladas, composiciones secas, o similares que se indica que son libres de gluten se reconocen por parte los expertos en el arte como que opcionalmente están sustancialmente libres de gluten.

<u>Kit de la invención</u>

En un aspecto adicional, la invención se refiere a un kit para la puesta en práctica de un método tal y como se define en las reivindicaciones 1 a 10 que comprende:

(i) Una sonda que hibrida específicamente con el exón 4 del gen UBE2L3,

(ii) Una sonda que hibrida específicamente con el exón 5 del gen UBE2L3 y, opcionalmente

(iii) Una sonda que hibrida de forma específica con un gen de expresión constitutiva

en donde los componentes (i) y (ii) constituyen al menos un 1% del total de las sondas presentes en el kit.

El kit se basa en el poder predictivo del método de diagnóstico de la presente invención. El valor de referencia de los niveles de expresión del gen UBE2L3 o de cualquiera de sus isoformas se puede determinar antes de llevar a cabo el procedimiento de la presente invención. Con la ayuda del kit, se puede calcular la expresión del gen UBE2L3 o de cualquiera de sus isoformas con respecto a las muestras de control. El control puede de esta manera estar comprendido también en el kit.

El kit de la invención más preferiblemente comprende los medios necesarios para cuantificar la expresión del gen UBE2L3 o cualquiera de sus isoformas y opcionalmente, para comparar la cantidad detectada con una cantidad de referencia. Dicho kit puede contener todos aquellos reactivos necesarios para analizar la cantidad del producto de expresión del gen UBE2L3 así como cualquiera de sus isoformas y cualquiera de los exones (definidos por las secuencias SEQ ID NO: 1 a 7) por medio de cualquiera de los métodos conocidos en estado de la técnica mencionados anteriormente en este documento. El kit además puede incluir, sin ningún tipo de limitación, tampones, agentes para prevenir la contaminación, inhibidores de la degradación del RNA, etc. Por otro lado el kit puede incluir todos los soportes y recipientes necesarios para su puesta en marcha y optimización. Preferiblemente, el kit comprende además las instrucciones para llevar a cabo el método de la invención.

Más preferiblemente, el kit o dispositivo de la invención comprende los cebadores y sondas obtenidos de las secuencias de la invención (Tabla 1). El kit puede contener la/s sonda/s y cebadores útiles para cuantificar la expresión de dicho gen, o cualquiera de los exones descritos en la tabla 1, así como las combinaciones de los mismos.

En realizaciones particulares, el kit se selecciona entre (a) un kit adecuado para la PCR, (b) un kit adecuado para la Transferencia Northern Blot y (c) un kit adecuado para los análisis de microarray Se pueden combinar cualquiera de dos o más de estas realizaciones, de tal manera que el kit puede comprender, por ejemplo, ambos de (a) y (c).

En el caso de (a) un kit adecuado para la PCR, esta PCR es normalmente la PCR cuantitativa en tiempo real PCR (RQ-PCR), una técnica de cuantificación de la expresión génica sensible y reproducible. En este caso se desea que el kit comprenda

adicionalmente cebadores y sondas y oligonucleótido(s) del kit. Estos reactivos pueden estar comprendidos opcionalmente en el kit.

Una Transferencia Northern implica el uso de electroforesis para separar las muestras de ARN por tamaño y la posterior detección con un(os) oligonucleótido(s) (sonda de hibridación) complementaria con (parte de) la secuencia diana del ARN de interés.

Es también posible que el(los) oligonucleótido(s) estén inmovilizados sobre una superficie (preferiblemente sólida). En una de sus realizaciones, el kit comprende un microarray. Un microarray de ARN o DNA es una matriz sobre un sustrato sólido (normalmente un porta de vidrio o una celda de una película fina de silicio) que evalúa grandes cantidades de diferentes ARN o DNA que son detectables mediante sondas específicas inmovilizadas sobre un sustrato sólido. Cada soporte sólido contiene una secuencia específica de ácido nucleico, normalmente una secuencia de ADN, como sondas (o indicadores). Aunque el número de sondas no está limitado de manera alguna, existe una realización preferida en la que el microarray se personaliza para los procedimientos de la invención. En una realización, dicho microarray personalizado comprende cincuenta sondas o menos, tal como treinta sondas o menos, incluyendo veinte sondas o menos. En kit puede contener tampones y soluciones de hibridación , potenciación , bloqueo, lavado y cualquier componente necesario para llevar a cabo el método de diagnóstico según la práctica común.

En una realización particular, el kit de la invención comprende adicionalmente uno o más componentes seleccionados del grupo formado por:

(i) Medios para la purificación de ARN a partir de una muestra de células,
(ii) Medios para la transcripción inversa de una preparación de ARNm y
(iii) Medios para amplificación de los exones 4 y 5 del gen UBE2L3.

El kit de la invención podría por tanto incluir cualquier medio conocido en el estado de la técnica para la purificación del ARN, tales como el aislamiento del ARN basado en la extracción con fenol, precipitación mediante el uso de soluciones de sal caotrópica o la adsorción en sílice entre otros, para la conversión del mRNA a cDNA o para la amplificación de los exones 4 y 5 del gen UBE2L3.

Método de diagnóstico y tratamiento

En otro aspecto, la invención se refiere a un método para el diagnóstico y tratamiento de un sujeto que comprende:

(i) diagnosticar en dicho sujeto la presencia de enfermedad celiaca mediante el método de acuerdo a las reivindicaciones 1 a 10 y

(ii) administrar al sujeto diagnosticado con enfermedad celiaca en la etapa (i) un tratamiento adecuado para dicha enfermedad

Formas preferidas de realización de la etapa (i) de diagnóstico del presente aspecto se han descrito en detalle en el contexto del método diagnóstico de la invención y son igualmente aplicables al método de diagnóstico y tratamiento.

La expresión “tratamiento adecuado”, en el contexto de la presente invención se refiere principalmente al seguimiento de una dieta libre de gluten, como se ha descrito previamente. Adicionalmente se podrían emplear tratamientos enzimáticos para la degradación y neutralización del gluten o vacunas específicas. Asimismo, se puede valorar el uso de medicamentos que ayuden a mantener la barrera intestinal. En caso de celiaquía refractaria, el tratamiento puede implicar el uso de esteroides.

En una realización particular, el tratamiento es una dieta libre de gluten.

La invención se describirá a través de los siguientes ejemplos que deben considerarse meramente ilustrativos y no limitativos del alcance de la invención.

EJEMPLOS

Métodos

Por lo general la utilidad clínica de los estudios de asociación de genoma completo (o GWAS, por sus siglas en inglés - *genome-wide association studies*) ha sido muy limitada, por lo que es necesario desarrollar estrategias que permitan identificar la variación verdaderamente importante y traducir esos hallazgos en aplicaciones. Por su parte, la Aleatorización Mendeliana (AM) es un método estadístico por el cual se pueden priorizar genes utilizando información de cómo estos se relacionan con polimorfismos genéticos o SNPs adyacentes y de cómo los SNPs se asocian a su vez a distintos fenotipos (Zhu Z *et al*., Nat Genet. 2016;48:481-7). Así, la AM permite

identificar genes que median la asociación SNP-enfermedad, dotándola de cierta causalidad y de un mecanismo plausible.

Con el fin de priorizar los genes causales implicados en las asociaciones identificadas en distintos estudios GWAS realizados en enfermedad celiaca (EC), se utilizó el software SMR para aplicar la AM (Zhu Z *et al*., Nat Genet. 2016;48:481-7). Más concretamente, se analizaron los resultados del GWAS más grande llevado a cabo en EC (9451 casos vs 16434 controles) (Dubois PC *et al*., Nat Genet. 2010;42:295-302) y se cruzaron con datos de expresión de genoma completo en relación con genotipo (*expression quantitative trait loci* o eQTLs) de la base de datos GTEx (GTEx Consortium. Genetic effects on gene expression across human tissues. Nature. 2017;550:204-13) (p<1e-5), provenientes de sangre (n=122) y de intestino delgado (n=369). También se aplicó el software SMR a los mismos datos de GWAS con la metilación del ADN como posible mediadora en las asociaciones, utilizando para ello la versión Lite de la base de datos de *methylation quantitative trait loci* (mQTL) de McRae et al (McRae A *et al*., Identification of 55,000 Replicated DNA Methylation QTL. bioRxiv 166710) (p<1e-5; n=1,366), disponible en la página web de SMR (https://cnsgenomics.com/software/smr/#DataResource). Por otro lado, se cruzaron los resultados obtenidos a partir de los dos experimentos independientes (SMR con eQTLs y mQTLs), con el fin de dar con SNPs asociados a enfermedad celiaca por medio de la expresión de genes con cambios en sus niveles de metilación. Después, se replicaron los resultados obtenidos con los datos de expresión en sangre por medio de los eQTLs de la base de datos CAGE (Lloyd-Jones LR *et al*., Am J Hum Genet. 2017; 100:228-237) (p<1e-5; n= 2,765), también disponibles en la web anteriormente mencionada.

Todos los análisis se llevaron a cabo con los parámetros por defecto del software SMR, incluyendo el descarte de señales de asociación derivadas del desequilibrio de ligamiento por medio del algoritmo Heidi (p<0.01) y un valor umbral de p empírico menor de 1e-4 para el test SMR, debido a la potencia limitada del estudio derivada sobre todo del tamaño muestral de los estudios eQTL/mQTL utilizados.

Finalmente, se interrogaron los candidatos seleccionados por medio de la priorización de genes vía AM en dos bases de datos totalmente independientes de expresión de genoma completo en células mononucleares de sangre periférica disponibles en GEO (https://www.ncbi.nlm.nih.gov/geo/): un análisis en la plataforma *Illumina HumanHT-12 V4.0 expression beadchip* en 17 pacientes de EC en dieta sin gluten y 20 controles sin

inflamación intestinal (GSE113469) (Sangineto M *et al*., PLoS One. 2018;13:e0197915), y un estudio en otro tipo de *microarray* de expresión-*Affymetrix Human Genome U133A Array* de 42 controles, 59 pacientes de Enfermedad de Crohn y 26 de Colitis Ulcerosa (GSE3365) (Burczynski ME *et al*., J Mol Diagn. 2006;8:51-61). Las comparaciones se llevaron a cabo mediante la aplicación de tests Mann-Whitney U y se construyeron curvas Característica Operativa del Receptor o *Receiver Operating Characteristic* (ROC), que utilizaron los niveles de expresión de los genes candidato y sus combinaciones como clasificadores binarios. Concretamente, la expresión relativa de *UBE2L3* en EC se calculó a partir de las unidades de intensidad del gen *house-keeping GAPDH* y de las sondas disponibles en el *array* de Illumina antes mencionado para *UBE2L3* tal y como se indica a continuación:

2^((UBE2L3.1 - UBE2L3.2)/GAPDH)

siendo UBE2L3.1=ILMN_1677877 (exón 5 común a todas las variantes génicas) y UBE2L3.2=ILMN_1796830 (exón 4 exclusivo de la variante no-codificante número 2).

Resultados y discusión

El análisis de AM con los datos de eQTLs de intestino delgado de GTEx no dio ningún resultado significativo, mientas que el mismo análisis con los datos de sangre de la misma base de datos resultó en la priorización de tres SNPs, asociados a EC por medio de la expresión de sus genes adyacentes *AHSA2*, *AC007278.2* y *UBE2L3* (**Tabla 2**).

Gene	topSNP	topSNP_chr	topSNP_bp	b_GWAS	p_GWAS	b_eQTL	p_eQTL	b_SMR	p_SMR
AHSA2	rs4672441	2	61638952	-0.126633	0.00000167	0.438683	9.51867E-38	-0.288666	7.63033E-06
AC007278.2	rs917997	2	103070568	-0.231905	5.97E-15	0.231665	1.48261E-15	-1.00104	2.4163E-08
UBE2L3	rs5754217	22	21939675	0.151862	0.000000692	0.162817	2.41949E-13	0.932718	4.09239E-05

Tabla 2: Resumen de los resultados obtenidos a partir del análisis de AM en el GWAS celiaco con datos de eQTLs de sangre de GTEx. Además de la localización cromosómica de cada SNP y cada gen, se muestra el tamaño del efecto de cada estudio utilizado (b), así como su valor de p (p).

Sin embargo, cuando se cruzaron estos resultados con los obtenidos a partir de la AM del GWAS celiaco con mQTLs, solamente se mantuvo el SNP rs5754217 situado en el primer intrón de *UBE2L3* en el cromosoma 22 (**Figura 1**). Además, este binomio SNP-gen volvió a pasar el test SMR cuando se utilizaron los eQTLs de sangre de CAGE en lugar de los de GTEx, de manera que se pudo replicar los resultados en un *dataset* mayor (b=0.165547, error estándar=0.0340045, p=1.13E-06).

Por otro lado, cuando se comparó la expresión en pacientes de EC en dieta sin gluten (n=17) con los controles (n=20) del *dataset* de expresión de *Illumina* en células mononucleares de sangre periférica, las dos sondas mostraron comportamientos opuestos pero muy significativos (p<0.0001) (**Figura 2A**). Además, la construcción de un *score* de expresión relativa (tal y como se determina en Métodos) dio lugar a una diferencia aún más grande entre EC y controles (**Figura 2B**).

Finalmente, se construyeron curvas ROC para valorar el potencial diagnóstico de la expresión de los diferentes exones de *UBE2L3* por separado y de la expresión relativa. Se observó que, si bien la expresión independiente de los exones es un clasificador eficiente (área bajo la curva o *area under the curve* – AUC=0.997 por sonda), la expresión relativa de *UBE2L3* tiene una sorprendente especificidad y sensibilidad del 100% y por tanto, un valor de AUC=1 (**Figura 3**). Además, esta capacidad de discriminación parece específica de EC ya que no fue observada en un *dataset* público e independiente de expresión en células mononucleares de sangre periférica de enfermedad de Crohn y de Colitis Ulcerosa para ninguna de las sondas localizadas en el candidato *UBE2L3*. No obstante, los valores de AUC resultantes fueron de 0,528-0,699 y 0,452-0,752 para Crohn y Colitis Ulcerosa, respectivamente.

El altísimo potencial predictivo del *score* de expresión relativa de *UBE2L3* lo sitúa en un lugar privilegiado de cara al desarrollo de un test diagnóstico en sangre periférica que podría eliminar las provocaciones con gluten del diagnóstico de la EC en ausencia de gluten de la dieta.

REIVINDICACIONES

1. Método *in vitro* para el diagnóstico de la enfermedad celiaca en un sujeto sospechoso de padecer enfermedad celiaca que comprende cuantificar en una muestra de dicho sujeto el contenido en:
   (i) el mRNA que codifica la isoforma 2 del gen UBE2L3 y
   (ii) el mRNA total del gen UBE2L3
   en donde un contenido relativo del mRNA que codifica la isoforma 2 del gen UBE2L3 con respecto al total de mRNA del gen UBE2L3 elevado con respecto a un valor de referencia es indicativo de que el sujeto sufre enfermedad celiaca.

2. Método según la reivindicación 1 en donde el valor de referencia se selecciona de:
   el contenido relativo del mRNA que codifica la isoforma 2 del gen UBE2L3 con respecto al total de mRNA del gen UBE2L3 en un sujeto que no sufre enfermedad celiaca o el valor medio del contenido relativo del mRNA que codifica la isoforma 2 del gen UBE2L3 con respecto al total de mRNA del gen UBE2L3 en una población de sujetos que no sufren enfermedad celiaca.

3. Método según la reivindicación 1 en donde el contenido relativo del mRNA que codifica la isoforma 2 del gen UBE2L3 con respecto al total de mRNA del gen UBE2L3 se determina mediante la fórmula (I):

$$SCORE = 2^{(\text{cantidad mRNA isoforma 2 UBE2L3} - \text{cantidad mRNA total UBE2L3})}$$

4. Método según cualquiera de las reivindicaciones 1 a 3 en donde los niveles de mRNA de la isoforma 2 del gen UBE2L3 o del mRNA total del gen UBE2L3 se normalizan frente al nivel de un gen de expresión constitutiva, en el que el gen de expresión constitutiva es GAPDH.

5. Método según cualquiera de las reivindicaciones 1 a 4, en el que la muestra en la que se cuantifica el mRNA comprende células mononucleares de sangre periférica.

6. Método según cualquiera de las reivindicaciones 1 a 5, en donde el mRNA que codifica la isoforma 2 del gen UBE2L3 se determina mediante el empleo de una sonda que hibrida de forma específica con una región del exón 4 del gen UBE2L3 y en donde el

mRNA total del gen UBE2L3 se determina mediante el empleo de una sonda que hibrida de forma específica con una región del exón 5 de las isoformas 1 a 5 del gen UBE2L3.

7. Método según cualquiera de las reivindicaciones 1 a 5 en donde la cuantificación del mRNA que codifica la isoforma 2 del gen UBE2L3 se lleva a cabo mediante cuantificación del cDNA cuya secuencia viene definida mediante número de acceso del NCBI NR_028436.2 de la versión del 4 de diciembre de 2018.

8. Método según cualquiera de las reivindicaciones 1 a 7, en la cuantificación del mRNA total del gen UBE2L3 se lleva a cabo mediante la cuantificación de la suma de los niveles de los mRNA de las isoformas 1 a 5 del gen UBE2L3 con números de acceso del NCBI NM_003347.3, NM_001256, NR_028436, NM_001256356.1, NM_001253655.4 y NR_046082.1 del 24 de febrero de 2019 .

9. Método según cualquiera de las reivindicaciones 1 a 4 en donde la cuantificación del ARNm se realiza mediante un microarray de expresión.

10. Método según cualquiera de las reivindicaciones 1 a 9 en donde el sujeto sospechoso de padecer enfermedad celiaca sigue una dieta sustancialmente libre de gluten.

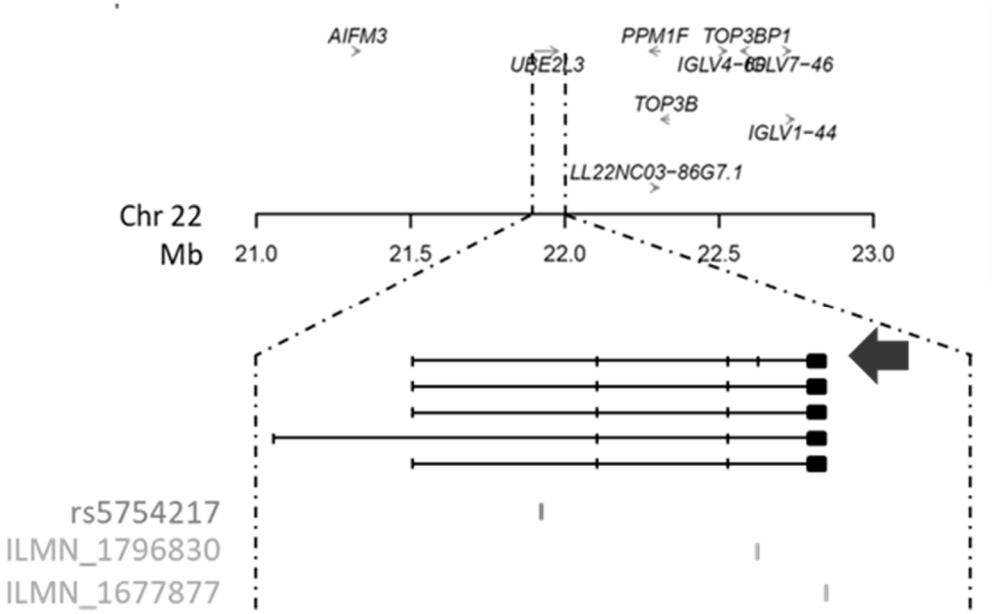
AIFM3
PPM1F
TOP3BP1
UBE2L3
IGLV4-69
IGLV7-46
TOP3B
IGLV1-44
LL22NC03-86G7.1
Chr 22
Mb
21.0
21.5
22.0
22.5
23.0
rs5754217
ILMN_1796830
ILMN_1677877

FIG. 1

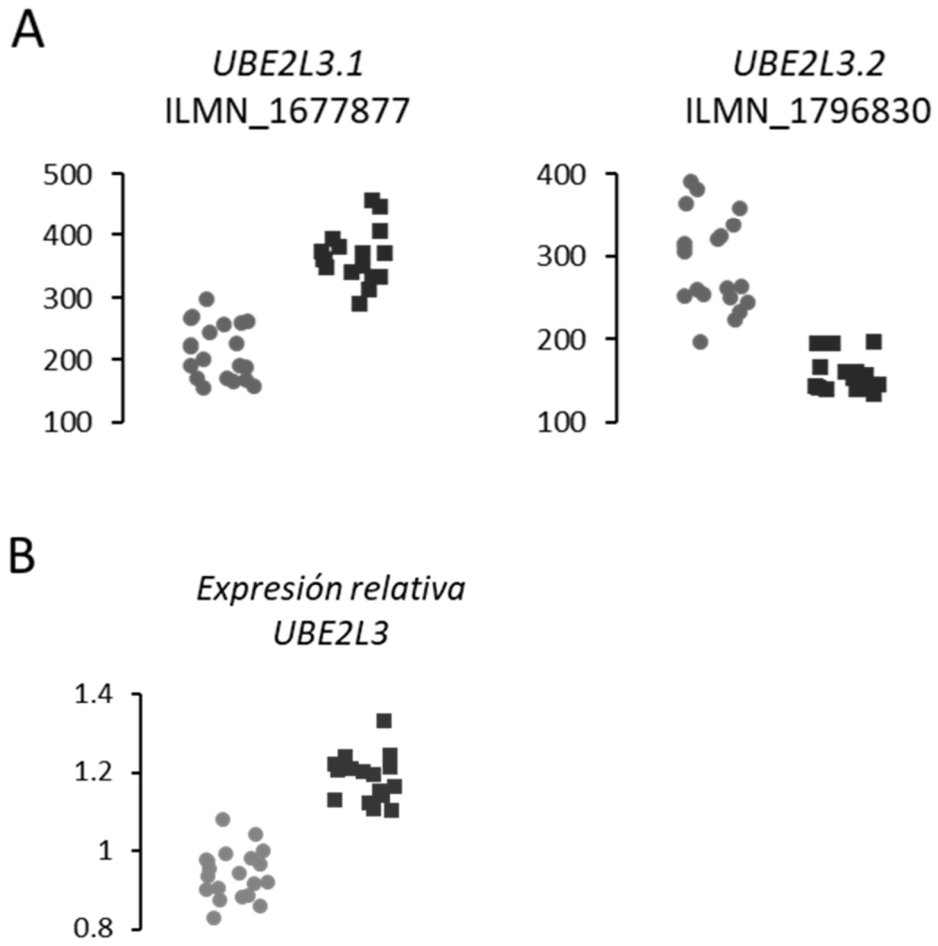
A
UBE2L3.1
ILMN_1677877
500
400
300
200
100
UBE2L3.2
ILMN_1796830
400
300
200
100
B
Expresión relativa
UBE2L3
1.4
1.2
1
0.8

FIG. 2

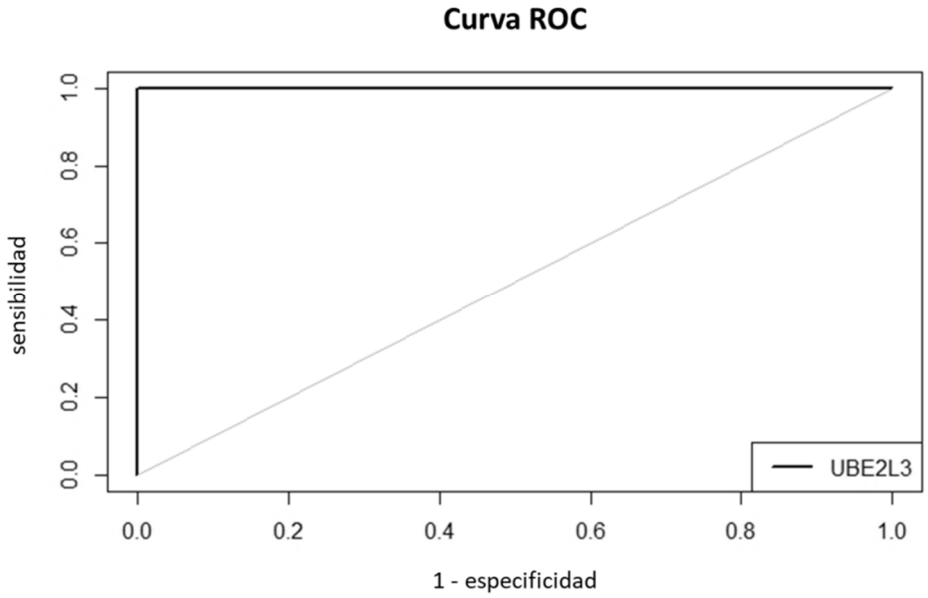
Curva ROC
sensibilidad
0.0
0.2
0.4
0.6
0.8
1.0
1 - especificidad
UBE2L3

FIG. 3