

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 636 671**

51 Int. Cl.:

C12Q 1/68 (2006.01)

G01N 33/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **26.01.2011** **PCT/US2011/022621**

87 Fecha y número de publicación internacional: **04.08.2011** **WO11094345**

96 Fecha de presentación y número de la solicitud europea: **26.01.2011** **E 11737602 (0)**

97 Fecha y número de publicación de la concesión europea: **24.05.2017** **EP 2529033**

54 Título: **Métodos para predicción del riesgo, diagnóstico, pronóstico de trastornos pulmonares**

30 Prioridad:

13.04.2010 US 323760 P

12.04.2010 US 323238 P

27.01.2010 US 298814 P

26.01.2010 US 298473 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
06.10.2017

73 Titular/es:

NATIONAL JEWISH HEALTH (100.0%)

1400 Jackson Street M206B

Denver, CO 80206, US

72 Inventor/es:

SCHWARTZ, DAVID y

SEIBOLD, MAX

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 636 671 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

Métodos para predicción del riesgo, diagnóstico, pronóstico de trastornos pulmonares

5 ANTECEDENTES DE LA INVENCION

Los trastornos de fibrosis pulmonar son una creciente preocupación en poblaciones humanas y no humanas. La fibrosis pulmonar está asociada con un número de trastornos complejos (por ejemplo Síndrome de Herman-Pudlak, esclerosis tuberosa, neurofibromatosis, y disqueratosis congénita). La neumonía intersticial idiopática (NII) representa una clase de trastorno fibrótico pulmonar crónico caracterizado por la cicatrización progresiva del intersticio alveolar que conduce a

10 disnea grave, hipoxemia, y muerte. La fibrosis pulmonar idiopática (FPI) es el tipo más común de NII y actualmente tiene la mortalidad más alta. A pesar de ser un área de investigación intensiva, la etiología de la FPI es en gran parte desconocida. La agrupación familiar de FPI y la susceptibilidad diferencial de los individuos a los polvos fibrogénicos ha implicado la genética en el desarrollo de este trastorno. Las variantes genéticas en los genes de la transcriptasa inversa de telomerasa (TERT), la proteína surfactante A1 y proteína surfactante C han sido implicados en el desarrollo de neumonía intersticial

15 familiar (NIF). Sin embargo, estas mutaciones solamente representan un pequeño porcentaje de casos de NIF. La asociación familiar con FPI es del 5-20%, y la herencia parece ser autosómica. La eficacia de los tratamientos actuales, tales como agentes fibrogénicos, es variable, indicando una necesidad de tratamiento más individualizado.

Las mucinas representan una familia de glicoproteínas asociadas con el epitelio de la mucosa. Las mucinas pueden estar asociadas con la membrana celular o ser secretadas, y típicamente forman un componente del moco. La expresión anormal de mutaciones en estas proteínas ha sido asociada con adenocarcinomas, así como trastornos pulmonares tales como asma y bronquitis.

Cheea et al describe los polimorfismos de MMP-1 y el riesgo de fibrosis pulmonar idiopática. El documento WO 2009/073167 describe la identificación y el diagnóstico de fibrosis pulmonar utilizando genes de mucina.

Los inventores actuales han encontrado que las variantes genéticas del gen MUC5B están asociadas con enfermedad pulmonar, y pueden proporcionar una herramienta útil para pronosticar el curso de la enfermedad y determinar un curso de

30 tratamiento. Además, el nivel incrementado de la expresión de MUC5B que resulta de las variantes genéticas descritas proporciona un objetivo terapéutico novedoso para enfermedades pulmonares tales como NII, FPI, y NIF.

BREVE SUMARIO DE LA INVENCION

Por consiguiente, en algunas realizaciones, la invención proporciona métodos y composiciones para el diagnóstico, predicción del riesgo, y pronóstico del curso de una afección fibrótica pulmonar. La divulgación proporciona además

35 métodos de tratamiento personalizados para enfermedades pulmonares.

En algunas realizaciones, la divulgación proporciona métodos para determinar si un sujeto tiene o está en riesgo de desarrollar una enfermedad pulmonar, comprendiendo dicho método determinar (detectar) si un sujeto expresa un nivel de ARN MUC5B elevado o un nivel de proteína MUC5B elevado con respecto a un control estándar (por ejemplo, normal), en

40 donde la presencia de dicho nivel de ARN MUC5B elevado o dicho nivel de proteína MUC5B elevado indica que dicho sujeto tiene o está en riesgo de desarrollar una enfermedad pulmonar. La enfermedad pulmonar es una afección fibrótica pulmonar.

El nivel de ARN o proteína MUC5B se puede determinar utilizando un ensayo in vitro o ensayo de formación de imágenes in vivo. Dicho nivel de proteína MUC5B elevado o dicho nivel de ARN MUC5B elevado se puede determinar a partir de una muestra biológica del sujeto, por ejemplo, un tejido pulmonar o fluido corporal de dicho sujeto. El fluido corporal puede ser, por ejemplo, sangre entera, plasma, suero, orina, esputo, saliva, una muestra de lavado broncoalveolar, o condensado de

45 aire exhalado. En algunas realizaciones, la muestra además se procesa, por ejemplo, para separar los componentes celulares o componentes subcelulares. Por ejemplo, la determinación puede además comprender separar las células de la muestra restante, o aislar los exosomas o las vesículas subcelulares.

El método puede además comprender administrar un tratamiento al sujeto, por ejemplo, un tratamiento de enfermedad pulmonar, o tratamiento de enfermedad pulmonar intersticial. El tratamiento puede ser un agente mucolítico. El tratamiento puede ser un antagonista de MUC5B. El método puede además comprender determinar un segundo nivel de ARN

50 MUC5B o nivel de proteína MUC5B después de administrar dicho tratamiento y comparar dicho segundo nivel con el nivel observado antes de administrar dicho tratamiento.

En algunas realizaciones, el nivel de expresión de al menos un marcador de enfermedad pulmonar adicional se determina y se compara con un control estándar. Por ejemplo, el al menos un marcador de enfermedad pulmonar adicional se puede seleccionar del grupo que consiste en Proteína Surfactante A, Proteína Surfactante D, KL-6/MUC1, CC16, CK-19, Ca 19-9, SLX, MCP-1, MIP-1a, ITAC, glutatión, péptido de procolágeno tipo III, sIL-2R, ACE, neopterina, beta-glucuronidasa, LDH, CCL-18, CCL-2, CXCL12, MMP7, y osteopontina. Un nivel de expresión aberrante del marcador de enfermedad pulmonar indica que el sujeto tiene o está en riesgo de desarrollar una enfermedad pulmonar. En algunas realizaciones, la expresión

60 aberrante es elevada con respecto a un control normal. En algunas realizaciones, la expresión aberrante es reducida con respecto a un control normal. En algunas realizaciones, el método comprende determinar si el genoma del sujeto comprende una variante genética del al menos un marcador de enfermedad pulmonar adicional seleccionado del grupo

65

que consiste en Proteína Surfactante A2, Proteína Surfactante B, Proteína Surfactante C, TERC, TERT, IL-1RN, IL-1 α , It-1 β , TNF, Linfotóxina α , TNF-RII, IL-10, IL-6, IL-12, IFN γ , TGF β , CR1, ACE, IL-8, CXCR1, CXCR2, MUC1 (KL6), y MUC5AC, en donde la presencia de una variante genética del al menos un marcador de enfermedad pulmonar adicional es indicativa de que el sujeto tiene o está en riesgo de desarrollar una enfermedad pulmonar. En algunas realizaciones, el método no comprende determinar si el genoma del sujeto comprende una variante genética de MUC5AC.

En algunas realizaciones, el control estándar se obtiene de muestra normal, no enferma. En algunas realizaciones, el control estándar es de un individuo diferente o grupo de individuos. En algunas realizaciones, el control estándar es un estándar obtenido de una población de individuos que no tienen una enfermedad pulmonar. En algunas realizaciones, el control estándar se obtiene del mismo individuo, por ejemplo, obtenido en un momento diferente, por ejemplo, antes de la exposición a un elemento estresante de las vías respiratorias. Típicamente, cuando se detecta o se determina el nivel de expresión de un ARN o proteína dada (por ejemplo, MUC5B), se detecta el mismo ARN o proteína en el control estándar. Sin embargo, en algunas realizaciones, se puede detectar un ARN o proteína diferente y la relación utilizada para determinar si el nivel de ARN o proteína del sujeto es elevado. Además, en algunas realizaciones, el método puede comprender la comparación con un control positivo, por ejemplo, de una muestra de enfermedad pulmonar conocida, o una muestra de un individuo o grupo de individuos que se sabe que portan un gen MUC5B de variante genética o tienen expresión de MUC5B elevada.

En algunas realizaciones, la invención proporciona métodos para determinar si un sujeto tiene o está en riesgo de desarrollar una afección fibrótica pulmonar, comprendiendo dicho método detectar en una muestra biológica de dicho sujeto si un genoma de dicho sujeto comprende un gen MUC5B de variante genética que es el alelo T de rs35705950, en donde la presencia de dicho gen MUC5B de variante genética que es el alelo T de rs35705950 indica que dicho sujeto tiene o está en riesgo de desarrollar una afección fibrótica pulmonar. La enfermedad pulmonar es una afección fibrótica pulmonar.

El gen MUC5B de variante genética que es el alelo T de rs35705950 en dicho sujeto da como resultado la expresión elevada de ARN MUC5B o proteína MUC5B. En algunas realizaciones, el sujeto es homocigoto para dicho gen MUC5B de variante genética que es el alelo T de rs35705950. En algunas realizaciones, el sujeto es heterocigoto para dicho gen MUC5B de variante genética que es el alelo T de rs35705950. En algunas realizaciones de la divulgación, el sujeto carece del gen MUC5B de variante genética. En algunas realizaciones de la divulgación, el gen MUC5B de variante genética es un gen MUC5B de región reguladora de variante genética, por ejemplo, un gen MUC5B promotor de variante genética. El gen MUC5B de variante genética tiene un polimorfismo de un solo nucleótido (SNP). El SNP es rs35705950. En algunas realizaciones, se determina la presencia de más de un SNP.

En algunas realizaciones, el gen MUC5B de variante genética comprende un primer polimorfismo de un solo nucleótido (SNP) y un segundo SNP. En algunas realizaciones, el primer SNP está presente dentro de una primera cadena de ADN y dicho segundo SNP está presente dentro de una segunda cadena de ADN. En algunas realizaciones, el primer y el segundo SNP están presentes dentro de la misma cadena de ADN.

En algunas realizaciones, la determinación comprende el uso de al menos una secuencia seleccionada del grupo que consiste en las SEC ID N^{os}:20-53, incluyendo la SEC ID N^o: 24, para determinar si el genoma del sujeto comprende un gen MUC5B de variante genética, por ejemplo, utilizando un ensayo de ácido nucleico apropiado para detectar el nucleótido variante en la secuencia seleccionada. Por ejemplo, la determinación puede comprender el uso de una o más de las secuencias de SEC ID N^{os}:20-53 en una RT-PCR, hibridación de arreglos, u otro método de detección de SNP apropiado como se describe en la presente memoria. En algunas realizaciones, la determinación comprende (i) poner en contacto una muestra del sujeto con una sonda de ácido nucleico que tiene al menos 10 nucleótidos contiguos de al menos una de las secuencias seleccionadas de las SEC ID N^{os}:20-53, incluyendo la SEC ID N^o: 24, o su complemento, donde dichos 10 nucleótidos contiguos abarcan el nucleótido de variante genética (i.e., la posición del SNP mostrada para cada secuencia), y (ii) determinar si la sonda de ácido nucleico se hibrida con un ácido nucleico en la muestra. La al menos una secuencia incluye la SEC ID N^o:24, en donde la presencia de una T en la posición 28 de la SEC ID N^o:24 indica un gen MUC5B de variante genética, y que el sujeto tiene o tendrá una forma atenuada de la enfermedad pulmonar. La presencia de una G en la posición 28 de la SEC ID N^o:24 indica que el sujeto tiene o tendrá una forma más grave de la enfermedad pulmonar (por ejemplo, donde el sujeto es homocigoto para G en la posición 28, o carece de un gen MUC5B promotor de variante genética).

En algunas realizaciones, el método comprende además determinar si dicho individuo expresa un nivel de ARN MUC5B elevado o un nivel de proteína MUC5B elevado con respecto a un control estándar, en donde la presencia de dicho nivel de ARN MUC5B elevado o dicho nivel de proteína MUC5B elevado indica además que dicho sujeto tiene o está en riesgo de desarrollar una enfermedad pulmonar. Dicho paso de determinación se puede llevar a cabo como se discutió anteriormente.

En algunas realizaciones, el método no comprende determinar si el individuo expresa un nivel elevado de ARN o proteína MUC5AC. En algunas realizaciones, el método no comprende determinar si dicho individuo expresa un nivel elevado de un segundo ARN o proteína distinto de un ARN o proteína MUC5B. En algunas realizaciones, el método no comprende determinar si dicho individuo expresa un nivel elevado de un segundo ARN o proteína distinta de un ARN o proteína MUC5B, a menos que dicho segundo ARN o proteína sea un ARN o proteína MUC5AC.

El método puede además comprender administrar un tratamiento al sujeto, por ejemplo, un tratamiento de enfermedad pulmonar, o tratamiento de enfermedad pulmonar intersticial. El tratamiento puede ser un agente mucolítico. El tratamiento puede ser un antagonista de MUC5B, por ejemplo, molécula pequeña que inhibe la producción o actividad de MUC5B. En algunas realizaciones, el método comprende además determinar un segundo nivel de ARN MUC5B o nivel de proteína MUC5B después de administrar dicho tratamiento y comparar dicho segundo nivel con el nivel observado antes de administrar dicho tratamiento.

En algunas realizaciones, el método comprende además determinar si el genoma del sujeto comprende al menos un gen marcador de enfermedad pulmonar de variante genética adicional. En algunas realizaciones, el al menos un marcador de enfermedad pulmonar adicional se puede seleccionar del grupo que consiste en Proteína Surfactante A2, Proteína Surfactante B, Proteína Surfactante C, TERC, TERT, IL-1RN, IL-1 α , IL-1 β , TNF, Linfotóxina α , TNF-RII, IL-10, IL-6, IL-12, IFN γ , TGF β , CR1, ACE, IL-8, CXCR1, CXCR2, MUC1 (KL6), o MUC5AC. La presencia de un gen marcador de enfermedad pulmonar de variante genética adicional puede indicar que el sujeto está en riesgo de o tiene una enfermedad pulmonar.

En algunas realizaciones, la presencia del gen MUC5B de variante genética que es el alelo de T de rs35705950 indica que el sujeto tiene una forma atenuada de la enfermedad pulmonar. Esto es, el sujeto tendrá una gravedad reducida de los síntomas, pérdida más gradual de la función pulmonar, o incremento de la supervivencia en comparación con la forma normal, no atenuada de la enfermedad pulmonar, i.e., en comparación con la enfermedad pulmonar como ésta ocurre en un individuo que no tiene un gen MUC5B de variante genética que es el alelo T de rs35705950.

Por tanto, en algunas realizaciones, la invención proporciona un método de pronóstico de una afección fibrótica pulmonar en un paciente, comprendiendo dicho método detectar en una muestra biológica de dicho paciente si un genoma de dicho paciente comprende un gen MUC5B de variante genética que es el alelo T de rs35705950, en donde la presencia de dicho gen MUC5B de variante genética que es el alelo T de rs35705950 indica un pronóstico menos grave de dicha afección fibrótica pulmonar en dicho paciente con respecto a la ausencia del alelo T del SNP en el gen MUC5B. La ausencia de un gen MUC5B de variante genética puede indicar que el paciente tiene una forma más agresiva de dicha enfermedad pulmonar. La enfermedad pulmonar es una afección fibrótica pulmonar.

En algunas realizaciones, el método comprende además establecer un curso de tratamiento para el sujeto, por ejemplo, en base a la presencia del gen MUC5B de variante genética en el sujeto. La presencia de un gen MUC5B de variante genética, o el nivel de la expresión del gen MUC5B, puede ser determinada en el sujeto, un tratamiento administrado al sujeto, y el progreso del sujeto monitorizado, por ejemplo, monitorizando la expresión de MUC5B con el tiempo u otros indicadores de diagnóstico pulmonar, y determinar si es necesario un tratamiento adicional. El método puede además comprender administrar tratamiento de enfermedad pulmonar o un tratamiento de enfermedad pulmonar intersticial al sujeto. El método puede además comprender determinar si el genoma del sujeto comprende un gen MUC5B de variante genética, en donde la presencia de un gen MUC5B de variante genética indica una forma atenuada de dicha enfermedad pulmonar intersticial en dicho sujeto.

En algunas realizaciones, la divulgación proporciona métodos para determinar si una enfermedad pulmonar está progresando en el paciente con enfermedad pulmonar, comprendiendo dicho método: (i) determinar un primer nivel de ARN MUC5B o primer nivel de proteína MUC5B en dicho paciente en un primer punto de tiempo; (ii) determinar un segundo nivel de ARN MUC5B o segundo nivel de proteína MUC5B en dicho paciente en un segundo punto de tiempo; y (iii) comparar el segundo nivel de ARN MUC5B con el primer nivel de ARN MUC5B o comparar el segundo nivel de proteína MUC5B con el primer nivel de proteína MUC5B, en donde si el segundo nivel de ARN MUC5B es mayor que el primer nivel de ARN MUC5B o si el primer nivel de proteína MUC5B es mayor que el primer nivel de proteína MUC5B, la enfermedad pulmonar está progresando en el paciente. La enfermedad pulmonar puede ser una enfermedad pulmonar intersticial, por ejemplo, una enfermedad pulmonar intersticial fibrótica, tal como fibrosis pulmonar idiopática o neumonía intersticial familiar.

En algunas realizaciones, el método comprende además determinar la velocidad de progresión en base a dicha comparación. Esto es, un rápido incremento en la expresión de MUC5B en un tiempo corto se correlaciona con una progresión más rápida de la enfermedad pulmonar. En algunas realizaciones, dicha determinación de dicho primer nivel de ARN MUC5B o primer nivel de proteína MUC5B y dicho segundo nivel de ARN MUC5B o segundo nivel de proteína MUC5B comprende normalizar dicho primer nivel de ARN MUC5B o primer nivel de proteína MUC5B y dicho segundo nivel de ARN MUC5B o segundo nivel de proteína MUC5B a un nivel de ARN o proteína expresado a partir de un gen estándar en dicho paciente con enfermedad pulmonar intersticial, por ejemplo, GAPDH, actina beta, HPRT1, tubulina beta, o beta-20 microglobulina.

En algunas realizaciones, la divulgación proporciona métodos para tratar, prevenir o mejorar una enfermedad pulmonar en un sujeto que lo necesite, comprendiendo el método administrar a dicho paciente una cantidad efectiva de un antagonista de MUC5B, en donde dicho antagonista reduce la expresión del gen MUC5B o reduce la actividad de la proteína MUC5B en comparación con la expresión o actividad en la ausencia de dicho antagonista de MUC5B, tratando, previniendo o mejorando de este modo la enfermedad pulmonar en el sujeto. El antagonista de MUC5B puede ser un ácido nucleico, por ejemplo, un pARN, ARNip, o secuencia antisentido, y reduce la expresión del gen MUC5B. El antagonista de MUC5B

puede ser una molécula pequeña, por ejemplo, que reduce la traducción de ARNm de MUC5B o el empaquetamiento o actividad de la proteína MUC5B. El antagonista de MUC5B se puede seleccionar del grupo que consiste en: un anticuerpo de MUC5B o fragmento de unión a MUC5B del mismo, un aptámero de unión a MUC5B, y un agente mucolítico. El ácido nucleico antagonista de MUC5B puede ser capaz de hibridarse al menos con una secuencia de 10 nucleótidos contiguos de una secuencia de ácidos nucleicos objetivo que codifica MUC5B. El método puede además comprender monitorizar el sujeto, por ejemplo, determinando el nivel de ARN o proteína MUC5B antes y después de dicha administración, o en uno o más puntos de tiempo después de dicha administración. Por tanto, el método de tratamiento puede incluir un paso de determinación de si el genoma del sujeto comprende un gen MUC5B de variante genética, y/o un paso de determinación de si el sujeto tiene un nivel elevado de ARN o proteína MUC5B, como se describe en la presente memoria.

En algunas realizaciones, la divulgación proporciona métodos de identificación de un compuesto de tratamiento de enfermedad pulmonar candidato, comprendiendo dicho método: (i) poner en contacto un compuesto de prueba con una proteína MUC5B; (ii) permitir que dicho compuesto de prueba inhiba la actividad de dicha proteína MUC5B; y (iii) seleccionar el compuesto de prueba que inhibe la actividad de dicha proteína MUC5B, identificando de este modo un compuesto de tratamiento de enfermedad pulmonar candidato. El método se puede llevar a cabo in vivo, por ejemplo, en un modelo animal para enfermedad pulmonar. El método se puede llevar a cabo in vitro.

En algunas realizaciones, la divulgación proporciona métodos de identificación de un compuesto de tratamiento de enfermedad pulmonar candidato, comprendiendo dicho método: (i) poner en contacto un compuesto de prueba con una célula secretora de MUC5B; (ii) permitir que dicho compuesto de prueba inhiba la secreción de proteína MUC5B de dicha célula secretora de MUC5B; y (iii) seleccionar el compuesto de prueba que inhibe la secreción de proteína MUC5B de dicha célula secretora de MUC5B, identificando de este modo un compuesto de tratamiento de enfermedad pulmonar candidato. Dicha célula secretora de MUC5B puede estar in vitro. Dicha célula secretora de MUC5B puede formar parte de un tejido pulmonar. Dicho tejido pulmonar puede formar parte de un organismo, i.e., el método se lleva a cabo in vivo. El organismo puede ser un mamífero, por ejemplo, un modelo animal o un humano.

La invención además proporciona equipos, por ejemplo, para determinar si un sujeto expresa un nivel elevado de ARN MUC5B o de proteína MUC5B, o porta un gen MUC5B de variante genética. El equipo puede comprender (a) un agente de unión a MUC5B capaz de unirse a una sustancia seleccionada del grupo que consiste en (i) una secuencia de gen MUC5B de variante genética; (ii) un ARN MUC5B o fragmento del mismo; y (iii) una proteína MUC5B o fragmento de la misma, y (b) un reactivo de detección o un aparato de detección capaz de indicar la unión de dicho agente de unión a MUC5B a dicha sustancia. El agente de unión a MUC5B puede ser marcado, por ejemplo, con un marcador fluorescente o radioisótopo. El equipo puede además comprender un dispositivo de recogida de muestras para recoger una muestra del sujeto. El agente de unión a MUC5B se puede unir a un gen de MUC5B de variante genética en la región promotora. El equipo puede además comprender al menos una muestra de control, por ejemplo, una secuencia de gen de MUC5B no variante o una muestra de un control normal, sin enfermedad.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

La Figura 1 representa un diagrama de flujo relacionado con el diseño de estudio genético descrito en la presente memoria.

La Figura 2 representa una gráfica de puntuación LOD Multipunto para el cribado de genoma completo (884 marcadores con una distancia de inter-marcador promedio de 4,2 centimorgans (CM)) en 82 familias con dos o más casos de NII.

La Figura 3 ilustra la representación gráfica del desequilibrio de enlace (LD) por pares para los SNP significativamente asociados con FPI o NIF mediante la prueba de asociación alélica en el cribado genético de mucinas formadoras de gel expresadas en el pulmón. Los valores de LD mostrados se calculan mediante la estadística r^2 para los sujetos con FPI de cribado genético de mucina ($n=492$). La gráfica multi-coloreada sobre la representación gráfica indica la ubicación aproximada de estos SNP dentro de la región de mucina formadora de gel. El SNP (rs35705950) del promotor de MUC5B altamente significativo y los valores de LD por pares correspondientes están resaltados en rojo. La región intergénica está abreviada como Int, y el Promotor de MUC5B está abreviado como Pr. Los patrones de LD fueron cualitativamente similares en los controles aunque en la mayoría de los casos el LD fue más débil entre los controles.

Las Figuras 4A-4C representan ilustraciones de la expresión del gen MUC5B en sujetos con FPI ($N=33$) y no afectados ($N=47$) estratificados por el genotipo de SNP (rs35705950) del promotor de MUC5B y el estado de hábito de fumar. A. La expresión del gen MUC5B entre sujetos no afectados y con FPI codificados con color en base a si los sujetos son de tipo silvestre (gris oscuro) o heterocigotos para el SNP del promotor de MUC5B (gris claro). B. Comparación de la expresión de MUC5B en sujetos no afectados, entre fumadores no afectados solamente, y entre no fumadores no afectados solamente, por el genotipo de SNP del promotor de MUC5B. C. Comparación de la expresión de MUC5B en todos los sujetos con FPI, entre los fumadores con FPI solamente, y entre los no fumadores con FPI solamente, por el genotipo de SNP del promotor de MUC5B. Las líneas representan las medias del grupo y la expresión de MUC5B se determina con respecto a la expresión de GAPDH.

Las Figuras 5A-5C representan la inmunohistoquímica de MUC5B de tejido no afectado y con FPI. Las secciones de tejido teñidas para la distribución de MUC5B tanto en el pulmón no afectado como con FPI muestran fuerte tinción citoplásmica específica dentro de las células columnares secretorias de los bronquios y los bronquiolos proximales más grandes (Fig. 5A). En sujetos con FPI, se observaron regiones de densa acumulación de MUC5B

en áreas de apanalamiento microscópicas e implicaron la tinción parcheada del revestimiento epitelial metaplástico de los quistes de panal (Fig. 5B), así como los tapones de moco dentro de los quistes (Fig. 5C).

DESCRIPCIÓN DETALLADA DE LA INVENCIÓN

Los inventores han encontrado que los individuos que portan variantes genéticas del gen MUC5B que tienen expresión elevada del gen tienen una probabilidad incrementada de desarrollar una enfermedad pulmonar, por ejemplo, una enfermedad pulmonar intersticial tal como enfermedad pulmonar intersticial fibrótica, fibrosis pulmonar idiopática, neumonía intersticial familiar, etc. La presencia de algunas variaciones genéticas en el gen MUC5B, aunque incrementa la probabilidad de una enfermedad pulmonar, son indicativas de una forma atenuada de la enfermedad, por ejemplo, una progresión más gradual de los síntomas y mejora de la supervivencia.

I. Definiciones

Los términos “enfermedad pulmonar”, “trastorno pulmonar”, “enfermedad de pulmón”, etc. se utilizan indistintamente en la presente memoria. El término se utiliza para referirse ampliamente a trastornos pulmonares caracterizados por dificultad para respirar, toser, incomodidad e inflamación de las vías respiratorias, aumento de moco y/o fibrosis pulmonar.

Las mucinas son una familia de proteínas de alto peso molecular, fuertemente glicosiladas (glicoproteínas) producidas por epitelios de mamífero. Las mucinas secretadas, formadoras de gel forman un componente del moco. Típicamente, los extremos terminales N- y C- de las proteínas de mucina son ligeramente glicosilados, pero ricos en residuos de cisteína formadores de enlaces de disulfuro.

La Mucina 5b (MUC5B) es una mucina formadora de gel expresada en el tejido epitelial de las vías respiratorias. Las mucinas formadoras de gel adicionales, MUC2, MUC5AC, y MUC6, se han mapeado a la misma región cromosómica en el cromosoma humano 11. MUC5B se caracteriza además en Desseyn et al. (1996) J. Biol Chem. 273:30157-64.

El término “variante genética”, en el contexto de un gen particular, se refiere a un gen con una secuencia de ácidos nucleicos variante (por ejemplo, no estándar o anormal). El gen incluye secuencias de codificación y de no codificación, tales como regiones reguladoras. Las variantes genéticas incluyen mutaciones y secuencias polimórficas. Por tanto, la variante genética puede afectar a la expresión o actividad del gen o producto génico. La variante genética puede ser una inserción de uno o más nucleótidos, supresión de uno o más nucleótidos, o una sustitución de uno o más nucleótidos. Un polimorfismo de un solo nucleótido (SNP) es un ejemplo de una variante genética.

El término “gen MUC5B de variante genética” se refiere a una variante genética de MUC5B (un gen MUC5B con una variación genética como se describió anteriormente). El término “gen MUC5B promotor de variante genética” se refiere a una variación que está específicamente en la región promotora del gen MUC5B. De manera similar, “gen MUC5B de región reguladora de variante genética,” y “gen MUC5B intrónico de variante genética” localizan la variación dentro del gen MUC5B. Un ejemplo de un gen MUC5B de variante genética es rs35705950, que incluye un SNP en la región promotora.

Una “muestra de mucosa de las vías respiratorias” se puede obtener utilizando métodos conocidos en la técnica, por ejemplo, un cepillo epitelial bronquial como se describe en la presente memoria. Los métodos adicionales incluyen biopsia endobronquial, lavado bronquial, lavado broncoalveolar, lavado pulmonar completo, biopsia transendoscópica y lavado transtraqueal.

Los términos “sujeto”, “paciente”, “individuo”, etc. no pretenden ser limitativos y pueden ser generalmente intercambiados. Esto es, un individuo descrito como un “paciente” no necesariamente tiene una enfermedad dada, sino que puede estar simplemente buscando consejo médico.

Una muestra o valor “de control” se refiere a una muestra que sirve como una referencia, normalmente una referencia conocida, para la comparación con una muestra de prueba. Por ejemplo, una muestra de prueba se puede tomar de un paciente que se sospecha que tiene una enfermedad pulmonar dada y comparada con muestras de un paciente que se sabe con enfermedad pulmonar, que se sabe portador de MUC5B de variante genética, o un individuo que se sabe normal (sin enfermedad). Un control también puede representar un valor promedio recogido de una población de individuos similares, por ejemplo, pacientes con enfermedad pulmonar o individuos sanos con antecedentes médicos similares, misma edad, peso, etc. También se puede obtener un valor de control del mismo individuo, por ejemplo, de una muestra obtenida anteriormente, antes de la enfermedad, o antes del tratamiento. Un experto reconocerá que los controles se pueden diseñar para la evaluación de cualquier número de parámetros.

Un experto en la técnica entenderá que los controles son valiosos en una situación dada y será capaz de analizar los datos en base a comparaciones con valores de control. Los controles también son valiosos para determinar el significado de los datos. Por ejemplo, si los valores para un parámetro dado son ampliamente variantes en los controles, la variación en las muestras de prueba no será considerada como significativa.

Como se utiliza en la presente memoria, los términos “farmacéuticamente” aceptable se utiliza de manera sinónima con fisiológicamente aceptable y farmacológicamente aceptable. Una composición farmacéutica comprenderá generalmente agentes para la regulación y la conservación en el almacenamiento, y puede incluir soluciones reguladoras y portadores para el suministro apropiado, dependiendo de la vía de administración.

Los términos “dosis” y “dosificación” se utilizan indistintamente en la presente memoria. Una dosis se refiere a la cantidad de ingrediente activo dado a un individuo en cada administración. Para la presente invención, la dosis se referirá generalmente a la cantidad de tratamiento de enfermedad pulmonar, agente antiinflamatorio, o antagonista de MUC5B. La dosis variará dependiendo de una serie de factores, incluyendo el intervalo de dosis normales para una terapia dada, frecuencia de administración; tamaño y tolerancia del individuo; gravedad de la afección; riesgo de efectos secundarios; y la vía de administración. Un experto reconocerá que la dosis se puede modificar dependiendo de los factores anteriores o en base al progreso terapéutico. El término “forma de dosificación” se refiere al formato particular del producto farmacéutico, y depende de la vía de administración. Por ejemplo, una forma de dosificación puede ser en una forma líquida para nebulización, por ejemplo, para inhalantes, en un comprimido o líquido, por ejemplo, para suministro oral, o una solución salina, por ejemplo, para inyección.

Como se utiliza en la presente memoria, los términos “tratar” y “prevenir” no pretenden ser términos absolutos. El tratamiento puede referirse a cualquier retraso en la aparición, reducción en la frecuencia o gravedad de los síntomas, mejora de los síntomas, mejora en la comodidad del paciente y/o la función respiratoria, etc. El efecto del tratamiento se puede comparar con un individuo o grupo de individuos que no reciben un tratamiento dado, o con el mismo paciente antes de, o después del cese del tratamiento.

El término “prevenir” se refiere a una disminución en la aparición de los síntomas de enfermedad pulmonar en un paciente. Como se indicó anteriormente, la prevención puede ser completa (sin síntomas detectables) o parcial, de tal manera que se observan menos síntomas de los que probablemente aparecerían en ausencia de tratamiento.

El término “cantidad terapéuticamente efectiva” como se utiliza en la presente memoria, se refiere a aquella cantidad del agente terapéutico suficiente para mejorar el trastorno, como se describió anteriormente. Por ejemplo, para el parámetro dado, una cantidad terapéuticamente efectiva mostrará un incremento o disminución de al menos el 5%, 10%, 15%, 20%, 25%, 40%, 50%, 60%, 75%, 80%, 90%, o al menos el 100%. La eficacia terapéutica también se puede expresar como incremento o disminución en “veces”. Por ejemplo, una cantidad terapéuticamente efectiva puede tener al menos 1,2 veces, 1,5 veces, 2 veces, 5 veces, o más efecto sobre un control.

El término “diagnóstico” se refiere a una probabilidad relativa de que una enfermedad pulmonar esté presente en el sujeto. De manera similar, el término “pronóstico” se refiere a una probabilidad relativa de que un cierto resultado futuro pueda ocurrir en el sujeto. Por ejemplo, en el contexto de la presente invención, pronóstico se puede referir a la probabilidad de que un individuo desarrollará una enfermedad pulmonar, o la gravedad probable de la enfermedad (por ejemplo, gravedad de los síntomas, velocidad de la declinación funcional, supervivencia, etc.). Los términos no pretenden ser absolutos, como será apreciado por cualquier experto en el campo del diagnóstico médico.

Los términos “correlación” y “asociado” en referencia a la determinación de un factor de riesgo de enfermedad pulmonar, se refiere a la comparación de la presencia o cantidad del factor de riesgo (por ejemplo, desregulación o variación genética en un gen de mucina) en un individuo con su presencia o cantidad en personas que se sabe que sufren de, o que se sabe que están en riesgo de, la enfermedad pulmonar, o en personas que se sabe que están libres de enfermedad pulmonar, y la asignación de una probabilidad incrementada o disminuida de tener/desarrollar la enfermedad pulmonar a un individuo en base a el/los resultado(s) del ensayo.

“Ácido nucleico” u “oligonucleótido” o “polinucleótido” o equivalentes gramaticales utilizados en la presente memoria significa al menos dos nucleótidos unidos covalentemente entre sí. Los oligonucleótidos son típicamente de aproximadamente 5, 6, 7, 8, 9, 10, 12, 15, 25, 30, 40, 50 o más nucleótidos de longitud, hasta aproximadamente 100 nucleótidos de longitud. Los ácidos nucleicos y polinucleótidos son unos polímeros de cualquier longitud, incluyendo longitudes más largas, por ejemplo, 200, 300, 500, 1000, 2000, 3000, 5000, 7000, 10,000, etc. Un ácido nucleico de la presente invención contendrá generalmente enlaces de fosfodiéster, aunque en algunos casos, se incluyen análogos de ácido nucleico que pueden tener cadenas principales alternativas, que comprenden, por ejemplo, enlaces de fosforamidato, fosforotioato, fosforoditioato, o O-metilfosforamidita (véase Eckstein, Oligonucleotides and Analogues: A Practical Approach, Oxford University Press); y cadenas principales y enlaces de ácido péptido nucleico. Otros ácidos nucleicos análogos incluyen aquellos con cadenas principales positivas; cadenas principales no iónicas y cadenas principales no de ribosa, incluyendo las descritas en las patentes de EE.UU. Nos. 5.235.033 y 5.034.506, y los Capítulos 6 y 7, ASC Symposium Series 580, Carbohydrate Modifications in Antisense Research, Sanghui & Cook, eds. Los ácidos nucleicos que contienen uno o más azúcares carbocíclicos se incluyen también dentro de una definición de ácidos nucleicos. Las modificaciones de la cadena principal de ribosa-fosfato se pueden hacer por una variedad de razones, por ejemplo, para incrementar la estabilidad y vida media de tales moléculas en ambientes fisiológicos o como sondas en un biochip. Se pueden hacer mezclas de ácidos nucleicos y análogos de origen natural; alternativamente, se pueden hacer mezclas de diferentes análogos de ácido nucleico, y mezclas de ácidos nucleicos y análogos de origen natural.

Los términos “idéntico” o porcentaje de “identidad”, en el contexto de dos o más ácidos nucleicos (por ejemplo, secuencias o subsecuencias genómicas, tal como se muestra en las SEC ID N^{os} :20-53, o secuencias de codificación) o secuencias de polipéptido, se refieren a dos o más secuencias o subsecuencias que son las mismas o tienen un porcentaje especificado de residuos de aminoácido o nucleótidos que son los mismos (i.e., 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o más identidad sobre una región especificada), cuando se

comparan y se alinean para correspondencia máxima sobre una ventana de comparación, o región designada medida utilizando uno de los siguientes algoritmos de comparación de secuencia o mediante alineación manual e inspección visual. Tales secuencias se dice entonces que son “sustancialmente idénticas”. Esta definición también se refiere al complemento de una secuencia de prueba. Opcionalmente, la identidad existe sobre una región que es al menos de aproximadamente 10 a aproximadamente 100, de aproximadamente 20 a aproximadamente 75, de aproximadamente 30 a aproximadamente 50 aminoácidos o nucleótidos de longitud.

Un ejemplo de algoritmos adecuados para determinar el porcentaje de identidad de secuencia y la similitud de secuencia son los algoritmos BLAST y BLAST 2.0, que se describen en Altschul et al., Nuc. Acids Res. 25 25:3389-3402 (1977) y Altschul et al., J. Mol. Biol. 215:403-410 (1990), respectivamente. Como será apreciado por un experto en la técnica, el software para realizar los análisis BLAST está públicamente disponible a través del sitio web del Centro Nacional para la Información Biotecnológica (ncbi.nlm.nih.gov).

Los términos “polipéptido”, “péptido” y “proteína” se utilizan indistintamente en la presente para referirse a un polímero de residuos de aminoácido. Los términos aplican a polímeros de aminoácido en los que uno o más residuos de aminoácido son un mimético químico artificial de un aminoácido de origen natural correspondiente, así como a polímeros de aminoácido de origen natural, aquéllos que contienen residuos modificados, y polímero de aminoácido de origen no natural.

El término “aminoácido” se refiere a aminoácidos de origen natural y sintéticos, así como análogos de aminoácido y miméticos de aminoácido que funcionan de manera similar a los aminoácidos de origen natural. Los aminoácidos de origen natural son aquéllos codificados por el código genético, así como aquellos aminoácidos que son modificados posteriormente, por ejemplo, hidroxiprolina, γ -carboxiglutamato, y O-fosfoserina. Análogos de aminoácido se refiere a compuestos que tienen la misma estructura química básica que un aminoácido de origen natural, por ejemplo, un carbono α que está unido a un hidrógeno, un grupo carboxilo, un grupo amino, y un grupo R, por ejemplo, homoserina, norleucina, sulfóxido de metionina, metionina metil sulfonio. Tales análogos pueden tener grupos R modificados (por ejemplo, norleucina) o cadenas principales de péptido modificado, pero conservan la misma estructura química básica que un aminoácido de origen natural. Los miméticos de aminoácido se refieren a compuestos químicos que tienen una estructura que es diferente de la estructura química general de un aminoácido, pero que funciona de manera similar a un aminoácido de origen natural.

Los aminoácidos pueden ser referidos en la presente memoria ya sea por sus símbolos de tres letras comúnmente conocidos o por los símbolos de una letra recomendados por la Comisión de Nomenclatura Bioquímica IUPAC-IUB. Los nucleótidos, asimismo, pueden ser referidos por sus códigos de una sola letra comúnmente aceptados.

“Variantes modificadas conservativamente” aplica tanto a secuencias de aminoácidos como de ácidos nucleicos. Con respecto a las secuencias de ácidos nucleicos particulares, variantes modificadas conservativamente se refiere a aquellos ácidos nucleicos que codifican secuencias de aminoácidos idénticas o esencialmente idénticas, o donde el ácido nucleico no codifica una secuencia de aminoácidos, a secuencias esencialmente idénticas o asociadas, por ejemplo, naturalmente contiguas. Debido a la degeneración del código genético, un gran número de ácidos nucleicos funcionalmente idénticos codifica la mayoría de las proteínas. Por ejemplo, los codones GCA, GCC, GCG y GCU todos codifican el aminoácido alanina. Por tanto, en cada posición donde una alanina es especificada por un codón, el codón se puede alterar a otro de los codones correspondientes descritos sin alterar el polipéptido codificado. Tales variaciones de ácido nucleico son “variaciones silenciosas”, que son una especie de variaciones modificadas conservativamente. Cada secuencia de ácidos nucleicos en la presente memoria que codifica un polipéptido también describe variaciones silenciosas del ácido nucleico. Un experto reconocerá que en ciertos contextos cada codón en un ácido nucleico (excepto AUG, que es normalmente el único codón para metionina, y TGG, que es normalmente el único codón para triptófano) se puede modificar para producir una molécula funcionalmente idéntica. Por consiguiente, a menudo las variaciones silenciosas de un ácido nucleico que codifica un polipéptido están implícitas en una secuencia descrita con respecto al producto de expresión, pero no con respecto a las secuencias de sonda reales.

En cuanto a las secuencias de aminoácidos, un experto reconocerá que las sustituciones, supresiones o adiciones individuales a un ácido nucleico, péptido, polipéptido, o secuencia de proteínas que altera, añade o suprime un único aminoácido o un pequeño porcentaje de aminoácidos en la secuencia codificada es una “variante modificada conservativamente” donde la alteración da como resultado la sustitución de un aminoácido con un aminoácido químicamente similar. Las tablas de sustitución conservativa que proporcionan aminoácidos funcionalmente similares son bien conocidas en la técnica. Tales variantes modificadas conservativamente son adicionales a y no excluyen las variantes polimórficas, homólogos de interespecie, y alelos de la invención. Sustituciones típicamente conservativas entre sí: 1) Alanina (A), Glicina (G); (2) Ácido aspártico (D), Ácido glutámico (E); 3) Asparagina (N), Glutamina (Q); 4) Arginina (R), Lisina (K); 5) Isoleucina (I), Leucina (L), Metionina (M), Valina (V); 6) Fenilalanina (F), Tirosina (Y), Triptófano (W); 7) Serina (S), Treonina (T); y 8) Cisteína (C), Metionina (M) (véase por ejemplo, Creighton, Proteins (1984)).

Un “marcador” o una “porción detectable” es una composición detectable mediante medios espectroscópicos, fotoquímicos, bioquímicos, inmunoquímicos, químicos u otros medios físicos. Por ejemplo, los marcadores útiles incluyen 32P, colorantes fluorescentes, reactivos electrodenso, enzimas (por ejemplo, como se utiliza comúnmente en una ELISA), biotina, digoxigenina o haptenos y proteínas u otras entidades que se pueden hacer detectables, por ejemplo,

incorporando un radiomarcador en un péptido o anticuerpo específicamente reactivo con un péptido objetivo. Se puede emplear cualquier método conocido en la técnica para conjugar un anticuerpo al marcador, por ejemplo, utilizando los métodos descritos en Hermanson, *Bioconjugate Techniques* 1996, Academic Press, Inc., San Diego.

Una "sonda de ácido nucleico u oligonucleótido marcado" es uno que está unido, ya sea covalentemente, a través de un enlazador o una unión química, o no covalentemente, a través de uniones iónicas, de Van der Waals, electrostáticas o de hidrógeno a un marcador tal que la presencia de la sonda se puede detectar detectando la presencia del marcador unido a la sonda. Alternativamente, el método que utiliza interacciones de alta afinidad puede lograr los mismos resultados donde uno de un par de parejas de unión se une al otro, por ejemplo, biotina, estreptavidina.

La frase "se hibrida selectivamente (o específicamente) con" se refiere a la unión, duplexación, o hibridación de una molécula solamente con una secuencia de nucleótidos particular con una mayor afinidad, por ejemplo, bajo condiciones más rigurosas, que con otras secuencias de nucleótidos (por ejemplo, DNA o ARN celular total o de biblioteca). Un experto en la técnica apreciará que la hibridación específica entre nucleótidos normalmente se basa en la unión de pares de Watson-Crick entre secuencias de nucleótidos complementarias.

El término "sonda" o "cebador", como se utiliza en la presente memoria, se define como uno o más fragmentos de ácido nucleico cuya hibridación específica con una muestra puede ser detectada. Una sonda o cebador puede ser de cualquier longitud dependiendo de la técnica particular para la que será utilizado. Por ejemplo, los cebadores de PCR son generalmente de entre 10 y 40 nucleótidos de longitud, mientras que las sondas de ácidos nucleicos para, por ejemplo, una transferencia de Southern, pueden ser de más de cien nucleótidos de longitud. La sonda puede estar sin marcar o marcada como se describe más adelante de modo que se puede detectar su unión al objetivo o muestra. La sonda se puede producir a partir de una fuente de ácidos nucleicos de una o más porciones (preseleccionadas) particulares de un cromosoma, por ejemplo, uno o más clones, un cromosoma completo o fragmento de cromosoma aislado, o una colección de productos de amplificación de reacción en cadena de la polimerasa (PCR). La longitud y la complejidad del ácido nucleico fijado sobre el elemento objetivo no es crítica para la invención. Un experto puede ajustar estos factores para proporcionar hibridación óptima y producción de señal para un procedimiento de hibridación dado, y para proporcionar la resolución requerida entre diferentes genes o ubicaciones genómicas.

La sonda también puede ser ácidos nucleicos aislados inmovilizados sobre una superficie sólida (por ejemplo, nitrocelulosa, vidrio, cuarzo, láminas de sílice fundida), como en un arreglo. La sonda puede ser un miembro de un arreglo de ácidos nucleicos como se describe, por ejemplo, en el documento WO 96/17958. Las técnicas capaces de producir arreglos de alta densidad también se pueden utilizar para este propósito (véase, por ejemplo, Fodor (1991) *Science* 767-773; Johnston (1998) *Curr. Biol.* 8: R171-R174; Schummer (1997) *Biotechniques* 23:1087-1092; Kern (1997) *Biotechniques* 23: 120-124; patente de EE.UU. N° 5.143.854). Un experto reconocerá que la secuencia precisa de las sondas particulares descritas en la presente memoria se puede modificar hasta un cierto grado para producir sondas que son "sustancialmente idénticas" a las sondas descritas, pero conservan la capacidad de unirse específicamente a (i.e., hibridarse específicamente con) los mismos objetivos o muestras que la sonda de la que se derivaron. Tales modificaciones están cubiertas específicamente por referencia a las sondas individuales descritas en la presente memoria.

"Anticuerpo" se refiere a un polipéptido que comprende una región estructural de un gen de inmunoglobulina o fragmentos del mismo que específicamente se une y reconoce un antígeno, por ejemplo, un antígeno bacteriano específico. Típicamente, la "región variable" contiene la región de unión al antígeno del anticuerpo (o su equivalente funcional) y es más crítico en cuanto a especificidad y afinidad de unión. Véase Paul, *Fundamental Immunology* (2003).

Una unidad estructural de inmunoglobulina (anticuerpo) ejemplar comprende un tetrámero. Cada tetrámero se compone de dos pares idénticos de cadenas polipeptídicas, teniendo cada par una cadena "ligera" (de aproximadamente 25 kD) y una "pesada" (de aproximadamente 50-70 kD). El N-terminal de cada cadena define una región variable de aproximadamente 100 a 110 o más aminoácidos responsables principalmente del reconocimiento del antígeno. Los términos cadena ligera variable (VL) y cadena pesada variable (VH) se refieren a estas cadenas ligera y pesada respectivamente.

Los anticuerpos pueden existir como inmunoglobulinas intactas o como cualquiera de una serie de fragmentos bien caracterizados que incluyen actividad de unión al antígeno específica. Tales fragmentos se pueden producir mediante digestión con diversas peptidasas. La pepsina digiere un anticuerpo debajo de los enlaces disulfuro en la región bisagra para producir $F(ab)_2$; un dímero de Fab que en sí mismo es una cadena ligera unida a VH-CH 1 por un enlace disulfuro. El $F(ab)_2$ se puede reducir bajo condiciones suaves para romper en enlace disulfuro en la región bisagra, convirtiendo así el dímero $F(ab)_2$ en un monómero Fab'. El monómero Fab' es esencialmente Fab con parte de la región bisagra (véase *Fundamental Immunology* (Paul ed., 3ª ed. 1993). Mientras que diversos fragmentos de anticuerpo se definen en términos de la digestión de un anticuerpo intacto, un experto apreciará que tales fragmentos pueden ser sintetizados de nuevo ya sea químicamente o utilizando metodología de ADN recombinante. Por tanto, el término anticuerpo, como se utiliza en la presente memoria, también incluye fragmentos de anticuerpo ya sea producidos por la modificación de anticuerpos completos, o aquellos sintetizados de nuevo utilizando metodologías de ADN recombinante (por ejemplo, Fv de cadena sencilla) o aquellos identificados utilizando bibliotecas de expresión en fago (véase, por ejemplo, McCafferty et al., *Nature* 348:552-554 (1990)).

II. Mucinas

Hay varias mucinas formadoras de gel que incluyen, pero no se limitan a, MUC6, MUC2, MUC5AC, y MUC5B. Estas proteínas son grandes filamentosas y altamente O-glicosiladas.

III. Enfermedades Pulmonares

Las enfermedades pulmonares contempladas en la presente memoria pueden incluir cualquier trastorno pulmonar, enfermedad de fibrosis pulmonar, enfermedad pulmonar intersticial, neumonía intersticial idiopática (NII), fibrosis pulmonar idiopática, neumonía intersticial familiar (NIF), síndrome de dificultad respiratoria aguda (SDRA), enfermedad pulmonar de esclerodermia, Sarcoidosis, Enfermedad de Berilio, trastorno pulmonar asociado a artritis reumatoide, trastorno pulmonar asociado a vascular de colágeno, trastorno pulmonar asociado a humo de cigarrillo, síndrome de Sjögren, enfermedad mixta del tejido conectivo, neumonía intersticial no específica (NINE), etc.

Las condiciones fibróticas pulmonares, por ejemplo, enfermedades pulmonares intersticiales (ILD) se caracterizan por dificultad para respirar, tos crónica, fatiga y debilidad, pérdida de apetito y rápida pérdida de peso. La fibrosis pulmonar esta comúnmente vinculada a enfermedades pulmonares intersticiales (por ejemplo, trastornos autoinmunes, infecciones virales u otras lesiones microscópicas), pero puede ser idiopática. La fibrosis implica el intercambio de tejido pulmonar normal con tejido fibrótico (tejido cicatricial) que conduce a capacidad de oxígeno reducida.

Las neumonías intersticiales idiopáticas (NII) son un subconjunto de enfermedades pulmonares intersticiales difusas de etiología desconocida (el término "idiopático" indica origen desconocido). Las NII se caracterizan por la expansión del compartimento intersticial (i.e., aquella porción del parénquima pulmonar intercalado entre las membranas basales epitelial y endotelial) con un infiltrado de células inflamatorias. El infiltrado inflamatorio está algunas veces acompañado por fibrosis, ya sea en forma de deposición de colágeno anormal o proliferación de fibroblastos capaces de la síntesis de colágeno.

La Fibrosis Pulmonar Idiopática (FPI) ocurre en miles de personas en todo el mundo con una duplicación de prevalencia durante los últimos 10 años. La aparición de FPI ocurre alrededor de los 50 a 70 años de edad y comienza con progresiva dificultad para respirar e hipoxemia. La supervivencia media de FPI es de alrededor de 3-5 años y es hasta la fecha intratable. La etiología y patogénesis de la afección no se entiende bien. Aproximadamente el 5-20 por ciento de todos los casos de FPI tienen antecedentes familiares y la herencia parece ser autosómica dominante.

Las enfermedades pulmonares fibróticas adicionales incluyen neumonía intersticial aguda (AIP), Bronquiolitis Respiratoria asociada a Enfermedad Intersticial Pulmonar (RBILD), Neumonía Intersticial Descamativa (NID), Neumonía Intersticial No Específica (NINE), Bronquiolitis obliterante, con Neumonía Organizativa (BONO).

La AIP es una forma rápidamente progresiva e histológicamente distinta de neumonía intersticial. El patrón patológico es una forma de organización de daño alveolar difuso (DAD) que también se encuentra en el síndrome de dificultad respiratoria aguda (SDRA) y otras neumonías intersticiales agudas de causas conocidas (véase Clinical Atlas of Interstitial Lung Disease (2006 ed.) pp61-63).

La RBILD se caracteriza por lesiones inflamatorias de los bronquiolos respiratorios en fumadores de cigarrillos. El aspecto histológico de RBILD se caracteriza por la acumulación de macrófagos pigmentados dentro de los bronquiolos respiratorios y los espacios aéreos circundantes, variablemente, engrosamiento septal alveolar fibrótico peribronquial, e inflamación mural asociada mínima (véase Wells et al. (2003) Sem Respir. Crit. Care Med. vol. 24).

La NID es una enfermedad pulmonar intersticial rara caracterizada por la acumulación de macrófagos en grandes cantidades en los espacios alveolares asociados con inflamación intersticial y/o fibrosis. Los macrófagos frecuentemente contienen pigmento de color marrón claro. Los nódulos linfoides son comunes, así como un infiltrado de eosinófilos disperso pero distinto. La NID es más común en fumadores (véase Tazelaar et al. (Sep. 21, 2010) Histopathology).

La NINE se caracteriza patológicamente por la inflamación intersticial uniforme y la fibrosis que aparecen durante un corto periodo de tiempo. La NINE difiere de otras enfermedades pulmonares intersticiales en que tiene un pronóstico generalmente bueno. Además, la uniformidad temporal de los cambios parenquimales observados en NINE contrasta enormemente con la heterogeneidad temporal de la neumonía intersticial usual (véase Coche et al. (2001) Brit J Radiol 74:189).

La BONO, a diferencia de la NINE, puede ser fatal en cuestión de días de los primeros síntomas agudos. Se caracteriza por la aparición rápida de síndrome de dificultad respiratoria aguda; por lo tanto, clínicamente, la BONO rápidamente progresiva puede ser indistinguible de la neumonía intersticial aguda. Las características histológicas incluyen agrupaciones de células inflamatorias mononucleares que forman tejido de granulación y taponan las vías respiratorias distales y los espacios alveolares. Estos tapones de tejido de granulación pueden formar pólipos que migran dentro de los conductos alveolares o pueden estar focalmente unidos a la pared (véase White & Ruth-Saad (2007) Crit. Care Nurse 27:53).

Detalles adicionales acerca de las características y terapias disponibles para estas enfermedades se pueden encontrar, por ejemplo, en el sitio web de la Asociación Americana del Pulmón en lungusa.org/lung-disease/pulmonary-fibrosis.

Los indicadores de diagnóstico de trastornos pulmonares incluyen la biopsia (por ejemplo, VATS o biopsia pulmonar quirúrgica), tomografía computarizada de alta resolución (TCAR) o métricas de respiración, tal como volumen espiratorio forzado (VEF1), capacidad vital (CV), capacidad vital forzada (CVF) y VEF1/CVF.

- 5 Los trastornos adicionales asociados con la expresión de MUC5B y/o los SNP asociados con MUC5B (por ejemplo, SNP rs35705950) pueden incluir, pero no se limitan a, trastornos de secreción mucosa, cánceres (por ejemplo, de ovario, de mama, de pulmón, pancreático, etc.), enfermedad ocular, colitis y cirrosis del hígado.

IV. Métodos de Diagnóstico y Pronóstico

- 10 Los métodos para detectar e identificar ácidos nucleicos y proteínas e interacciones entre tales moléculas implican técnicas de biología molecular convencional, microbiología y ADN recombinante dentro de la experiencia de la técnica. Tales técnicas se explican completamente en la literatura (véase, por ejemplo, Sambrook, Fritsch & Maniatis, Molecular Cloning: A Laboratory Manual, Segunda Edición 1989, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.; Animal Cell Culture, R. I. Freshney, ed., 1986).

15 A. Muestras biológicas

- Para la detección de una variante genética utilizando ADN genómico, se puede obtener una muestra biológica de casi cualquier tejido. Un experto en la técnica entenderá que una muestra de sangre o una torunda de mejilla se espera que porte la misma información de secuencia genética que una célula pulmonar. Para la detección de un nivel de expresión
20 dado, se utilizan típicamente muestras de tejido pulmonar y otros fluidos biológicos.

Las muestras biológicas pueden incluir una muestra de mucosa pulmonar o fluido biológico tal como sangre o componentes sanguíneos (plasma, suero), esputo, moco, orina, saliva, etc.

- 25 Una muestra de mucosa pulmonar se puede obtener utilizando métodos conocidos en la técnica, por ejemplo, un cepillo epitelial bronquial o condensado de aire exhalado. Los métodos adicionales incluyen biopsia bronquial, lavado bronquial, lavado broncoalveolar, lavado pulmonar completo, biopsia transendoscópica, catéter tras laringoscópico y lavado transtraqueal. Una revisión de técnicas comúnmente utilizadas, incluyendo comparaciones y cuestiones de seguridad, se proporciona en Busse et al. (2005) Am J Respir Crit Care Med 172:807-816.

- 30 Para las técnicas de lavado, se puede insertar un broncoscopio hasta el nivel deseado de las vías respiratorias. Un volumen pequeño de fluido estéril, fisiológicamente aceptable (por ejemplo, solución salina regulada) se libera, e inmediatamente se aspira. El material de lavado contiene células de la mucosa y epitelios superiores (Riise et al. (1996) Eur Resp J 9:1665).

- 35 Para el uso de un cepillo epitelial bronquial, se puede utilizar un cepillo de citología estéril, no irritante (por ejemplo, nylon). Se pueden tomar múltiples cepillados para asegurar un muestreo representativo. El cepillo se agita entonces en fluido fisiológicamente aceptable, y las células y los desechos se separan utilizando métodos rutinarios (Riise et al. (1992) Eur Resp J 5: 382).

- 40 Los componentes celulares se pueden aislar utilizando métodos conocidos en la técnica, por ejemplo, centrifugación. De manera similar, los componentes subcelulares (por ejemplo, exosomas o vesículas) se pueden aislar utilizando métodos conocidos o productos de separación comerciales (disponibles de BioCat, System Bio, Bioscientific, etc.). Un método ejemplar es descrito, por ejemplo, por They et al. (2006) Current Prot. Cell Biol.

45 B. Detección de variantes genéticas

- Los inventores han encontrado que las variaciones genéticas en los genes de mucina están asociadas con enfermedades pulmonares. Estas variaciones genéticas se pueden encontrar en cualquier parte del gen, por ejemplo, en las regiones reguladoras, intrones, o exones. Las variaciones genéticas relevantes también se pueden encontrar en las regiones de intergén, por ejemplo, en secuencias entre genes de mucina. Las inserciones, sustituciones y supresiones se incluyen en las variantes genéticas. Los polimorfismos de un solo nucleótido (los SNP) son variantes genéticas ejemplares.

- 50 En particular, 14 SNP independientes están asociados con trastornos pulmonares (por ejemplo, NIF o FPI). Los estudios descritos en la presente memoria demuestran que la presencia de uno o más de estos SNP asociados con MUC5B pueden conducir a la predisposición a una enfermedad pulmonar. Además, si están presentes, algunos de estos SNP están relacionados con un sitio de unión del factor de transcripción. El sitio de unión del factor de transcripción puede efectuar la modulación de la expresión de MUC5B, por ejemplo pérdida de E2F3, y la generación de HOXA9 y PAX-2.

- 60 La invención proporciona por tanto métodos para evaluar la presencia o ausencia de SNP en una muestra de un sujeto sospechoso de tener o desarrollar un trastorno pulmonar (por ejemplo, debido a antecedentes familiares). En ciertas realizaciones de la divulgación, se criban uno o más SNP en una o más muestras de un sujeto. Los SNP pueden ser asociados con uno o más genes, por ejemplo, uno o más genes MUC u otros genes asociados con la secreción mucosa. Un SNP asociado al gen MUC está asociado con MUC5B y/u otro gen MUC, tal como MUC5AC o MUC1. Los SNP contemplados para diagnóstico, tratamiento o pronóstico pueden incluir los SNP encontrados dentro de un gen MUC y/o
65 dentro de una región reguladora o promotora asociada con un gen MUC. Por ejemplo, uno o más SNP pueden incluir, pero no se limitan a, la detección de los SNP de MUC5B mostrados en la Tabla 4 (SEC ID N^{os}. 20-53), por ejemplo, SNP

rs35705950 (SEC ID N°: 24), solos o en combinación con otras variaciones genéticas o SNP y/u otros métodos de diagnóstico o pronóstico.

Se conocen en la técnica métodos para detectar variantes genéticas tales como un SNP, por ejemplo, transferencia de Southern o de Northern, arreglo de nucleótidos, métodos de amplificación, etc. Los cebadores o sondas se diseñan para hibridarse con una secuencia objetivo. Por ejemplo, el ADN genómico se puede cribar para determinar la presencia de un elemento genético identificado de utilización de una sonda en base a una o más secuencias, por ejemplo, utilizando una sonda con una identidad sustancial con una subsecuencia del gen MUC5B, tal como una de las subsecuencias mostradas en la Tabla 4 (SEC ID N°s: 20-53). Las secuencias genómicas de MUC5B humanas ejemplares que se pueden utilizar para la referencia y el diseño de sonda y cebador se encuentran en las Entradas de GenBank Nos. AF107890.1 y AJ004862.1. El ARN expresado también se puede cribar, pero puede no incluir todas las variaciones genéticas relevantes. Se pueden emplear diversos grados de rigurosidad de hibridación en el ensayo. A medida que las condiciones para la hibridación se hacen más rigurosas, debe haber un mayor grado de complementariedad entre la sonda y el objetivo para que ocurra la formación de dúplex. Por tanto, las condiciones de alta rigurosidad se utilizan típicamente para detectar un SNP.

Por tanto, en algunas realizaciones, un gen MUC5B de variante genética en un sujeto se detecta poniendo en contacto un ácido nucleico en una muestra del sujeto con una sonda que tiene una identidad sustancial con una subsecuencia de gen MUC5B, y determinando si el ácido nucleico indica que el sujeto tiene un gen MUC5B de variante genética. En algunos casos, la muestra se puede procesar antes de la amplificación, por ejemplo, para separar el ADN genómico de otros componentes de la muestra. En algunos casos, la sonda tiene al menos el 90, 92, 94, 95, 96, 98, 99, o 100% de identidad con la subsecuencia de gen MUC5B. Típicamente, la sonda es de entre 10-500 nucleótidos de longitud, por ejemplo, 10-100, 10-40, 10-20, 20-100, 100-400, etc. En el caso de detectar un SNP, la sonda puede ser incluso más corta, por ejemplo, de 8-20 nucleótidos de longitud. En algunos casos, la secuencia de gen MUC5B a detectar incluye al menos 8 nucleótidos contiguos, por ejemplo, al menos 10, 15, 20, 25, 30, 35 o más nucleótidos contiguos de una de las secuencias mostradas en la SEC ID N°: 20-53. En algunas realizaciones, la secuencia a detectar incluye 8 nucleótidos contiguos, por ejemplo, al menos 10, 15, 20, 25, 30, 35 o más nucleótidos contiguos de la SEC ID N°: 24. En algunos aspectos, los nucleótidos contiguos incluyen el nucleótido 28 de la SEC ID N°: 24.

El grado de rigurosidad se puede controlar por temperatura, fuerza iónica, pH y/o la presencia de un solvente parcialmente desnaturante tal como formamida. Por ejemplo, la rigurosidad de la hibridación se varía convenientemente cambiando la concentración de formamida dentro del intervalo de hasta y aproximadamente 50%. El grado de complementariedad (identidad de secuencia) requerido para la unión detectable variará de acuerdo con la rigurosidad del medio de hibridación y/o el medio de lavado. En ciertas realizaciones, en particular para la detección de un SNP particular, el grado de complementariedad es de aproximadamente el 100 por ciento. En otras realizaciones, las variaciones de secuencia pueden dar como resultado <100% de complementariedad, <90% de sondas de complementariedad, <80% de sondas de complementariedad, etc., en particular, en una secuencia que no implica un SNP. En algunos ejemplos, por ejemplo, la detección de homólogos de especie, los cebadores se pueden compensar reduciendo la rigurosidad del medio de hibridación y/o de lavado.

Las condiciones de alta rigurosidad para la hibridación de ácido nucleico son bien conocidas en la técnica. Por ejemplo, las condiciones pueden comprender condiciones de poca sal y/o alta temperatura, tal como es proporcionado por NaCl de aproximadamente 0,02 M a aproximadamente 0,15 M a temperaturas de aproximadamente 50°C a aproximadamente 70°C. Otras condiciones ejemplares se describen en los siguientes Ejemplos. Se entiende que la temperatura y la fuerza iónica de una rigurosidad deseada se determinan en parte por la longitud de el/los ácido(s) nucleico(s) particular(es), la longitud y contenido de nucleótidos de la/las secuencia(s) objetivo, la composición de carga de el/los ácido(s) nucleico(s), y por la presencia o concentración de formamida, cloruro de tetrametilamonio u otro(s) solvente(s) en una mezcla de hibridación. Los ácidos nucleicos pueden ser completamente complementarios a una secuencia objetivo o exhibir uno o más desapareamientos.

Los ácidos nucleicos de interés (por ejemplo, ácidos nucleicos que comprenden, o comprendidos dentro de, las SEC ID N°s: 20-53) también se pueden amplificar utilizando una variedad de técnicas de amplificación conocidas. Por ejemplo, la tecnología de reacción en cadena de la polimerasa (PCR) se puede utilizar para amplificar las secuencias objetivo (por ejemplo, variantes genéticas) directamente a partir de ADN, ARN, o ADNc. En algunas realizaciones, se amplifica un tramo de ácidos nucleicos utilizando cebadores a ambos lados de una variación genética objetivo, y el producto de amplificación se secuencia entonces para detectar la variación genética objetivo (utilizando, por ejemplo, secuenciación de Sanger, Pirosecuenciación, tecnologías de secuenciación de Nextgen®). Por ejemplo, los cebadores se pueden diseñar para hibridarse a ambos lados de la región reguladora aguas arriba del gen MUC5B, y la secuencia intermedia determinada para detectar un SNP en la región promotora. En algunas realizaciones, uno de los cebadores se puede diseñar para hibridarse con variante genética objetivo. En algunos casos, se puede identificar un nucleótido de variante genética utilizando la RT-PCR, por ejemplo, utilizando monómeros de nucleótido marcados. De esta manera, se puede detectar la identidad del nucleótido en una posición dada a medida que se añade al ácido nucleico de polimerización. El sistema Scorpion™ es un ejemplo comercialmente disponible de esta tecnología.

Por tanto, en algunas realizaciones, un gen MUC5B de variante genética en un sujeto se puede detectar amplificando un ácido nucleico en una muestra del sujeto para formar un producto de amplificación, y determinando si el producto de amplificación indica un gen MUC5B de variante genética. En algunos casos, la muestra se puede procesar antes de la

amplificación, por ejemplo, para separar el ADN genómico de otros componentes de la muestra. En algunos casos, la amplificación comprende poner en contacto la muestra con cebadores de amplificación que tienen identidad sustancial con las subsecuencias genómicas de MUC5B, por ejemplo, al menos 90, 92, 94, 95, 96, 98, 99, o 100% de identidad. Típicamente, la secuencia a amplificar es de entre 30-1000 nucleótidos de longitud, por ejemplo, 50-500, 50-400, 100-400, 50-200, 100-300, etc. En algunos casos, la secuencia a amplificar o detectar incluye al menos 8 nucleótidos contiguos, por ejemplo, al menos 10, 15, 20, 25, 30, 35 o más nucleótidos contiguos de una de las secuencias mostradas en las SEC ID N^{os}: 20-53. En algunas realizaciones, la secuencia a amplificar o detectar incluye 8 nucleótidos contiguos, por ejemplo, al menos 10, 15, 20, 25, 30, 35 o más nucleótidos contiguos de la SEC ID N^o: 24. En algunos aspectos, los nucleótidos contiguos incluyen el nucleótido 28 de la SEC ID N^o: 24.

Las técnicas de amplificación también pueden ser útiles para clonar secuencias de ácidos nucleicos, para hacer ácidos nucleicos para uso como sondas para detectar la presencia de un ácido nucleico objetivo en las muestras, para secuenciación de ácidos nucleicos, para muestras de control o para otros propósitos. Las sondas y cebadores también están fácilmente disponibles de fuentes comerciales, por ejemplo, de Invitrogen, Clontech, etc.

C. Detección de niveles de expresión

La expresión de un gen dado, por ejemplo, MUC5B u otra mucina, marcador de enfermedad pulmonar, o estándar (control), se detecta típicamente detectando la cantidad de ARN (por ejemplo, ARNm) o proteína. Los niveles de muestra se pueden comparar con un nivel de control.

Los métodos para detectar ARN son en gran medida acumulativos con los ensayos de detección de ácido nucleico descritos anteriormente. El ARN a detectar puede incluir ARNm. Se puede llevar a cabo una reacción de transcriptasa inversa y la secuencia objetivo se puede entonces amplificar utilizando la PCR estándar. La PCR cuantitativa (qPCR) o PCR en tiempo real (RT-PCR) es útil para determinar los niveles de expresión relativos, cuando se compara con un control. Las técnicas y plataformas de PCR cuantitativa son conocidas en la técnica, y están disponibles comercialmente (véase, por ejemplo, el sitio web del Symposium de qPCR, disponible en qpcrsymposium.com). Los arreglos de ácido nucleico son también útiles para detectar la expresión de ácido nucleico. Los arreglos personalizables están disponibles de, por ejemplo, Affimatrix. Una secuencia de ARNm de MUC5B humana ejemplar, por ejemplo, para diseño de sonda y cebador, se puede encontrar en la Entrada de GenBank No AF086604.1.

Los niveles de proteína se pueden detectar utilizando anticuerpos o fragmentos de anticuerpo específicos para esa proteína, ligandos naturales, moléculas pequeñas, aptámeros, etc. Una secuencia de MUC5B humana ejemplar, por ejemplo, para cribar un agente de fijación de objetivo, se puede encontrar en la Entrada de UniProt No 000446.

Las técnicas basadas en anticuerpo se conocen en la técnica, y se describen, por ejemplo, en Harlow & Lane (1988) *Antibodies: A Laboratory Manual* y Harlow (1998) *Using Antibodies: A Laboratory Manual*; Wild, *The Immunoassay Handbook*, 3^a edición (2005) and Law, *Immunoassay: A Practical Guide* (1996). El ensayo se puede dirigir a la detección de un objetivo molecular (por ejemplo, proteína o antígeno), o una célula, tejido, muestra biológica, muestra líquida o superficie sospechosa de portar un anticuerpo u objetivo de anticuerpo.

Una lista no exhaustiva de inmunoensayos incluye: formatos competitivos y no competitivos, ensayos de inmunoabsorción ligados a enzimas (ELISA), ensayos de microspot, transferencias de Western, filtración en gel y cromatografía, inmunocromatografía, inmunohistoquímica, citometría de flujo o clasificación celular activada por fluorescencia (FACS), microarreglos y más. Tales técnicas también se pueden utilizar in situ, ex vivo, in vivo, por ejemplo, para formación de imágenes de diagnóstico.

Los aptámeros son ácidos nucleicos que se diseñan para unirse a una amplia variedad de objetivos de manera no de Watson Crick. Un aptámero se puede por tanto utilizar para detectar o de otra manera dirigirse a casi cualquier molécula de interés, incluyendo una proteína asociada con enfermedad pulmonar. Los métodos de construcción y determinación de las características de unión de los aptámeros son bien conocidos en la técnica. Por ejemplo, tales técnicas se describen en las patentes de EE.UU. Nos 5.582.981, 5.595.877 y 5.637.459. Los aptámeros son típicamente de al menos 5 nucleótidos, 10, 20, 30 o 40 nucleótidos de longitud, y se pueden componer de ácidos nucleicos modificados para mejorar la estabilidad. Se pueden añadir secuencias flanqueantes para estabilidad estructural, por ejemplo, para formar estructuras tridimensionales en el aptámero.

Los agentes de detección de proteína descritos en la presente memoria también se pueden utilizar como un tratamiento y/o diagnóstico de enfermedad pulmonar o predictor de la progresión de la enfermedad, por ejemplo, propensión para la supervivencia, en un sujeto que tiene o sospechoso de desarrollar un trastorno pulmonar. Los anticuerpos de MUC5B se pueden utilizar para evaluar los niveles de proteína MUC5B en un sujeto que tiene o sospechoso de desarrollar un trastorno pulmonar. Se contempla en la presente memoria que los anticuerpos o fragmentos de anticuerpo se pueden utilizar para modular la producción de MUC5B en un sujeto que tiene o sospechoso de desarrollar una enfermedad pulmonar. Se pueden utilizar uno o más agentes capaces de modular MUC5B para tratar a un sujeto que tiene o sospechoso de desarrollar un trastorno pulmonar. Se pueden generar uno o más anticuerpos o fragmentos de anticuerpo para detectar uno o más de los SNP descritos en la presente mediante cualquier método conocido en la técnica.

Las pruebas de diagnóstico de MUC5B pueden incluir, pero no se limitan a, solas o en combinación, análisis del SNP rs35705950 en el gen MUC5B, niveles de ARNm de MUC5B y/o niveles de proteína MUC5B.

D. Marcadores de enfermedad pulmonar adicionales

Los métodos de detección anteriores se pueden aplicar a marcadores de enfermedad pulmonar adicionales. Esto es, se puede determinar el nivel de expresión o la presencia de variantes genéticas de al menos un gen marcador de enfermedad pulmonar adicional, o se puede determinar la actividad de la proteína marcadora, y comparar con un control estándar para el marcador de enfermedad pulmonar. El examen de marcadores de enfermedad pulmonar adicionales se puede utilizar para confirmar un diagnóstico de enfermedad pulmonar, monitorizar la progresión de la enfermedad, o determinar la eficacia de un curso de tratamiento en un sujeto.

En algunos casos, la enfermedad pulmonar está indicada por un incremento del número de linfocitos, por ejemplo, células CD4+CD28 (Moeller et al. (2009) Am. J. Resp. Crit Care. Med. 179:588; Gilani (2010) PLoS One 5:e8959).

Las variaciones genéticas en los siguientes genes están asociadas con enfermedad pulmonar; Proteína Surfactante A2, Proteína Surfactante B, Proteína Surfactante C, TERC, TERT, IL-1RN, IL-1 α , It-1 β , TNF, Linfotóxina α , TNF-RII, IL-10, IL-6, IL-12, IFN γ , TGF β , CR1, ACE, IL-8, CXCR1, CXCR2, MUC1 (KL6), o MUC5AC. Por tanto, la divulgación incluye además métodos para determinar si el genoma de un sujeto comprende una variante genética de al menos un gen seleccionado de entre estos genes. La presencia de una variante genética indica que el sujeto tiene o está en riesgo de desarrollar enfermedad pulmonar. Dicha determinación se puede combinar opcionalmente con la determinación de si el genoma del sujeto comprende un gen MUC5B de variante genética, o la determinación de si el sujeto tiene un nivel elevado de ARN o proteína MUC5B para confirmar o reforzar el diagnóstico o pronóstico.

La expresión anormal en los siguientes genes también puede ser indicativa de enfermedad pulmonar: Proteína Surfactante A, Proteína Surfactante D, KL-6/MUC1, CC16, CK-19, Ca 19-9, SLX, MCP-1, MIP-1a, ITAC, glutatión, péptido de procolágeno tipo III, sIL-2R, ACE, neopterin, beta-glucuronidasa, LDH, CCL-18, CCL-2, CXCL12, MMP7, y osteopontina. Por tanto, la expresión de uno de estos genes se puede detectar y comparar con un control, en donde un nivel de expresión anormal indica que el sujeto tiene o está en riesgo de desarrollar enfermedad pulmonar. Dicha determinación se puede combinar opcionalmente con la determinación de si el genoma del sujeto comprende un gen MUC5B de variante genética, o la determinación de si el sujeto tiene un nivel elevado de ARN o proteína MUC5B para confirmar o reforzar el diagnóstico o pronóstico.

E. Indicaciones

Los métodos de detección descritos en la presente memoria se pueden utilizar para diagnóstico, pronóstico, predicción del riesgo, determinación de un curso de tratamiento, monitorización de la eficacia terapéutica y monitorización de la progresión de la enfermedad. Un experto apreciará que cada uno de los métodos de detección se pueden utilizar solos o en combinación.

Por ejemplo, la presencia de un gen MUC5B de variante genética se puede determinar en un sujeto sospechoso de tener o en riesgo de desarrollar un trastorno pulmonar. En el caso de que se observe un gen MUC5B de variante genética, el sujeto puede opcionalmente someterse a pruebas adicionales, por ejemplo, para determinar el nivel de expresión del gen MUC5B, o detectar una forma de variante genética de al menos un marcador de enfermedad pulmonar adicional. El sujeto puede ser prescrito con un curso de tratamiento en base a los resultados de una o más pruebas. Tal tratamiento puede incluir la administración de un antagonista de MUC5B, o un tratamiento de enfermedad pulmonar estándar tal como un fármaco mucolítico. El nivel de expresión del gen MUC5B se puede detectar de nuevo después del tratamiento, o periódicamente durante el curso del tratamiento, para determinar la eficacia terapéutica del tratamiento. Por ejemplo, si se prescribe un tratamiento de enfermedad pulmonar para la administración periódica (por ejemplo, diariamente, dos veces al día, semanalmente, etc.), el nivel de expresión del gen MUC5B se puede monitorizar periódicamente después (por ejemplo, mensualmente).

Los métodos de detección de la invención se pueden utilizar para determinar si el sujeto tiene una forma atenuada de la enfermedad pulmonar. Los inventores han demostrado que los individuos que portan el gen MUC5B de variante genética rs35705950 tienen un mejor pronóstico de enfermedad pulmonar que los individuos que no portan un gen MUC5B de variante genética. Por tanto, la determinación de si un individuo porta el gen MUC5B de variante genética se puede utilizar para diseñar un curso de tratamiento para el individuo.

V. Métodos de Tratamiento

A. Tratamientos de enfermedad pulmonar

Una serie de tratamientos de enfermedad pulmonar están disponibles para tratar la inflamación de las vías respiratorias y/o el exceso de secreción de moco. Estos incluyen agentes que se pueden clasificar en líneas generales, por ejemplo, como agentes mucolíticos, agentes mucorreguladores, agentes mucocinéticos y expectorantes (véase, por ejemplo, Balsamo et al. (2010) Eur. Respir. Rev. 19:127-33), aunque hay cierto solapamiento en las categorías. Tales agentes son útiles para tratar las enfermedades pulmonares descritas en la presente memoria, por ejemplo, como parte de un curso de tratamiento y monitorización, o después de la detección de ARN o proteína MUC5B elevada, o la detección de un gen MUC5B de variante genética.

Los fármacos mucolíticos son aquellos que disminuyen la viscosidad del moco, ya sea mediante despolimerización de glicoproteínas de mucina o mediante despolimerización de redes poliméricas de ADN y F-actina. El primer modo de acción puede ser de particularmente útil para tratar el exceso de MUC5B. Los mucolíticos ejemplares incluyen N-acetilcisteína, N-acistelina, erdosteína, dornasa alfa, timosina beta4, dextrano, pulmozyme, heparina y bronquitol (manosa inhalada).

Los mucorreguladores son aquellos agentes que regulan la secreción de moco, o interfieren con la red de ADN/F-actina. Ejemplos de mucorreguladores incluyen, por ejemplo, carbocisteína, agentes anticolinérgicos, glucocorticoides y antibióticos macrólidos.

Los agentes mucocinéticos incrementan la eliminación de moco actuando sobre el revestimiento ciliar de las vías respiratorias. Los agentes mucocinéticos ejemplares incluyen, por ejemplo, broncodilatadores, surfactantes y ambroxol.

Los expectorantes son agentes que inducen la descarga de moco de las vías respiratorias o del tracto respiratorio. Algunos ejemplos incluyen solución salina hipertónica, guaifenesina, dornasa/pulmozyme y bronquitol (manosa inhalada).

El tratamiento de enfermedad pulmonar, tal como los agentes descritos anteriormente, se puede utilizar solo, secuencialmente, o en combinación según los métodos descritos en la presente memoria. Un tratamiento de enfermedad pulmonar se puede utilizar en combinación con un inhibidor más dirigido de la expresión de MUC5B.

B. Antagonistas de MUC5B

Los resultados descritos en la presente memoria indican que la expresión elevada del gen MUC5B está asociada con enfermedad pulmonar. La divulgación incluye por tanto métodos y composiciones para inhibir la expresión, secreción y/o actividad de MUC5B. Los inhibidores ejemplares incluyen ARNip y antisentido, pRNA (ARN asociado a promotor, véase, por ejemplo, Schmitz et al. (2010) Genes Dev. 24:2264-69), anticuerpos específicos de MUC5B y fragmentos de los mismos, y aptámeros específicos de MUC5B. Se puede inhibir la actividad de MUC5B o se puede incrementar la eliminación de MUC5B, por ejemplo, utilizando agentes mucolíticos, inhibidores de glicosilación, o inhibidores de secreción de proteína. Los términos "inhibidor" y "antagonista" y términos similares se utilizan de manera sinónima en la presente memoria.

Por tanto, una secuencia de nucleótidos que interfiere específicamente con la expresión del gen MUC5B en el nivel de transcripción o de traducción se puede utilizar para tratar o prevenir la enfermedad pulmonar. Este enfoque puede utilizar, por ejemplo, ARNip y/u oligonucleótidos antisentido para bloquear la transcripción o traducción de un ARNm específico (por ejemplo, un ARN de variante genética), ya sea induciendo la degradación del ARNm con un ARNip o enmascarando el ARNm con un ácido nucleico antisentido. El ARNip o constructo antisentido puede no bloquear significativamente la expresión de otros genes de mucina.

El ARNip de doble cadena que corresponde al gen MUC5B se puede utilizar para silenciar la transcripción y/o traducción induciendo la degradación de las transcripciones de ARNm de MUC5B, y por tanto tratar o prevenir la enfermedad pulmonar (por ejemplo, la enfermedad pulmonar asociada con MUC5B de variante genética). El ARNip es típicamente de aproximadamente 5 a aproximadamente 100 nucleótidos de longitud, más típicamente de aproximadamente 10 a aproximadamente 50 nucleótidos de longitud, lo más típicamente de aproximadamente 15 a aproximadamente 30 nucleótidos de longitud. Las moléculas de ARNip y los métodos para generarlas se describen en, por ejemplo, Bass, 2001, Nature, 411, 428-429; Elbashir et al., 2001, Nature, 411, 494-498; el documento WO 00/44895; el documento WO 01/36646; el documento WO 99/32619; el documento WO 00/01846; el documento WO 01/29058; el documento WO 99/07409; y el documento WO 00/44914. Una molécula de ADN que transcribe ARNbc o ARNip (por ejemplo, como un dúplex de horquilla) también proporciona ARNi. Las moléculas de ADN para transcribir ARNbc se describen en la patente de EE.UU. N° 6.573.099, y en la publicación de solicitud de patente de EE.UU. N° 2002/0160393 y 2003/0027783, y Tuschl y Borkhardt, Molecular Interventions, 2: 158 (2002). Por ejemplo, los oligonucleótidos de ARNbc que se hibridan específicamente con las secuencias de ácidos nucleicos de MUC5B descritas en la presente memoria se pueden utilizar en los métodos. Se puede utilizar una disminución en la gravedad de los síntomas de enfermedad pulmonar en comparación con los síntomas detectados en la ausencia del ARN interferente para monitorizar la eficacia del ARNip.

Los oligonucleótidos antisentido que se hibridan específicamente con secuencias de ácidos nucleicos que codifican polipéptidos MUC5B se pueden utilizar también para silenciar la transcripción y/o traducción, y por tanto tratar o prevenir la enfermedad pulmonar. Por ejemplo, se pueden utilizar oligonucleótidos antisentido que se hibridan específicamente con una secuencia de polinucleótidos de MUC5B. Se puede utilizar una disminución en la gravedad de los síntomas de enfermedad pulmonar en comparación con los síntomas detectados en la ausencia de los ácidos nucleicos antisentido para monitorizar la eficacia de los ácidos nucleicos antisentido.

Los ácidos nucleicos antisentido son moléculas de ADN o ARN que son complementarias a al menos una porción de una molécula de ARNm específica (véase, por ejemplo, Weintraub, Scientific American, 262:40 (1990)). Típicamente, los oligonucleótidos antisentido sintéticos son generalmente de entre 15 y 25 bases de longitud. Los ácidos nucleicos antisentido pueden comprender nucleótidos de origen natural o nucleótidos modificados tales como, por ejemplo, nucleótidos de cadena principal modificada con fosforotioato, metilfosfonato, y azúcar-fosfato anomérica.

En la célula, los ácidos nucleicos antisentido se hibridan con el ARNm correspondiente, formando una molécula de doble cadena. Los ácidos nucleicos antisentido, interfieren con la traducción del ARNm, puesto que la célula no traducirá un ARNm que es de doble cadena. Se prefieren los oligómeros antisentido de aproximadamente 15 nucleótidos, puesto que son fácilmente sintetizados y es menos probable que causen problemas que las moléculas más grandes cuando se introducen en la célula productora de nucleótidos mutante objetivo. El uso de métodos antisentido para inhibir la traducción in vitro de genes es bien conocido en la técnica (Marcus-Sakura, *Anal. Biochem.*, 172:289, (1988)). De forma menos común, se pueden utilizar moléculas antisentido que se unen directamente al ADN.

El ARNip y antisentido se pueden suministrar al sujeto utilizando cualquier medio conocido en la técnica, incluyendo mediante inyección, inhalación o ingestión oral. Otro sistema de suministro adecuado es un sistema de dispersión coloidal tal como, por ejemplo, complejos de macromoléculas, nanocápsulas, microesferas, glóbulos, y sistemas basados en lípidos que incluyen emulsiones de aceite en agua, micelas, micelas mezcladas y liposomas. El sistema coloidal preferido es un liposoma. Los liposomas son vesículas de membrana artificial que son útiles como vehículos de suministro in vitro e in vivo. Los ácidos nucleicos, incluyendo ARN y ADN dentro de los liposomas y ser suministrados a las células en una forma biológicamente activa (Fraley, et al., *Trends Biochem. Sci.*, 6:77, 1981). Los liposomas se pueden dirigir a tipos de célula o tejidos específicos utilizando cualquier medio conocido en la técnica.

La divulgación también proporciona anticuerpos que se unen específicamente a la proteína MUC5B. Tales anticuerpos se pueden utilizar para secuestrar el MUC5B secretado, por ejemplo, para prevenir la actividad formadora de gel y la formación de exceso de moco.

Un anticuerpo que detecta específicamente MUC5B, y no otras proteínas de mucina, se puede aislar utilizando las técnicas estándar descritas en la presente memoria. Las secuencias de proteínas para MUC5B en una serie de especies, por ejemplo, humanos, primates no humanos, ratas, perros, gatos, caballos, bovinos, etc., están públicamente disponibles.

Los anticuerpos monoclonales se obtienen mediante diversas técnicas familiares para los expertos en la técnica. Brevemente, las células del bazo de un animal inmunizado con un antígeno deseado se immortalizan, comúnmente mediante fusión con una célula de mieloma (véase, por ejemplo, Kohler & Milstein, *Eur. J. Immunol.* 6: 511-519 (1976)). Los métodos alternativos de immortalización incluyen transformación con Virus de Epstein-Barr, oncogenes o retrovirus, u otros métodos bien conocidos en la técnica. Las colonias que surgen de células immortalizadas individuales se criban para la producción de anticuerpos de la especificidad y afinidad deseadas para el antígeno, y la producción de los anticuerpos monoclonales producidos por tales células se puede incrementar mediante diversas técnicas, incluyendo la inyección en la cavidad peritoneal de un huésped vertebrado. Alternativamente, se pueden aislar las secuencias de ADN que codifican un anticuerpo monoclonal o un fragmento de unión del mismo cribando una biblioteca de ADN de células B humanas según el protocolo general indicado por Huse et al., *Science* 246: 1275-1281 (1989).

Los anticuerpos monoclonales se recogen y se valoran contra el MUC5B en un inmunoensayo, por ejemplo, un inmunoensayo en fase sólida con el inmunógeno inmovilizado sobre un soporte sólido. Los anticuerpos monoclonales habitualmente se unirán con un K_d de al menos aproximadamente 0,1mM, y más habitualmente de al menos aproximadamente 1 μ M, y se pueden diseñar a menudo para unirse con un K_d de 1 nM o menos.

Las inmunoglobulinas, incluyendo fragmentos de unión a MUC5B y derivados de los mismos, se pueden producir fácilmente mediante una variedad de técnicas de ADN recombinante, incluyendo mediante la expresión en células transfectadas (por ejemplo, células eucariotas immortalizadas, tales como células de mieloma o de hibridoma) o en ratones, ratas, conejos, u otro vertebrado capaz de producir anticuerpos mediante métodos bien conocidos. Las células fuente adecuadas para las secuencias de ADN y las células huésped para la expresión y secreción de inmunoglobulina se pueden obtener de una serie de fuentes, tales como la American Type Culture Collection (Catalogue of Cell Lines and Hybridomas, Quinta edición (1985) Rockville, Md).

El anticuerpo es un anticuerpo humanizado, i.e., un anticuerpo que conserva la reactividad de un anticuerpo no humano mientras que es menos inmunogénico en humanos. Esto se puede lograr, por ejemplo, conservando las regiones de CDR no humanas que son específicas para MUC5B, y reemplazando las partes restantes del anticuerpo con sus homólogas humanas. Véase, por ejemplo, Morrison et al., *PNAS USA*, 81:6851-6855 (1984); Morrison y Oi, *Adv. Immunol.*, 44: 65- 92 (1988); Verhoeyen et al., *Science*, 239: 1534-1536 (1988); Padlan, *Molec. Immun.*, 28:489-498 (1991); Padlan, *Molec. Immun.*, 31(3):169-217 (1994). Las técnicas para humanizar anticuerpos son bien conocidas en la técnica y se describen en, por ejemplo, las patentes de EE.UU. Nos 4.816.567; 5.530.101; 5.859.205; 5.585.089; 5.693.761; 5.693.762; 5.777.085; 6.180.370; 6.210.671; y 6.329.511; y el documento WO 87/02671; la Solicitud de Patente Europea 0173494; Jones et al. (1986) *Nature* 321 :522; y Verhoyen et al. (1988) *Science* 239: 1534. Los anticuerpos humanizados se describen además en, por ejemplo, Winter y Milstein (1991) *Nature* 349:293. Por ejemplo, los polinucleótidos que comprenden una primera secuencia que codifica las regiones de estructura de inmunoglobulina humanizadas y un segundo conjunto de secuencias que codifica las regiones determinantes de complementariedad de inmunoglobulina deseadas se pueden producir sintéticamente o combinando segmentos de ADNc o de ADN genómico apropiados. Las secuencias de ADN de región constante humanas se pueden aislar de acuerdo con procedimientos bien conocidos de una variedad de células humanas.

La actividad de la proteína MUC5B se puede inhibir, o la eliminación de MUC5B se puede incrementar, utilizando agentes mucolíticos que desintegran el moco y proteolizan las mucinas. Los agentes mucolíticos se describen en la presente memoria. Los inhibidores adicionales de proteína MUC5B incluyen inhibidores de glicosilación e inhibidores de secreción de proteína de las células epiteliales. Un inhibidor de glicosilación ejemplar incluye bencil-O-N-acetil-D galactosamina (específica para O-glucanos) y. Inhibidores adicionales de glicosilación de proteínas se describen, por ejemplo, en Jacob (1995) Curr. Opin. Structural Biol. 5:605-11 y Patsos et al. 2005 Biochem Soc. Trans. 33 :721-23. Los inhibidores de secreción incluyen Brefeldina A, conchicina, y moléculas pequeñas tales como las descritas en Stockwell (2006) Nat. Chem. Biol. 2: 7-8. La actividad de MUC5B también se puede modular dirigiendo la proteína MARCKS (Adler et al. (2000) Chest 117: Supp 1 266S-267S).

C. Métodos de Identificación de Antagonistas de MUC5B

La invención además proporciona métodos para identificar antagonistas adicionales de la expresión, secreción y/o actividad de MUC5B. Los métodos para cribar antagonistas pueden implicar medir la habilidad de los antagonistas potenciales para reducir una actividad de MUC5B identificable o competir por la unión con un agente de unión conocido (por ejemplo, anticuerpo específico de MUC5B). Por ejemplo, los agentes candidatos se pueden cribar por su capacidad para reducir la formación de gel de MUC5B, reducir la secreción de MUC5B, reducir la glicosilación de MUC5B, etc.

Los métodos de cribado se pueden realizar como ensayos in vitro o basados en células. Los ensayos basados en células se pueden realizar en cualquier célula en la que se exprese MUC5B, ya sea endógenamente o a través de métodos recombinantes. Los ensayos basados en células pueden implicar células completas o fracciones de células que contienen MUC5B para cribar la unión del agente o la modulación de la actividad de MUC5B por el agente. Los ensayos basados en células adecuados se describen en, por ejemplo, DePaola et al., Annals of Biomedical Engineering 29: 1-9 (2001).

Los agentes que se identifican inicialmente como inhibidores de MUC5B se pueden probar además para validar la actividad aparente. Preferiblemente tales estudios se conducen con modelos adecuados de enfermedad pulmonar basados en células o animales. El formato básico de tales métodos implica administrar un compuesto principal identificado durante un cribado inicial a un animal que sirve como un modelo y luego determinar si de hecho se mejora la enfermedad pulmonar. Los modelos animales utilizados en estudios de validación generalmente son mamíferos de cualquier clase. Ejemplos específicos de animales adecuados incluyen, pero no se limitan a, primates (por ejemplo, chimpancés, monos y similares) y roedores (por ejemplo, ratones, ratas, cobayas, conejos y similares).

Los agentes probados como antagonistas potenciales de MUC5B pueden ser cualquier compuesto químico pequeño, o una entidad biológica, tal como un polipéptido, azúcar, ácido nucleico o lípido. Alternativamente, los moduladores pueden ser versiones genéticamente alteradas de MUC5B, por ejemplo, formas que no son glicosiladas. Esencialmente cualquier compuesto químico se puede utilizar como un modulador potencial o ligando en los ensayos, aunque con la mayor frecuencia se utilizan compuestos que se pueden disolver en soluciones acuosas u orgánicas (especialmente basadas en DMSO). Los ensayos se diseñan para cribar grandes bibliotecas químicas automatizando los pasos del ensayo y proporcionando compuestos de cualquier fuente conveniente para los ensayos, que se ejecutan típicamente en paralelo (por ejemplo, en formatos de microtitulación sobre placas de microtitulación en ensayos robotizados).

Los métodos de cribado de alto rendimiento pueden implicar proporcionar una biblioteca química o peptídica combinatoria que contiene un número grande de compuestos terapéuticos potenciales (compuestos moduladores o de ligando potenciales). Tales "bibliotecas químicas combinatorias" o "bibliotecas de ligando" se criban luego en uno o más ensayos, como se describe en la presente memoria, para identificar esos miembros de la biblioteca (especies químicas particulares o subclases) que muestran una actividad característica deseada. Los compuestos identificados de este modo pueden servir como "compuestos principales" convencionales o se pueden utilizar por sí mismos como productos terapéuticos potenciales o reales.

Una biblioteca química combinatoria es una colección de diversos compuestos químicos generados ya sea por síntesis química o síntesis biológica, combinando una serie de "bloques de construcción" químicos tales como reactivos. Por ejemplo, una biblioteca química combinatoria lineal tal como una biblioteca de polipéptido se forma combinando un conjunto de bloques de construcción químicos (aminoácidos) de todas las maneras posibles para una longitud de compuesto dada (i.e., el número de aminoácidos en un compuesto polipeptídico). Se pueden sintetizar millones de compuestos químicos a través de tal mezcla combinatoria de bloques de construcción químicos.

La preparación y cribado de bibliotecas químicas combinatorias es bien conocida para los expertos en la técnica. Tales bibliotecas químicas combinatorias incluyen, pero no se limitan a, bibliotecas peptídicas (véase, por ejemplo, la Patente de EE.UU. 5.010.175, Furka, Int. J. Pept. Prot. Res. 37:487-493 (1991) y Houghton et al., Nature 354:84-88 (1991)). También se pueden utilizar otras sustancias químicas para generar bibliotecas de diversidad química. Tales sustancias químicas incluyen, pero no se limitan a: peptoides (por ejemplo, Publicación de PCT No WO 91/19735), péptidos codificados (por ejemplo, Publicación de PCT WO 93/20242), bio-oligómeros aleatorios (por ejemplo, Publicación de PCT No WO 92/00091), benzodiazepinas (por ejemplo, Patente de EE.UU. N°. 5.288.514), diversómeros tales como hidantoínas, benzodiazepinas y dipéptidos (Hobbs et al., Proc. Nat. Acad. Sci. USA 90:6909-6913 (1993)), polipéptidos vinílicos (Hagihara et al., J. Amer. Chem. Soc. 114:6568 (1992)), peptidomiméticos no peptídicos con andamiaje de glucosa (Hirschmann et al., J. Amer. Chem. Soc. 114:9217-9218 (1992)), síntesis orgánicas análogas de bibliotecas de compuestos pequeños (Chen et al., J. Amer. Chem. Soc. 116:2661 (1994)), oligocarbamatos (Cho et al., Science 261:1303 (1993)), y/o

peptidilfosfonatos (Campbell et al., J. Org. Chem. 59:658 (1994)), bibliotecas de ácidos nucleicos (ee Ausubel, Berger y Sambrook, all supra), bibliotecas de ácidos nucleicos peptídicos (véase, por ejemplo, la Patente de EE.UU. N° 5.539.083), bibliotecas de anticuerpos (véase, por ejemplo, Vaughn et al., Nature Biotechnology, 14(3):309-314 (1996) y PCT/US96/10287), bibliotecas de carbohidratos (véase, por ejemplo, Liang et al., Science, 274:1520-1522 (1996) y la Patente de EE.UU. 5.593.853), bibliotecas de moléculas orgánicas pequeñas (véase, por ejemplo, benzodiazepinas, Baum C&EN, Jan 18, page 33 (1993); isoprenoides, Patente de EE.UU. 5.569.588; tiazolidinonas y metatiazanonas, Patente de EE.UU. 5.549.974; pirrolidinas, Patentes de EE.UU. 5.525.735 y 5.519.134, compuestos de morfolino, Patente de EE.UU. 5.506.337, benzodiazepinas, y patente de EE.UU. N° 5.288.514).

D. Composiciones farmacéuticas

Las composiciones descritas en la presente memoria se pueden administrar por cualquier medio conocido en la técnica. Por ejemplo, las composiciones pueden incluir la administración a un sujeto por vía intravenosa, intradérmica, intraarterial, intraperitoneal, intralesional, intracraneal, intraarticular, intraprostical, intrapleurar, intratraqueal, intranasal, intravital, intravaginal, intarectal, tópica, intratumoral, intramuscular, intratecal, subcutánea, subconjuntiva, intravascular, mucosal, intrapericardial, intraumbilical, intraocular, oral, local, por inhalación, por inyección, por infusión, por infusión continua, por perfusión localizada, a través de un catéter, mediante un lavado, en una crema, o en una composición lipídica. La administración puede ser local, por ejemplo, a la mucosa pulmonar, o sistémica.

Las soluciones de los compuestos activos como base libre o sal farmacológicamente aceptable se pueden preparar en agua mezclada adecuadamente con un surfactante, tal como hidroxipropilcelulosa. Las dispersiones se pueden también preparar en glicerol, polietilenglicoles líquidos, y mezclas de los mismos y en aceites. En condiciones normales de almacenamiento y uso, estas preparaciones pueden contener un conservador para prevenir el crecimiento de microorganismos.

Las composiciones farmacéuticas se pueden suministrar mediante soluciones o pulverizadores intranasales o inhalables, aerosoles o inhaladores. Las soluciones nasales pueden ser soluciones acuosas diseñadas para ser administradas a los conductos nasales en gotas o pulverizadores. Las soluciones nasales se pueden preparar de modo que sean similares en muchos aspectos a las secreciones nasales. Por tanto, las soluciones nasales acuosas normalmente son isotónicas y ligeramente reguladas para mantener un pH de 5,5 a 6,5. Además, si se requiere, se pueden incluir en la formulación conservadores antimicrobianos, similares a los utilizados en preparaciones oftálmicas, y estabilizadores de fármaco apropiados. Se conocen diversas preparaciones nasales comerciales y pueden incluir, por ejemplo, antibióticos y antihistamínicos.

Las formulaciones orales pueden incluir excipientes como, por ejemplo, grados farmacéuticos de manitol, lactosa, almidón, estearato de magnesio, sacarina sódica, celulosa, carbonato de magnesio y similares. Estas composiciones toman la forma de soluciones, suspensiones, comprimidos, píldoras, cápsulas, formulaciones de liberación sostenida o polvos. Las composiciones farmacéuticas orales pueden comprender un diluyente inerte o portador comestible asimilable, o se pueden encerrar en cápsula de gelatina de cubierta dura o blanda, o se pueden comprimir en comprimidos, o se pueden incorporar directamente con el alimento de la dieta. Para administración terapéutica oral, los compuestos activos se pueden incorporar con excipientes y utilizados en forma de comprimidos ingeribles, comprimidos bucales, trociscos, cápsulas, elixires, suspensiones, jarabes, pastillas y similares. Tales composiciones y preparaciones deben contener al menos el 0,1% de compuesto activo. El porcentaje de las composiciones y preparaciones se puede, por supuesto, variar y puede estar convenientemente entre aproximadamente el 2 a aproximadamente el 75% del peso de la unidad, o preferiblemente entre el 25-60%. La cantidad de compuestos activos en tales composiciones es tal que se puede obtener una dosificación adecuada.

Para administración parenteral en una solución acuosa, por ejemplo, la solución debe ser adecuadamente regulada y el diluyente líquido primero hecho isotónico con suficiente solución salina o glucosa. Las soluciones acuosas, en particular, medios acuosos estériles, son especialmente adecuadas para la administración intravenosa, intramuscular, subcutánea e intraperitoneal. Por ejemplo, una dosificación se podría disolver en 1 ml de solución de NaCl isotónica y bien añadida a 1000 ml de fluido de hipodermoclasia o bien inyectada en el sitio de infusión propuesto.

Las soluciones inyectables estériles se pueden preparar incorporando los compuestos activos o constructos en la cantidad requerida en el solvente apropiado seguido de esterilización por filtración. Generalmente, las dispersiones se preparan incorporando los diversos ingredientes activos esterilizados en un vehículo estéril que contiene el medio de dispersión básico. Las técnicas de secado al vacío y secado por congelación, que producen un polvo del ingrediente activo más cualquier ingrediente deseado adicional, se pueden utilizar para preparar polvos estériles para la reconstitución de soluciones inyectables estériles. La preparación de soluciones más, o altamente, concentradas para la inyección directa también se contempla. El DMSO se puede utilizar como solvente para una penetración extremadamente rápida, suministrando altas concentraciones de los agentes activos a un área pequeña.

E. Regímenes de tratamiento

La divulgación proporciona métodos para tratar, prevenir y/o mejorar un trastorno pulmonar en un sujeto que lo necesite, opcionalmente en base a los métodos de diagnóstico y de predicción descritos en la presente memoria. El curso de tratamiento se determina mejor sobre una base individual dependiendo de las características particulares del sujeto y el tipo de tratamiento seleccionado. El tratamiento, tal como los descritos en la presente memoria, se puede administrar al

sujeto diariamente, dos veces al día, quincenalmente, mensualmente, o en cualquier base aplicable que sea terapéuticamente efectiva. El tratamiento se puede administrar solo o en combinación con cualquier otro tratamiento descrito en la presente memoria o conocido en la técnica. El tratamiento adicional se puede administrar simultáneamente con el primer tratamiento, en un momento diferente, o en un programa terapéutico completamente diferente (por ejemplo, el primer tratamiento puede ser diariamente, mientras que el tratamiento adicional es semanalmente).

La administración de una composición para mejorar la enfermedad pulmonar, por ejemplo, tratando la expresión elevada del gen MUC5B, puede ser una administración sistémica o localizada. Por ejemplo, el tratamiento de un sujeto que tiene un trastorno pulmonar puede incluir la administración de una forma inhalable o intranasal de agente anti-MUC5B (antagonista de MUC5B) diariamente o un programa de otra manera regular. El tratamiento puede ser solamente en una base de según sea necesario, por ejemplo, tras la aparición de síntomas de enfermedad pulmonar.

VI. Equipos

La divulgación proporciona equipos para la detección de marcadores de enfermedad pulmonar en un sujeto. El equipo puede ser para uso personal o proporcionado a profesionales médicos. El equipo puede ser un equipo para diagnosticar o pronosticar un trastorno pulmonar, o para monitorizar la progresión de la enfermedad o la eficacia del tratamiento.

El equipo puede incluir componentes para evaluar la expresión del gen MUC5B que comprenden, por ejemplo, un ácido nucleico capaz de detectar ARN MUC5B o un agente de enlace de proteína MUC5B, opcionalmente marcado. Un experto apreciará que la expresión del gen MUC5B se puede determinar midiendo el ARN o proteína MUC5B. El equipo puede además incluir recipientes de ensayo (tubos), soluciones reguladoras, o enzimas necesarios para llevar a cabo el ensayo de detección.

El equipo puede incluir componentes para determinar si el genoma del sujeto porta un gen MUC5B de variante genética, por ejemplo, un ácido nucleico que se hibrida específicamente con una secuencia de gen MUC5B de variante genética. Otros componentes en un equipo pueden incluir, componentes de ensayo de secuenciación de ADN, componentes de ensayo de genotipado Taqman®, Metaanálisis, uno o más sistema(s) de detección, una o más muestras de control o una combinación de los mismos. Los equipos pueden además incluir uno o más agentes donde al menos uno de los agentes es capaz de asociarse con el SNP rs35705950.

El equipo puede incluir componentes para examinar más de un marcador de enfermedad pulmonar. Por ejemplo, el equipo puede incluir agentes de detección de marcador, tales como cebadores o sondas específicos de marcador unidos a un arreglo dirigible. Los marcadores ejemplares incluyen los SNP en los genes MUC5B, o variantes genéticas en otros genes, por ejemplo, Proteína Surfactante A2, Proteína Surfactante B, Proteína Surfactante C, TERC, TERT, IL-1RN, IL-1 α , It-1 β , TNF, Linfotóxina α , TNF-RII, IL-10, IL-6, IL-12, IFN γ , TGF β , CR1, ACE, IL-8, CXCR1 o CXCR2. El nivel de expresión de los marcadores se puede detectar en lugar de o además de la secuencia genética. En este caso, los marcadores de enfermedad pulmonar útiles con expresión aberrante incluyen: Proteína Surfactante A, Proteína Surfactante D, KL-6/MUC1, CC16, CK-19, Ca 19-9, SLX, MCP-1, MIP-1a, ITAC, glutatión, péptido de procolágeno tipo III, sIL-2R, ACE, neopterina, beta-glucuronidasa, LDH, CCL-18, CCL-2, CXCL12, MMP7, y osteopontina. Los marcadores de enfermedad pulmonar adicionales pueden incluir los otros genes MUC, por ejemplo, MUC2, MUC5AC, y MUC6.

El equipo generalmente incluirá al menos un vial, tubo de ensayo, matraz, botella, jeringa u otro medio de recipiente, en el que el agente de pruebas, se puede hacer reaccionar o dividido en alícuotas adecuadamente. Los equipos pueden también incluir componentes para comparar los resultados tal como una muestra de control adecuada, por ejemplo un control positivo y/o negativo. El equipo puede también incluir un dispositivo de recogida para recoger y/o guardar la muestra del sujeto. El dispositivo de recogida puede incluir una torunda o aguja estéril (para recoger sangre), y/o un tubo estéril (por ejemplo, para guardar la torunda o una muestra de fluido corporal).

VII. Ejemplos

Ejemplo 1: Secuenciación de mucinas pulmonares, formadoras de gel y asociación con la enfermedad

Poblaciones de estudio: se identificaron y se fenotiparon sujetos con NIF o FPI. El diagnóstico de NII se estableció según criterios convencionales. Los sujetos elegibles tenían al menos 38 años de edad y tenían síntomas de NII durante al menos 3 meses. Se requirió una exploración de tomografía computarizada de alta resolución (HRCT) para mostrar la NII definida o probable según los criterios predefinidos, y se obtuvo una biopsia de pulmón quirúrgica en el 46% de los sujetos afectados. Las familias de FPI se definieron por la presencia de dos o más casos de NII definida o probable dentro de tres grados, con al menos un caso de NII establecido como FPI definido/probable. Los criterios de exclusión incluyeron la exposición significativa a agentes fibrogénicos conocidos o una etiología alternativa para ILD. Se adquirieron sujetos de control para el análisis genético (Fig. 1).

Análisis de enlace: se completó un cribado de enlace de amplio genoma en 82 familias multiplex utilizando un panel de enlace DeCode que consiste en un total de 884 marcadores con una distancia de inter-marcador promedio de 4.2 CM. El análisis de enlace no paramétrico multipunto se realizó utilizando Merlin, previamente descrito. Las puntuaciones LOD de Kong y Cos se calcularon utilizando la estadística Spares en un modelo exponencial; los intervalos de soporte se determinaron utilizando el método de una puntuación LOD hacia abajo.

Mapeo fino del cromosoma 11: para interrogar la región enlazada en el extremo p del cromosoma 11 (8.4 Mb unido por rs702966 y rs1136966), se realizó el mapeo fino genotipando 306 SNP de etiquetado en 145 casos no relacionados de NIF, 152 casos de FPI, y 233 controles Caucásicos. Las pruebas de asociación que comparan los casos de NIF y los casos de FPI con los controles se calcularon bajo un modelo aditivo para el alelo menor.

Resequenciación de MUC2 y MUC5AC: se diseñaron pares de cebadores para generar amplicones solapantes para la resecuenciación del promotor proximal y la mayoría de los exones de MUC2 y MUC5AC sobre secuencias enmascaradas para elementos repetitivos, los SNP, y la homología con otras regiones del genoma.

Cribado genético de mucinas formadoras de gel, expresadas en el pulmón: se realizó un estudio de asociación de control de casos en una población independiente de NIF (N=83), FPI esporádica (N=492), y sujetos de control (N=322) (Tabla 2) utilizando etiquetado y otros SNP localizados a través de los genes de mucina expresadas en el pulmón, formadoras de gel en el cromosoma 11. Se genotiparon exitosamente 175 SNP utilizando los ensayos Sequenom iPLEX, y se utilizó haploview para probar los SNP para la asociación alélica con NIF y FPI. Para aquellos SNP que permanecieron significativos después de la corrección de Bonferroni, se estimaron las relaciones de probabilidad bajo un modelo aditivo para el alelo raro después del ajuste para la edad y el género mediante regresión logística. Se calcularon pruebas de bondad de ajuste Chi-cuadrada para evaluar la evidencia de las explicaciones de modelo de enfermedad para desviaciones genotípicas del Equilibrio de Hardy Weinberg (HWE) entre los casos. Para el SNP más altamente asociado, se utilizó el modelado de enlace y asociación en pedigrees para probar si, en las familias de enlace originales, el SNP se enlazaba al locus de la enfermedad, estaba en desequilibrio de enlace con el locus de la enfermedad, y podría explicar la señal de enlace.

Se produjo una fuerte evidencia para el enlace basado en las 82 familias de NIF en el cromosoma 11 donde la puntuación LOD multipunto máxima fue 3,3 ($p=0.00004$, D11S1318; Fig. 3). El intervalo de soporte de 1-LOD para esta región enlazada se unió mediante los marcadores D11S4046 y D11S1760, que abarcan 3,4 Mb. Puesto que D11S4046 fue el marcador telomérico más tipificado, la región de interés era inclusiva del extremo p del cromosoma 11. Dentro de la región más grande de 8,4 Mb, se seleccionaron 306 SNP de etiquetado para mapeo fino en un análisis de asociación de control de casos (145 casos de NIF, 152 casos de FPI, y 233 controles). Las pruebas de asociación alélica revelaron 7 SNP dentro del gen mucina 2 (MUC2) significativamente asociados con bien NIF o bien FPI. MUC2 está contenido en una región genómica que alberga 4 genes de mucina formadora de gel (telómero a centrómero: MUC6, MUC2, MUC5AC, y MUC5B). Aunque hay reportados puntos calientes de recombinación localizados entre MUC6 y MUC2, y dentro de la porción proximal de MUC5B, los marcadores dentro de MUC2 y MUC5AC exhiben fuerte desequilibrio de enlace (LD) 17. Por tanto, MUC2 y MUC5AC se seleccionaron para la resecuenciación utilizando los cebadores de oligonucleótidos. El análisis de resecuenciación identificó 330 variantes genéticas en MUC2 y 195 variantes genéticas en MUC5AC. Las pruebas de asociación alélica entre estas variantes genéticas y el estado de enfermedad produjeron 7 SNP independientes tanto en MUC2 como en MUC5AC significativamente asociados con el estado de enfermedad de bien NIF o bien FPI.

Los inventores diseñaron un cribado genético para la variación genética común a través de la región genómica que contiene los 3 genes de mucina formadora de gel expresados en el pulmón (MUC2, MUC5AC, y MUC5B) en una población independiente de sujetos con NIF (NIF=83 y FPI=492) y controles ($n=322$) (Fig. 1, Tabla 2). Se observó que 19 SNP independientes estaban significativamente asociados mediante la prueba alélica con cualquiera o ambas de NIF o FPI después de la corrección de Bonferroni para comparaciones múltiples (Tabla 1). De estos 19 SNP, 6 ocurrieron en MUC2, uno en la región intergénica de MUC2-MUC5AC, 4 en MUC5AC, 3 en la región intergénica de MUC5AC-MUC5B, y 5 en el promotor de MUC5B putativo, dentro de 4kb del sitio 18, 19 de inicio de la transcripción de MUC5B (Tabla 1).

De importancia, se encontró que un SNP en el promotor putativo de MUC5B, 3kb aguas arriba del sitio (rs35705950) de inicio de la transcripción tiene el efecto más importante tanto sobre NIF como sobre FPI. El alelo menor de este SNP estaba presente en una frecuencia de 33,8% en los casos de NIF, 37,5% en los casos de FPI, y 9,1% entre los controles (asociación alélica; NIF $P=1,2 \times 10^{-15}$, FPI $P=2,5 \times 10^{-37}$). En particular, las frecuencias de genotipo para rs35705950 eran consistentes con HWE en los controles, pero no entre los casos de FPI ($P=6,0 \times 10^{-11}$) y casi así entre los casos de NIF ($P=0,11$). Comparando las frecuencias de genotipo observadas en los casos y los controles con las esperadas si rs35705950 es un locus de riesgo real, los datos demuestran que estas frecuencias de genotipo son consistentes con un efecto genotípico aditivo en el riesgo de enfermedad conferido por rs35705950 ($P=0,88$ y $P=0,77$, respectivamente para NIF e FPI para rechazar el efecto aditivo). Además, la frecuencia del alelo de la enfermedad y las estimaciones de penetrancia sugieren un modelo de enfermedad similar tanto para NIF como para FPI. La relación de probabilidad para la enfermedad para sujetos heterocigotos y homocigotos para el alelo raro de este SNP fue de 6,8 (IC 95% 3,9-12,0) y 20,8 (IC 95% 3,8-113,7) para NIF, and 9,0 (IC 95% 6,2-13,1) y 21,8 (IC 95% 5,1-93,5) para FPI (Tabla 1). Para asegurar que este SNP no estuviera etiquetando otro SNP en la región promotora de MUC5B, la región de 4kb se resecuenció aguas arriba del sitio de inicio de la transcripción de MUC5B en 48 casos de FPI y 48 controles (Tabla 3). Se observó que 34 variantes genéticas pero ninguna tenía un valor de LD de r^2 por pares con rs35705950 por encima de 0,2 (Tabla 4). Finalmente, entre las familias de enlace originales, se encontró que rs35705950 estaba tanto enlazado a ($P=0,04$) como en desequilibrio de enlace con ($P=1,5 \times 10^{-9}$) el locus de la enfermedad. Aunque hay alguna evidencia de otras variantes enlazadas en la región ($P=0,054$), estos resultados verifican la relevancia de este SNP para la enfermedad en estas familias.

Tabla 1. Resultados de la asociación genotípica asumiendo un modelo aditivo del cribado genético de las mucinas formadoras de gel, expresadas en el pulmón en sujetos con NII (NIF=83 e FPI=492) y controles (n=322).

SNP	Carga de Aminoácido de Nucleótido	Región de Mucina	Posición Hg19	Frecuencia del Alelo Menor			Prueba de Asociación Genotípica por el Grupo Clasificado			
				NIF (n=83)	FPI (n=492)	Controles (n=322)	Relación de Probabilidad (IC 95%)	Valor P	Relación de Probabilidad (IC 95%)	Valor P
rs10902081	C/T	MUC2 Int7	1079809	37,2	38,6	47,9	0,6(0,4-0,9)	0,011	0,7(0,5-0,8)	4,3x10 ⁻⁴
rs7127117*	T/C	MUC2 Int7	1079879	49,3	60	47,4	1,0 (0,7-1,5)	0,826	1,6(1,3-2,0)	6,9x10 ⁻⁵
rs41453346	C/T Tyr426Tyr	MUC2 Ex10	1080894	5	65	22	1,9 (0,8-4,3)	0,124	2,8(1,6-5,2)	0,001
rs41480348	G/ATh618Thr	MUC2 EX15	1082605	84	65	12,1	0,7(0,4-1,2)	0,188	0,5(0,4-0,8)	0,001
r37934606*	C/T	MC2 Int31	1093945	49,4	54	40,5	1,4(1,0-2,0)	0,055	1,7(1,4-2,2)	3,8x10 ⁻⁶
rs10902089*	A/G	MUC2 Int31	1094357	57,9	58,8	48,5	1,5(1,0-2,1)	0,031	1,5(1,2-1,9)	2,9x10 ⁻⁴
rs9667239	C/T	MUC2-5AC Intergénico	1143101	22,5	21	12,5	2,2(1,4-3,6)	0,001	1,9(1,4-2,7)	5,6x10 ⁻⁵
rs55846509	G/AAArg47Gn	MUC5AC Ex2	1154294	31	5,5	1,6	1,7 (0,6-5,1)	0,316	3,6(1,7-7,3)	0,001
rs28403537	C/TAla497Va	MUC5AC Ex12	1161315	89	13	34	2,7(1,3-5,3)	0,006	4,6(2,8-7,6)	32x10 ⁻⁹
MUC5AC025447*	C/T	MUC5AC Int26	826476*	20,1	21	13,8	1,6(1,0-2,5)	0,053	1,6(1,2-2,2)	0,003
rs36288961	G/T	MUC5AC Int46	1220462	28,8	26,6	15,9	2,2(1,4-3,5)	3,2x10 ⁻⁴	2,0(1,5-2,6)	3,7x10 ⁻⁶
rs35671223	C/T	MUC5AG5B Intergénico	1227069	42,6	42,4	33,4	1,4(1,0-2,0)	0,05	1,5(1,2-1,9)	0,001
rs28654232	C/T	MUC5AG5B Intergénico	1229227	21,6	22,8	32,9	0,6(0,4-0,9)	0,009	0,6(0,5-0,8)	1,1x10 ⁻⁴
rs34595903*	C/T	MUC5AG5B Intergénico	1230393	21,5	23,3	34,8	0,5(0,3-0,7)	0,001	0,5(0,4-0,7)	2,4x10 ⁻⁶
rs2672794	C/T	MUC5B Prm	1241005	27,2	27,5	40,4	0,5 (0,3-0,8)	0,001	0,5(0,4-0,7)	1,9x10 ⁻⁴
rs35705950	G/T	MUC5B Prm	1241221	33,8	37,5	9,1	6,2(3,7-104) ¹²	3,7x10 ⁻¹²	8,3(5,8-11,9)	4,6x10 ⁻³
rs35619543*	G/T	MUC5B Prm	1242250	40,3	39	23,8	2,4(1,6-3,6)	3,3x10 ⁻⁵	2,1(1,6-2,8)	1,5x10 ⁻⁸

rs12804004	G/T	MUC5B Prm	1242299	39,2	39,4	48,9	0,6(0,4-0,9)	0,019	0,6(0,5-0,8)	1,2x10 ⁻⁴
rs868903*	T/C	MUC5B Prm	1242690	65,4	61	49,5	1,8(1,3-2,6)	0,001	1,6(1,3-2,1)	2,8x10 ⁻⁵

*Para estos SNP, el ADN estaba disponible para 304 controles. ** Posición de nucleótido en base a NW_001838016.1

Tabla 2. Características demográficas de los sujetos en los análisis de re-secuenciación y cribado genético de mucinas.

	Re-secuenciación de sujetos			Cribado Genético de Sujetos con Mucinas Formadoras de Gel Expresadas en el Pulmón		
	NIF	FPI	Control	NIF	FPI	Control
Número de sujetos	69	96	54	83	492	322*
Género Masculino	41 (60%)	61 (64%)	18 (34%)	44 (53.0%)	352 (71.5%)	147 (45.7%)
Caucásicos	68 (99%)	89 (93%)	53 (98%)	83 (100%)	492 (100%)	322 (100%)
Edad en el momento del diagnóstico	66 ± 10	65 ± 8	68 ± 8	66.3±11.2	67.2 ± 8.1	60.3 ± 12.6
Ha fumado alguna vez	44 (64%)	71 (74%)	25 (47%)	46 (56.8%)	342 (69.9%)	245 (76.6%)
*Se incluyeron 325 sujetos de control en los análisis de asociación alélica pero solamente 322 en los análisis de regresión genotípica ya que las variables demográficas necesarias para la regresión faltaban para 3 sujetos. Adicionalmente, en algunos multiplexes de genotipado para las mucinas formadoras de gel expresadas en el pulmón, 18 de los 322 controles no se clasificaron debido a la falta de disponibilidad del ADN.						

ES 2 636 671 T3

Tabla 3: Oligos utilizados en la resecuenciación del promotor de MUC5B.

Amplificación de Promotor de MUC5B	Cebador Directo 5'>3'	Cebador Inverso 5'>3'	Tamaño del Amplificación (bp)	Coordenadas de Hg19
MUC5B-Ceb-1	GGTTCTCCTTGTCTTGCAGCCCT (SEQ ID NO:1)	ATGGGCTCTTGGTCTGCTCAGAG (SEQ ID NO:2)	616	Chr11:1239997-1240612
MUC5B-Ceb-2	GGGCCTGGCTCTGAGTACACATCCT (SEQ ID NO:3)	AAGGAAAGGGACACAGCCGGTTCC (SEQ ID NO:4)	644	Chr11:1240556-1241199
MUC5B-Ceb-3	GGGTCCCCATTATGGCAGGATT (SEQ ID NO:5)	TTTCTCCATGGCAGAGCTGGGACC (SEQ ID NO:6)	601	Chr11:1240957-1241557
MUC5B-Ceb-4	CTAGTGGGAGGGACGAGGGCAAAGT (SEQ ID NO:7)	CTCGTGGCTGTGACTGCACCCAG (SEQ ID NO:8)	610	Chr11:1241386-1241995
MUC5B-Ceb-5	TTGGCTAAGGTGGGAGACCT (SEQ ID NO:9)	AGCTTGGGAATGTGAGAACG (SEQ ID NO:10)	700	Chr11:1241791-1242490
MUC5B-Ceb-6	CATGAGGGGTGACAGGTGGCAA (SEQ ID NO:11)	CCCGCGTTTGTCTTTCTGAAGTT (SEQ ID NO:12)	676	Chr11:1242392-1243067
MUC5B-Ceb-7	GGTCAGAAGCTTTGAAGATGGGC (SEQ ID NO:13)	CTTGTCGAATGCCAGCCCTGATC (SEQ ID NO:14)	607	Chr11:1242985-1243591
MUC5B-Ceb-8	CTGCCAGGGTTAATGAGGAG (SEQ ID NO:14)	GGATCAGGAAGGATTGCGAG (SEQ ID NO:16)	663	Chr11:1243491-1244153
MUC5B-Ceb-9	AGGCAGGCTGGCTGACCACTGTTT (SEQ ID NO:17)	CGTGAAGACAGCATCGAGAGGGG (SEQ ID NO:18)	501	Chr11:1243966-1244466
MUC5B-Ceb-5 Sec Pr.	TTGGCTAAGGTGGGAGACCT (SEQ ID NO:19)			Chr11:1241791-1241810

Tabla 4. SNP identificados en la resecuenciación del promotor de MUC5B

Posición de Hg19	Nombre del SNP	Cambio de base	SEC ID N°:	Secuencia Flanqueante
240338	rs2672792	T/C	20	GTCACCTGCCAGGTCCCCGAGGCC[T/C]GGAACACCTTCTGCTGGGCCACC
240485	rs72636989	G/A	21	CCACCCAGGAGTTGGGGGGCCCCGT[G/A]CCAGGGAGCAGGAGGCTGCCGAGG
240925	Muc5B-Prm1	C/T	22	GTGGCCCTGATCACTGGTGCCTGGA[C/T]GGCCTCTGAAGGGGTCTGTGGGGTC
241005	rs2672794	C/T	23	AACCCCCCTCGGGTTCTGTGTGGT[C/T]AGGCCGCCCTTTGTCTCCACTGCC
241221	rs35705950	G/T	24	TTTCTTCTTTATCTTCTGTTTTCAGC[G/T]CCTTCAACTGTGAAGAAGTGA
241361	MUC5B-Prm2	A/G	25	TGCCCGGACCCAGCCAGTTCCCA[A/G]TGGGCCCTCTGCCGGGGAGGTGC
241762	MUC5B-Prm3	C/T	26	GGTGGGCATCGGCTTGTGAGCTGGAGCCG[C/T]GGGCAGGGAGGGGGATGTACAGAG
241821	rs11042491	G/A	27	GGCTAAGGTGGGAGACCTGGGCGGGTGC[G/A]TCGGGGGGACGTCTGCAGCAGAGGC
241848	rs2735726	T/C	28	TGCGTCGGGGGGACGTCTGCAGCAGAGGCC[T/C]GGGCAGCAGGCACACCCCTCTGCCAG
241993	MUC5B-Prm4	G/A	29	GGGGCTGGGTGCAGTACAGCCAC[G/A]AGCCCAGGGGTGGGACTCTGGCC
242092	MUC5B-Prm5	C/T	30	CCCCCACCAGCTGCCGTGCTGCAG[C/T]GGGTCTACCGGCCTGGATGTGAAA
242101	MUC5B-Prm6	C/T	31	CCGTGCCGTGCTGCAGCGGGTCTAC[C/T]GGCCTGGATGTGAAAGAGAGCTTG
242227	rs11042646	C/T	32	AGTCCCGGAAGTGAGCGGGGAGCTA[C/T]GCTGAGATCTGGGAGACCCCTGTC
242244	rs55974837	C/T	33	GGGAGCTACGCTGAGATCTGGGAGA[C/T]CCCCGCCCCACCCAGGTACAGG
242250	rs35619543	G/T	34	TACGCTGAGATCTGGGAGACCCCT[G/T]CCCCCACCAGGTACAGGGCCAGG
242299	rs12804004	T/G	35	GCAGAAGCCCGAGGTGTGCCCTGAG[T/G]TAAAGAAACCGTCACAAAGAACA
242472	rs56031419	G/A	36	TGTCTCCGCCCTCCATCTCCAGAAC[G/A]TTCTCACATTCCCAAGCTGAACCC
242508	rs868902	C/A	37	CCCAAGCTGAACCCCTGTCCCATG[C/A]AACACCAGCTCACCATCCCTCTGCC
242567	MUC5B-Prm7	C/T	38	GGCGCCACCGTCCACACTCCGTCT[C/T]TGGGGTTTCATGACTCCAGGGGAG
242599	MUC5B-Prm8	G/A	39	TTTCATGACTCCAGGGGACACAC[G/A]AGTGGCCCTCTGCCTTTGTCTCTC
242607	MUC5B-Prm9	C/T	40	CTCCAGGGGACACACGAGTGGCC[C/T]CTCCTGCCTTTGTCTCTGTGTTCA
242690	rs868903	C/T	41	CCCCCATGGAGCAGCTGGGCCAGCC[C/T]CTCCTTTTCAGCGCTGAACCGTAT
242910	MUC5B-Prm10	G/A	42	ACCCCAACAGCAGGGACAGGGCTCC[G/A]GGTCCCCACGTCTCTGCCAACACTT
242977	MUC5B-Prm11	G/A	43	CTTGATCCCCGCCATCCTATTGAGC[G/A]TGAGACAGGTGAGAAGCTTTGAAG
243218	MUC5B-Prm12	G/A	44	GTCTGCGCCACGGAGCATTCAGGAC[G/A]CTGGTGACCAGGGAGCCAGGAGGT

(Continuación)

Posición de Hg19	Nombre del SNP	Cambio de base	SEC ID N°:	Secuencia Flanqueante
243378	rs885455	A/G	45	CGTCAAGGAGGTTTACACATAGCCCCC[A/G]GGAAGCCCCACCCGACACAGCCGGA
243391	rs885454	G/A	46	TTTACCACATAGCCCCCGGAAGCCCCACCC[G/A]ACACAGCCGAGGTGCTAGGCTTC
243409	MUC5B-Prm13	T/C	47	CCCACCCGACACAGCCGAGGTGC[T/C]AGGCTTCTGCGGCTCCACCTGGG
243911	MUC5B-Prm14	G/A	48	GGACCCATGGTCAGTGGCTGGGGGT[G/A]CTGCCAGAGGCTGGGATTCCTTC
244060	rs7115457	G/A	49	GCCATCTAGGACGGGTGCCAGGTGG[G/A]GTAGCCCTTCTCTCCCTCCGATT
244080	rs7118568	C/G	50	GGTGGGGTAGGCCCTTCTCTCCCTT[C/G]CGATTCTCAGAAGCTGCTGGGGGTG
244197	rs56235854	G/A	51	AGCCCCCTCCCCGAGAGCAAACACAC[G/A]TGGCTGGAGCGGGGAAGAGCATGGTGC
244219	rs2735738	T/C	52	CACGTGGCTGGAGCGGGGAAGAGCA[T/C]GGTGCCCTGCGTGGCTGGCTGGC
244438	MUC5B-Prm15	C/T	53	GCCGCAGGCAGGTAAGAGCCCCCA[C/T]TCCGCCCTCTCGATGCTGTCTT

A continuación, se analizó la relación entre el SNP rs35705950 y los 18 de otros SNP significativamente asociados con NII. Probando el LD por pares entre estos SNP mediante la estadística r^2 , se encontró que 10 de los 18 SNP exhibían bajo nivel de LD ($r^2=0,15-0,27$) con rs35705950 entre los casos de FPI, sugiriendo que la significación de estos SNP es debida a LD con rs35705950 (Fig. 3). Utilizando modelos de regresión logística genotípica para ajustar los efectos de rs35705950, los inventores observaron que los coeficientes y los valores P correspondientes eran sustancialmente reducidos para todos los 18 SNP que se asociaron previamente con NIF y/o FPI (Tabla 5). Después del control para rs35705950, solamente un SNP conservó la significación nominal para FPI (rs41480348, $P=0,04$). Se demostró que la significación del SNP rs35705950 fue en gran parte no afectada por el ajuste para cualquiera de los 18 SNP probados (el valor P para todos los modelos de SNP fue menor que $1,7 \times 10^{-9}$ para NIF y $1,1 \times 10^{-24}$ para FPI; Tabla 5). Estos resultados demuestran un fuerte efecto independiente del SNP rs3570590 tanto sobre NIF como sobre FPI.

Tabla 5. Modelos de regresión logística genotípica para los 19 SNP significativos en el cribado de mucinas formadoras de gel expresadas en el pulmón solos, y después del ajuste para rs35705950, en pacientes con FPI o NIF.

#Modelo	SNP	Modelo de SNP único de FPI		NIF rs35705950		Modelo de SNP único de NIF		NIF rs35705950	
		Relación de Probabilidad (IC 95%)	Valor P	Relación de Probabilidad (IC 95%)	Valor P	Relación de Probabilidad (IC 95%)	Valor P	Relación de Probabilidad (IC 95%)	Valor P
1	rs10902081	0,7 (0,5-0,8)	$4,3 \times 10^{-4}$	0,9 (0,7-1,2)	0,429	0,6 (0,4-0,9)	0,011	0,8 (0,5-1,2)	0,25
	rs35705950	x	x	8,3 (5,7-11,9)	$1,5 \times 10^{-29}$	x	x	5,9 (3,5-10,1)	$6,6 \times 10^{-11}$
2	rs7127117	1,6 (1,3-2,0)	$6,9 \times 10^{-5}$	1,1 (0,8-1,4)	0,509	1,0 (0,7-1,5)	0,826	0,7 (0,4-1,1)	0,094
	rs35705950	x	x	7,9 (5,4-11,6)	$1,3 \times 10^{-25}$	x	x	6,3 (3,5-11,4)	$7,3 \times 10^{-10}$
3	rs41453346	2,8 (1,6-5,2)	0,001	1,1 (0,6-2,2)	0,72	1,9 (0,8-4,3)	0,124	1,2 (0,5-3,0)	0,653
	rs35705950	x	x	8,1 (5,6-11,9)	$2,7 \times 10^{-28}$	x	x	6,1 (3,6-10,3)	$1,2 \times 10^{-11}$
4	rs41480348	0,5 (0,4-0,8)	0,001	0,6 (0,4-1,0)	0,04	0,7 (0,4-1,2)	0,188	0,9 (0,5-1,7)	0,75
	rs35705950	x	x	7,9 (5,5-11,3)	$2,1 \times 10^{-29}$	x	x	6,1 (3,6-10,2)	$1,0 \times 10^{-11}$
5	rs7934606	1,7 (1,4-2,2)	$3,8 \times 10^{-5}$	1,0 (0,7-1,3)	0,876	1,4 (1,0-2,0)	0,055	0,9 (0,6-1,3)	0,473
	rs35705950	x	x	8,7 (5,8-12,9)	$1,4 \times 10^{-26}$	x	x	6,7 (3,8-11,9)	$7,5 \times 10^{-11}$
6	rs10902089	1,5 (1,2-1,9)	$2,9 \times 10^{-4}$	0,9 (0,7-1,2)	0,69	1,5 (1,0-2,1)	0,031	1,0 (0,7-1,6)	0,813
	rs35705950	x	x	8,3 (5,6-12,2)	$1,3 \times 10^{-26}$	x	x	6,1 (3,6-10,5)	$6,2 \times 10^{-11}$
7	rs9667239	1,9 (1,4-2,7)	$5,6 \times 10^{-5}$	0,8 (0,5-1,2)	0,3	2,2 (1,4-3,6)	0,001	1,1 (0,6-2,0)	0,668
	rs35705950	x	x	8,9 (6,0-13,3)	$8,2 \times 10^{-27}$	x	x	5,8 (3,3-10,2)	$6,0 \times 10^{-10}$
8	rs55846509	3,6 (1,7-7,3)	0,001	1,0 (0,5-2,3)	0,96	1,7 (0,6-5,1)	0,32	0,8 (0,3-2,5)	0,706
	rs35705950	x	x	8,3 (5,7-12,1)	$2,7 \times 10^{-28}$	x	x	6,4 (3,8-10,7)	$4,8 \times 10^{-12}$
9	rs28403537	4,6 (2,8-7,6)	$3,2 \times 10^{-9}$	1,5 (0,8-2,6)	0,2	2,7 (1,3-5,3)	0,006	0,8 (0,3-1,8)	0,53
	rs35705950	x	x	7,6 (5,2-11,2)	$1,1 \times 10^{-24}$	x	x	6,7 (3,8-11,8)	$4,7 \times 10^{-11}$
10	MUC5AC-025447	1,6 (1,2-2,2)	0,003	1,1 (0,8-1,6)	0,49	1,6 (1,0-2,5)	0,053	1,4 (0,8-2,4)	0,19
	RS35705950	x	x	7,7 (5,3-11,2)	$3,1 \times 10^{-27}$	x	x	6,0 (3,5-10,3)	$4,7 \times 10^{-11}$
11	rs35288961	2,0 (1,5-2,6)	$3,7 \times 10^{-5}$	1,1 (0,8-1,5)	0,58	2,2 (1,4-3,5)	$3,2 \times 10^{-4}$	1,3 (0,7-2,1)	0,384
	rs35705950	x	x	7,9 (5,4-11,5)	$6,6 \times 10^{-27}$	x	x	5,7 (3,3-10,0)	$1,3 \times 10^{-9}$
12	rs35671223	1,5 (1,2-1,9)	0,001	0,9 (0,7-1,2)	0,46	1,4 (1,0-2,0)	0,05	0,9 (0,6-1,4)	0,61
	rs35705950	x	x	8,5 (5,8-12,4)	$1,1 \times 10^{-28}$	x	x	6,3 (3,6-10,9)	$5,4 \times 10^{-11}$
13	rs28654232	0,6 (0,5-0,8)	$1,1 \times 10^{-4}$	0,9 (0,7-1,1)	0,29	0,6 (0,4-0,9)	0,009	0,7 (0,5-1,1)	0,167
	rs35705950	x	x	8,0 (5,5-11,5)	$5,7 \times 10^{-29}$	x	x	5,7 (3,4-9,6)	$5,8 \times 10^{-11}$
14	rs34585903	0,5 (0,4-0,7)	$2,4 \times 10^{-4}$	0,8 (0,6-1,1)	0,116	0,5 (0,3-0,7)	0,001	0,6 (0,4-1,0)	0,041
	rs35705950	x	x	7,4 (5,1-10,8)	$7,0 \times 10^{-25}$	x	x	5,1 (3,0-8,6)	$1,7 \times 10^{-9}$
15	rs2672794	0,5 (0,4-0,7)	$1,9 \times 10^{-7}$	0,9 (0,7-1,2)	0,442	0,5 (0,3-0,8)	0,001	0,7 (0,4-1,1)	0,152
	rs35705950	x	x	8,0 (5,5-11,6)	$2,5 \times 10^{-27}$	x	x	5,5 (3,2-9,3)	$3,2 \times 10^{-10}$
16	rs35619543	2,1 (1,6-2,8)	$1,5 \times 10^{-5}$	1,3 (0,9-1,7)	0,145	2,4 (1,6-3,6)	$3,3 \times 10^{-5}$	1,3 (0,8-2,1)	0,296
	rs35705950	x	x	7,6 (5,2-11,2)	$7,0 \times 10^{-25}$	x	x	6,1 (3,4-10,9)	$6,8 \times 10^{-10}$
17	rs12804004	0,6 (0,5-0,8)	$1,2 \times 10^{-4}$	0,8 (0,6-1,0)	0,07	0,6 (0,4-0,9)	0,019	0,7 (0,5-1,1)	0,159
	rs35705950	x	x	7,9 (5,5-11,3)	$6,4 \times 10^{-29}$	x	x	5,9 (3,5-10,0)	$3,6 \times 10^{-11}$
18	rs868903	1,6 (1,3-2,1)	$2,8 \times 10^{-5}$	1,0 (0,8-1,4)	0,753	1,8 (1,3-2,6)	0,001	1,4 (0,9-2,0)	0,145
	rs35705950	x	x	7,8 (5,3-11,5)	$8,6 \times 10^{-26}$	x	x	5,6 (3,2-9,6)	$4,4 \times 10^{-10}$

Ejemplo 2: El polimorfismo rs35705950 de un solo nucleótido da como resultado una expresión incrementada del gen MUC5B

El alelo G de tipo silvestre del SNP rs35705950 se conserva entre las especies de primates. El SNP está directamente 5' a una región altamente conservada entre las especies de vertebrados, y está en el medio de la secuencia que se predice que esté implicada en la regulación del gen MUC5B. Un análisis bioinformático del efecto del SNP rs35705950 predice una alteración de un sitio de enlace de E2F y la creación de al menos dos nuevos sitios de enlace (por ejemplo, HOX9 y PAX-2).

En base a estos análisis, se examinó el efecto de rs35705950 en la expresión del gen MUC5B. En el tejido pulmonar de 33 sujetos con FPI y 47 sujetos no afectados, la RT-PCR cuantitativa reveló que la expresión del gen MUC5B estaba sobre regulada 14,1 veces entre los sujetos con FPI en comparación con los sujetos no afectados ($P=0,0001$, Fig. 4A). Se observó un incremento de 37,4 veces en la expresión de MUC5B entre los sujetos no afectados que portan al menos una copia del alelo variante en comparación con los sujetos de tipo silvestre homocigotos ($P=0,0003$, Fig. 4B). Por el contrario, no se observó diferencia significativa en la expresión del gen MUC5B entre los sujetos con FPI con al menos un alelo variante de rs35705950 (Fig. 4C). Fumar, un factor de confusión potencial de la expresión de MUC5B, pareció tener poco efecto sobre la asociación entre el alelo variante rs35705950 y la expresión de MUC5B entre ya sea los sujetos no afectados o afectados por FPI (Figs. 4B y 4C).

La tinción inmunohistoquímica de MUC5B en el tejido pulmonar mostró tinción citoplásmica en células columnares secretoras de los bronquios y los bronquiolos proximales más grandes ($>200\text{ }\mu\text{m}$) en casos de FPI y controles (Fig. 5A). En sujetos con FPI, se observaron regiones de densa acumulación de MUC5B en áreas de aplanamiento microscópicas e implicaron la tinción parcheada del revestimiento epitelial metaplástico de los quistes de panal (Fig. 5B), así como los tapones de mucosa dentro de los quistes (Fig. 5C). No se observaron diferencias obvias en las características de tinción de MUC5B en casos de FPI con el polimorfismo del promotor de MUC5B.

Los sujetos con FPI tienen significativamente más expresión del gen pulmonar MUC5B que los controles, y la proteína MUC5B se expresa en lesiones patológicas de FPI. Los presentes resultados muestran que el riesgo de desarrollar NIF o FPI está sustancialmente correlacionado con el polimorfismo del promotor rs35705950, que causa expresión de MUC5B incrementada. En conjunto, los datos muestran que la expresión de MUC5B en el pulmón desempeña un papel en la patogénesis de enfermedad pulmonar.

En base a la relación entre el SNP y la producción en exceso de MUC5B, demasiado MUC5B puede perjudicar la defensa del hospedador mucosal a la lesión pulmonar excesiva de sustancias inhaladas, y con el tiempo, conducir al desarrollo de NII. Además del SNP del promotor de MUC5B, las exposiciones comunes y los procesos biológicos básicos pueden influenciar ya sea la expresión o la eliminación de MUC5B. Por ejemplo, la expresión de MUC5B puede ser incrementada en el pulmón por el humo de cigarrillo, acroleína, estrés oxidativo, IL-6, IL-8, IL-13, IL-17, 17β -estradiol, nucleótidos extracelulares, o cambios epigenéticos que alteran la metilación del ADN o la estructura de la cromatina. Además, la eliminación del moco pulmonar es dependiente de movimiento ciliar efectivo, la hidratación adecuada de la capa de líquido periciliar y una tos intacta. Independientemente de la causa, los presentes resultados indican que el exceso de MUC5B puede comprometer la defensa del hospedador mucosal, reducir la eliminación del pulmón de las partículas inhaladas, sustancias químicas disueltas y microorganismos. Dada la importancia de las exposiciones ambientales, tales como asbestos, sílice y otros contaminantes en el desarrollo de otras formas de enfermedad pulmonar intersticial, es lógico especular que las partículas inhaladas comunes, tales como las asociadas con el humo de cigarrillo o la contaminación del aire, podría conducir a o contribuir a la lesión intersticial exagerada en individuos quienes tienen defectos en la defensa del hospedador mucosal.

Además, el exceso de MUC5B en los bronquiolos respiratorios puede interferir con la reparación alveolar. La lesión alveolar puede conducir al colapso de las unidades broncoalveolares y esta lesión pulmonar focal es reparada a través de la re-epitelización del alveolo por las células epiteliales alveolares de tipo II. Por tanto, MUC5B puede impedir la reparación alveolar ya sea interfiriendo con la interacción entre las células epiteliales alveolares de tipo II y la matriz subyacente, o interfiriendo con las propiedades de tensión superficial del surfactante. Cualquier fallo de re-epitelizar la lámina basal del alveolo o la biología de surfactante subóptima podría incrementar el colapso en curso y la fibrosis de unidades broncoalveolares adyacentes, y finalmente dar como resultado NII.

Las lesiones de FPI son espacialmente heterogéneas, sugiriendo que FPI es multifocal, originándose en unidades broncoalveolares individuales. Puesto que el SNP rs35705950 ocurre en la región promotora de MUC5B y se predice que altera los sitios de unión del factor de transcripción, la producción ectópica de MUC5B en células o ubicaciones que causan lesión a la unidad broncoalveolar puede ser un agente causante. Las variantes genéticas no cribadas (especialmente en las regiones de mucina repetitivas inaccesibles) pueden estar en desequilibrio de enlace con el SNP del promotor de MUC5B y afectar a la función de otras mucinas pulmonares.

Las presentes observaciones proporcionan un enfoque clínico novedoso para los trastornos pulmonares tales como NII. La invocación de las mucinas de las vías respiratorias secretadas en la patogénesis de la fibrosis pulmonar sugiere que el espacio aéreo desempeña un papel en la patogénesis de NII. Aunque que el SNP (rs35705950) en el promotor de MUC5B se puede utilizar para identificar individuos en riesgo de desarrollar NII, la observación de que la biología de la mucina es

importante en la etiología de NII reorienta el enfoque de los estudios patogénicos y terapéuticos en la enfermedad pulmonar intersticial a las mucinas pulmonares y al espacio aéreo. Además, las causas genéticas de NII (por ejemplo, MUC5B, Proteína Surfactante C, Proteína Surfactante A2, y los dos genes de telomerasa TERC y TERT) proporcionan entendimiento en las manifestaciones clínicas únicas de este complejo proceso de enfermedad, y consecuentemente, conduce a la detección más temprana, pronóstico más predecible y estrategias terapéuticas personalizadas.

Ejemplo 3: MUC5B de variante genética asociado con la forma atenuada de enfermedad pulmonar

Los datos descritos en la presente memoria demuestran que el rs35705950 de MUC5B de variante genética está asociado con el desarrollo de enfermedad pulmonar. Los inventores a continuación examinaron si la variante genética de rs35705950 está asociada con la gravedad y el pronóstico de la enfermedad. Los inventores encontraron que los sujetos de tipo silvestre homocigotos (GG), i.e., aquellos que tienen secuencia del gen MUC5B normal, mostraron una disminución más pronunciada en la capacidad vital forzada (CVF) a lo largo del tiempo en comparación con los sujetos con al menos un alelo T ($P=0,0006$). Esencialmente, aunque la CVF disminuye para ambos grupos, la disminución es más gradual en los que portan el polimorfismo $G \rightarrow T$. Para los sujetos GG, el valor absoluto de la CVF descendió desde aproximadamente 3,4 litros hasta aproximadamente 3,1 litros durante los años 1-3 del estudio. Para los sujetos GT y TT, el valor absoluto de la CVF todavía descendió, pero comenzó a más de 3,5 litros y descendió hasta aproximadamente 3,4 litros durante los años 1-3 del estudio.

Adicionalmente, los inventores observaron una asociación entre la muerte con los sujetos que tienen al menos un alelo T que tienen una mortalidad inferior (OR (IC 95%) = ,37(,20-,67); $p=0.001$) después del ajuste para el género, antecedentes de tabaquismo y DLCO (capacidad pulmonar de difusión de CO₂). Los inventores también observaron una asociación con el tiempo hasta la muerte y el alelo T; tasa de riesgo (IC 95%) = ,50(,30-,83) $p=.0069$ después del ajuste para el género, antecedentes de tabaquismo y DLCO. Estos resultados sugieren que además de ser un fuerte factor de riesgo para el desarrollo de enfermedad pulmonar, el SNP rs35705950 puede indicar un pronóstico menos grave para la enfermedad pulmonar.

LISTADO DE SECUENCIAS

<110> National Jewish Health
Schwartz, David
Seibold, Max

<120> Métodos y Composiciones para la Predicción del Riego, Diagnóstico, Pronóstico, y Tratamiento de Trastornos Pulmonares

<130> 028463-000110PC

<150> EE.UU. 61/298,473
<151> 2010-01-26

<150> EE.UU. 61/298,814
<151> 2010-01-27

<150> EE.UU. 61/323,238
<151> 2010-04-12

<150> EE.UU. 61/323,760
<151> 2010-04-13

<160> 53

<170> Patente En La versión 3.5

<210> 1
<211> 24
<212> ADN
<213> Artificial

<220>
<223> Cebador directo de ADN sintético

<400> 1
ggttcctctt gtcttcgagc ccct 24

<210> 2
<211> 23
<212> ADN

<213> Artificial
 <220>
 <223> Cebador inverso de ADN sintético
 <400> 2
 5 atgggctctt ggtctgctca gag 23
 <210> 3
 <211> 25
 <212> ADN
 <213> Artificial
 10 <220>
 <223> Cebador directo de ADN sintético
 <400> 3
 ggcctggct ctgagtacac atcct 25
 <210> 4
 15 <211> 24
 <212> ADN
 <213> Artificial
 <220>
 <223> Cebador inverso de ADN sintético
 20 <400> 4
 aaggaaagg acacagccgg ttcc 24
 <210> 5
 <211> 23
 <212> ADN
 25 <213> Artificial
 <220>
 <223> Cebador directo de ADN sintético
 <400> 5
 ggtcccat tcatggcagg att 23
 30 <210> 6
 <211> 24
 <212> ADN
 <213> Artificial
 <220>
 35 <223> Cebador inverso de ADN sintético
 <400> 6
 ttctccatg gcagagctgg gacc 24
 <210> 7
 <211> 25
 40 <212> ADN
 <213> Artificial
 <220>
 <223> Cebador directo de ADN sintético
 <400> 7
 45 ctagtgggag ggacgagggc aaagt 25
 <210> 8
 <211> 23
 <212> ADN
 <213> Artificial
 50 <220>

<223> Cebador inverso de ADN sintético

<400> 8
ctcgtggctg tgactgcacc cag 23

5 <210> 9
<211> 20
<212> ADN
<213> Artificial

<220>
<223> Cebador directo de ADN sintético

10 <400> 9
ttggctaagg tgggagacct 20

<210> 10
<211> 20
<212> ADN
15 <213> Artificial

<220>
<223> Cebador inverso de ADN sintético

<400> 10
agcttgggaa tgtgagaacg 20

20 <210> 11
<211> 23
<212> ADN
<213> Artificial

<220>
25 <223> Cebador directo de ADN sintético

<400> 11
catgaggggt gacaggtggc aaa 23

<210> 12
<211> 23
30 <212> ADN
<213> Artificial

<220>
<223> Cebador inverso de ADN sintético

<400> 12
35 cccgcgttg tcttctgaa gtt 23

<210> 13
<211> 23
<212> ADN
<213> Artificial

<220>
40 <223> Cebador directo de ADN sintético

<400> 13
ggtcagaagc ttgaagatg ggc 23

<210> 14
45 <211> 23
<212> ADN
<213> Artificial

<220>
<223> Cebador inverso de ADN sintético

50 <400> 14

cttgtccaat gccagccctg atc 23

5 <210> 15
<211> 20
<212> ADN
<213> Artificial

<220>
<223> Cebador directo de ADN sintético

<400> 15
ctgccagggt taatgaggag 20

10 <210> 16
<211> 20
<212> ADN
<213> Artificial

<220>
15 <223> Cebador inverso de ADN sintético

<400> 16
ggatcaggaa ggatttgag 20

20 <210> 17
<211> 24
<212> ADN
<213> Artificial

<220>
<223> Cebador directo de ADN sintético

<400> 17
25 aggcaggctg gctgaccact gtt

<210> 18
<211> 23
<212> ADN
<213> Artificial

30 <220>
<223> Cebador inverso de ADN sintético

<400> 18
cgtgaagaca gcatcgagag ggg 23

35 <210> 19
<211> 20
<212> ADN
<213> Artificial

<220>
<223> Cebador directo de ADN sintético

40 <400> 19
ttggctaagg tgggagacct 20

<210> 20
<211> 51
<212> ADN
45 <213> Artificial

<220>
<223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B

<400> 20
gtcacctgcc caggccccg aggccyggaa caccttctg ctgggccac c 51

50 <210> 21

<211> 52
 <212> ADN
 <213> Artificial

 <220>
 5 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B

 <400> 21
 ccaccccagg agttgggggg cccccgtrcc agggagcagg aggctgccga gg 52

 <210> 22
 <211> 51
 10 <212> ADN
 <213> Artificial

 <220>
 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B

 <400> 22
 15 gtggccctga tcaactgtgc ctggayggcc tctgaagggg tctgtggggt c 51

 <210> 23
 <211> 51
 <212> ADN
 <213> Artificial

 20 <220>
 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B

 <400> 23
 aacccccctc ggggtctgtg tggctcyaggc cgccccttg tctccactgc c 51

 <210> 24
 25 <211> 49
 <212> ADN
 <213> Artificial

 <220>
 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B

 <400> 24
 30 ttcttcctt tatctctgt ttcagckcc ttcaactgtg aagaagtga 49

 <210> 25
 <211> 50
 <212> ADN
 35 <213> Artificial

 <220>
 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B

 <400> 25
 tgccccggac ccagcccagt tcccartggg ccctctgccc ggggaggtgc 50

 <210> 26
 <211> 55
 <212> ADN
 <213> Artificial

 <220>
 45 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B

 <400> 26
 ggtgggcatc ggctgtgag ctggagccgy gggcagggag gggggatgtc acgag 55

 <210> 27
 <211> 54
 50 <212> ADN
 <213> Artificial

<220>
 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B

<400> 27
 ggctaagggtg ggagacctgg gcgggtgcrt cgggggggacg tctgcagcag aggc 54

5 <210> 28
 <211> 58
 <212> ADN
 <213> Artificial

<220>
 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B

10 <400> 28
 tgcgtcgggg ggagctctgc agcagaggcc ygggcagcag gcacaccct cctgccag 58

<210> 29
 <211> 50
 15 <212> ADN
 <213> Artificial

<220>
 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B

20 <400> 29
 ggggcctggg tgcagtcaca gccacragcc caggggtggg gactctggcc 50

<210> 30
 <211> 50
 <212> ADN
 <213> Artificial

25 <220>
 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B

<400> 30
 cccctccac cgtgccgtgc tgcagygggt ctaccggcct ggatgtgaaa 50

30 <210> 31
 <211> 50
 <212> ADN
 <213> Artificial

<220>
 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B

35 <400> 31
 ccgtgccgtg ctgcagcggg tctacyggcc tggatgtgaa agagagcttg 50

<210> 32
 <211> 50
 <212> ADN
 40 <213> Artificial

<220>
 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B

<400> 32
 agtcccggaa gtgagcgggg agctaygctg agatctggga gaccccctgc 50

45 <210> 33
 <211> 50
 <212> ADN
 <213> Artificial

<220>
 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B

50

<400> 33
 gggagctacg ctgagatctg ggagaycccc tgccccacc caggtacagg 50
 <210> 34
 <211> 50
 5 <212> ADN
 <213> Artificial
 <220>
 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B
 <400> 34
 10 tacgctgaga tctgggagac cccctkcccc caccaggtg cagggccagg 50
 <210> 35
 <211> 50
 <212> ADN
 <213> Artificial
 15 <220>
 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B
 <400> 35
 gcagaagccc gaggtgtgcc ctgagktaa gaaaccgtca caaagaacaa 50
 <210> 36
 20 <211> 50
 <212> ADN
 <213> Artificial
 <220>
 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B
 25 <400> 36
 tgtctcgcc ctccatctcc agaacttct cacattcca agctgaaacc 50
 <210> 37
 <211> 52
 30 <212> ADN
 <213> Artificial
 <220>
 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B
 <400> 37
 cccaagctga aacctgtcc ccatgmaaca ccagctcacc atcccctctg cc 52
 35 <210> 38
 <211> 52
 <212> ADN
 <213> Artificial
 <220>
 40 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B
 <400> 38
 ggcgcccacc gtccacactc cgtctytgcg ggttcatga ctccaggggc ag 52
 <210> 39
 <211> 51
 45 <212> ADN
 <213> Artificial
 <220>
 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B
 <400> 39
 50 ttcatgact ccaggggcag cacacragtg gccctcctg cctttgtct c 51
 <210> 40

<211> 51
 <212> ADN
 <213> Artificial

 <220>
 5 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B

 <400> 40
 ctccaggggc agcacacgag tggccyctcc tgcctttgtc ctctgtgtcc a 51

 <210> 41
 <211> 51
 10 <212> ADN
 <213> Artificial

 <220>
 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B

 <400> 41
 15 ccccatgga gcagcctggg ccagccyctc ttttcacgg ctgaaccgta t 51

 <210> 42
 <211> 53
 <212> ADN
 <213> Artificial

 20 <220>
 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B

 <400> 42
 acccccacca gcagggcaca gggctccrgg tccccacgtc tctgccaaca ctt 53

 <210> 43
 25 <211> 50
 <212> ADN
 <213> Artificial

 <220>
 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B

 <400> 43
 30 ctgatcccc gccatcctat tgagcrtgag acaggtcaga agctttgaag 50

 <210> 44
 <211> 50
 <212> ADN
 35 <213> Artificial

 <220>
 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B

 <400> 44
 gtctgcgcca cggagcattc aggacrtgg tgaccaggga gccaggaggt 50

 40 <210> 45
 <211> 54
 <212> ADN
 <213> Artificial

 <220>
 45 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B

 <400> 45
 cgtcaaggag gtttaccaca tagccccrg gaagcccacc cgacaccagc cgga 54

 <210> 46
 <211> 56
 50 <212> ADN
 <213> Artificial

<220>
 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B

<400> 46
 ttaccacat agccccrpg aagcccaccc racaccagcc ggaggtgcta ggcttc 56

5 <210> 47
 <211> 50
 <212> ADN
 <213> Artificial

<220>
 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B

10 <400> 47
 cccacccgac accagccgga ggtgcyagc ttctgcggct cccacctggg 50

15 <210> 48
 <211> 51
 <212> ADN
 <213> Artificial

<220>
 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B

20 <400> 48
 ggacccatgg tcagtggctg ggggrctgc ccagaggctg ggattccctt c 51

<210> 49
 <211> 51
 <212> ADN
 <213> Artificial

25 <220>
 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B

<400> 49
 gccatctagg acgggtgcca ggtggrtag gcccttctt ccctccgat t 51

30 <210> 50
 <211> 51
 <212> ADN
 <213> Artificial

<220>
 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B

35 <400> 50
 ggtgggtag gcccttctt ccctscgat tctcagaagc tgctgggggt g 51

<210> 51
 <211> 53
 <212> ADN
 <213> Artificial

40 <220>
 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B

<400> 51
 agccctccc cgagagcaaa cacactggc tggagcggg aagagcatgg tgc 53

45 <210> 52
 <211> 51
 <212> ADN
 <213> Artificial

<220>
 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B

50

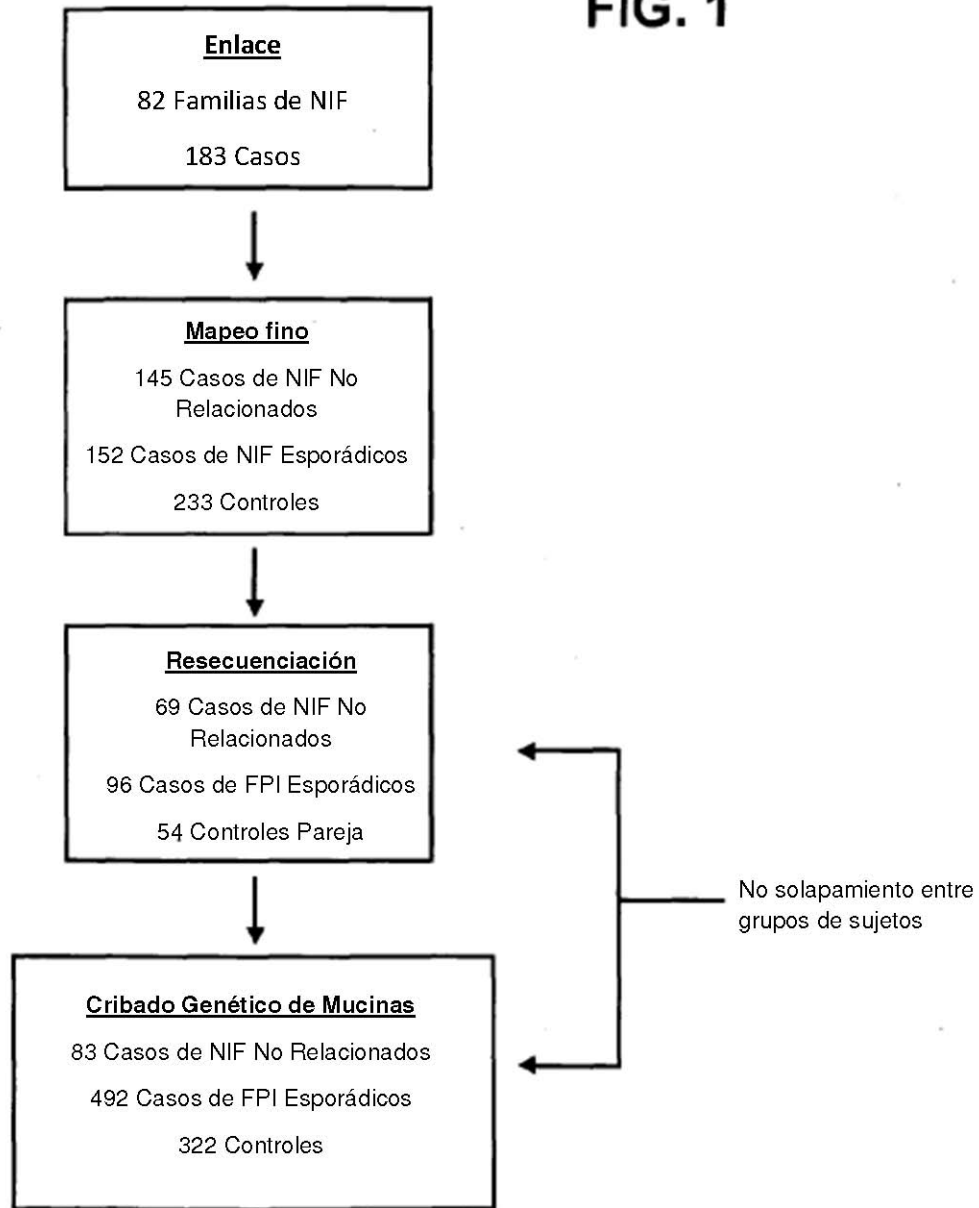
<400> 52
 cacgtggctg gagcggggaa gagcayggtg ccctgcgtgg cctggcctgg c 51
 <210> 53
 <211> 50
 5 <212> ADN
 <213> Artificial
 <220>
 <223> ADN sintético del SNP identificado en la resecuenciación del promotor de MUC5B
 <400> 53
 10 gccgcaggca ggtaagagcc cccaytccg cccctctcg atgctgttt 50

REIVINDICACIONES

- 5 1. Un método para determinar si un sujeto tiene o está en riesgo de desarrollar una afección fibrótica pulmonar, comprendiendo dicho método detectar en una muestra biológica de dicho sujeto si un genoma de dicho sujeto comprende el alelo T de rs35705950 y en donde la presencia del alelo T indica que dicho sujeto tiene o está en riesgo de desarrollar una afección fibrótica pulmonar.
- 10 2. El método de la reivindicación 1, en donde dicha afección fibrótica pulmonar es una enfermedad pulmonar intersticial fibrótica.
3. El método de la reivindicación 2, en donde dicha enfermedad pulmonar intersticial fibrótica es fibrosis pulmonar idiopática o neumonía intersticial familiar.
- 15 4. El método de la reivindicación 1, que comprende además detectar un segundo SNP en el gen MUC5B en una muestra biológica de dicho sujeto, en donde dicho segundo SNP se selecciona del grupo que consiste en rs2672792, rs72636989, MUC5B-Prm1, rs2672794, MUC5B-Prm2, rs11042491, rs2735726, rs868902, MUC5B-Prm3, MUC5B-Prm4, MUC5B-Prm5, rs868903, MUC5B-Prm6, rs885455, rs885454, MUC5B-Prm7, rs7115457, rs7118568 rs56235854 y rs2735738.
- 20 5. El método de la reivindicación 1, en donde dicho sujeto es homocigoto o heterocigoto para el alelo T.
6. El método de la reivindicación 1, en donde la ausencia del alelo T indica que dicho sujeto es menos probable que tenga o esté en riesgo de desarrollar una afección fibrótica pulmonar, y en donde dicho sujeto es negativo para el alelo T.
- 25 7. El método de la reivindicación 1, en donde dicho método comprende además detectar en una muestra biológica de dicho sujeto si dicho sujeto expresa un nivel de ARN MUC5B elevado o un nivel de proteína MUC5B elevado con respecto a un control estándar.
- 30 8. El método de la reivindicación 7, en donde dicho nivel de proteína MUC5B elevado o dicho nivel de ARN MUC5B elevado se detecta a partir de un tejido pulmonar o fluido corporal de dicho sujeto.
- 35 9. El método de la reivindicación 8, en donde dicho fluido corporal es sangre entera, plasma, suero, orina, esputo, saliva, muestra de lavado broncoalveolar, o condensado de aire exhalado.
10. El método de la reivindicación 8, en donde dicho fluido corporal es plasma, sangre entera o suero.
- 40 11. El método de la reivindicación 8, en donde el método comprende además aislar los exosomas o las vesículas subcelulares de dicho fluido corporal.
12. El método de la reivindicación 7, que comprende además determinar un segundo nivel de ARN MUC5B en una muestra de dicho sujeto después de que se haya administrado un tratamiento para una afección fibrótica pulmonar a dicho sujeto y comparar dicho segundo nivel al nivel observado antes de dicho tratamiento.
- 45 13. Un método de pronóstico de una afección fibrótica pulmonar en un paciente, comprendiendo dicho método detectar en una muestra biológica de dicho paciente si un genoma de dicho paciente comprende el alelo T de rs35705950 y en donde la presencia del alelo T indica un pronóstico menos grave de dicha afección fibrótica pulmonar en dicho paciente con respecto a la ausencia del alelo T del SNP en el gen MUC5B.
- 50 14. El método de la reivindicación 13, que comprende además detectar un segundo SNP en el gen MUC5B en una muestra biológica de dicho sujeto, en donde dicho segundo SNP se selecciona del grupo que consiste en rs2672792, rs72636989, MUC5B-Prm1, rs2672794, MUC5B-Prm2, rs11042491, rs2735726, rs868902, MUC5B-Prm3, MUC5B-Prm4, MUC5B-Prm5, rs868903, MUC5B-Prm6, rs885455, rs885454, MUC5B-Prm7, rs7115457, rs7118568 rs56235854 y rs2735738.
- 55 15. El método de la reivindicación 13, en donde dicho paciente es homocigoto o heterocigoto para el alelo T.
16. El método de la reivindicación 1 o 13 que comprende además establecer un curso de tratamiento para una afección fibrótica pulmonar.
- 60 17. El método de la reivindicación 13 en donde la ausencia del alelo T indica un pronóstico más grave de dicha afección fibrótica pulmonar en dicho paciente con respecto a la presencia del alelo T del SNP en el gen MUC5B, y en donde dicho paciente es negativo para el alelo T.

65

FIG. 1



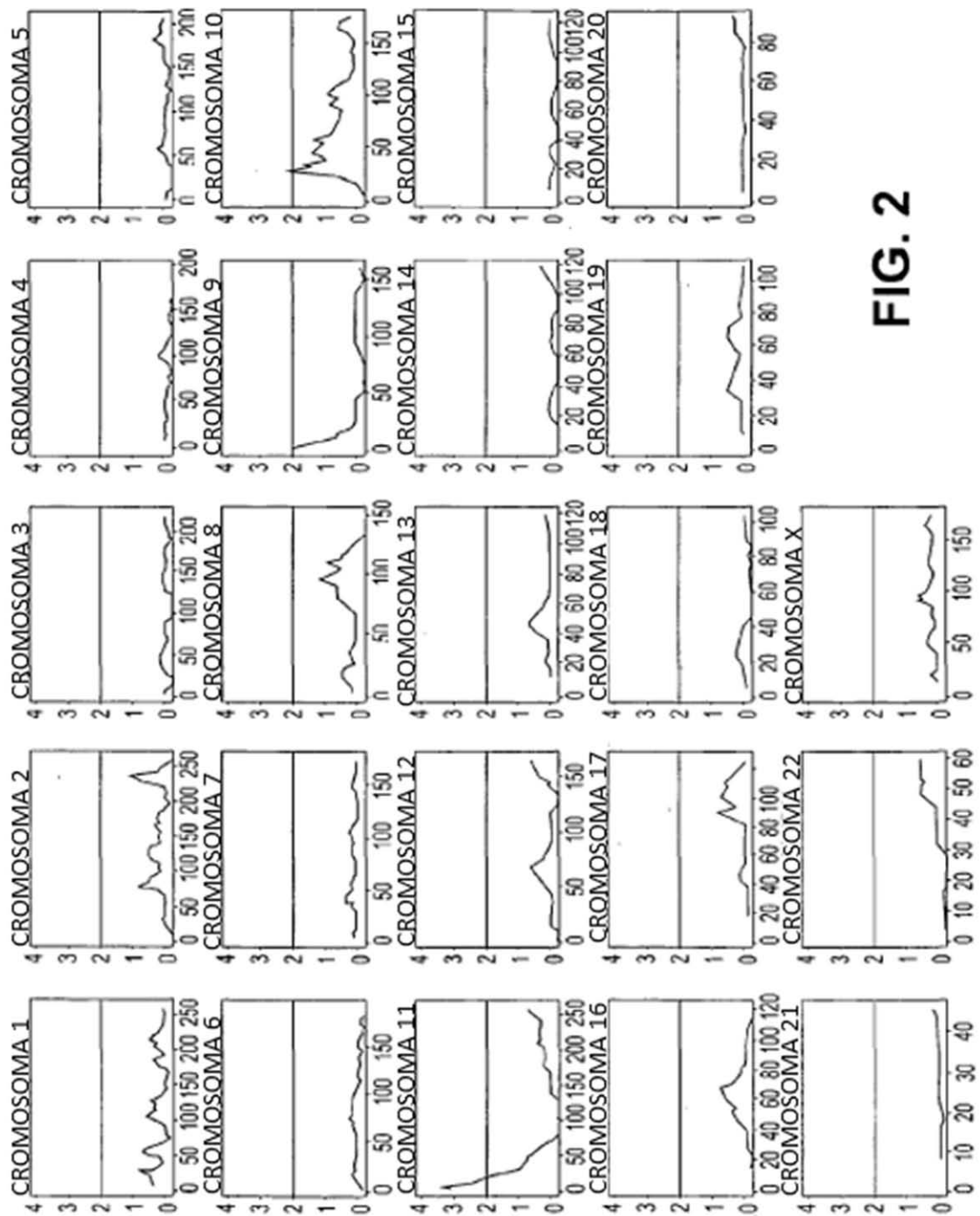
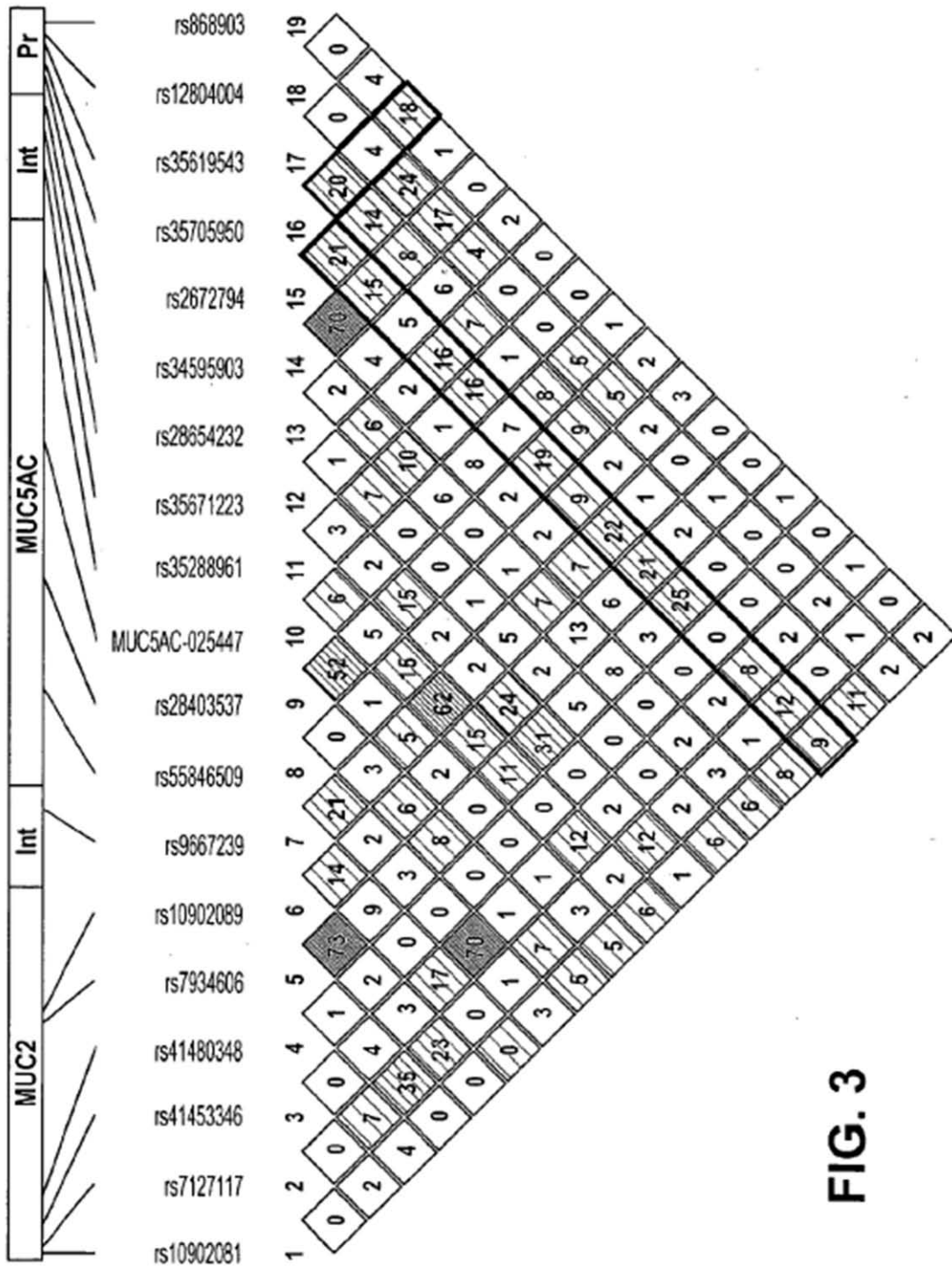


FIG. 2



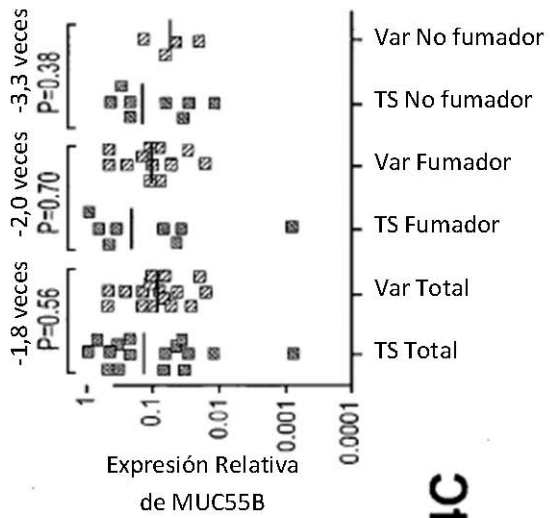
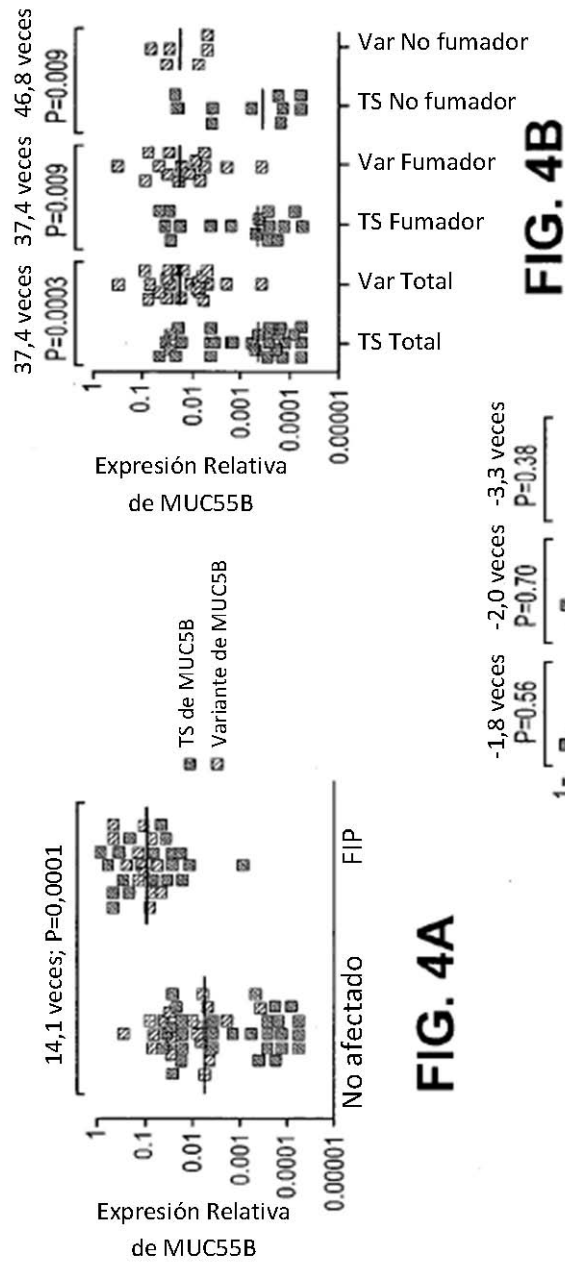


FIG. 4C

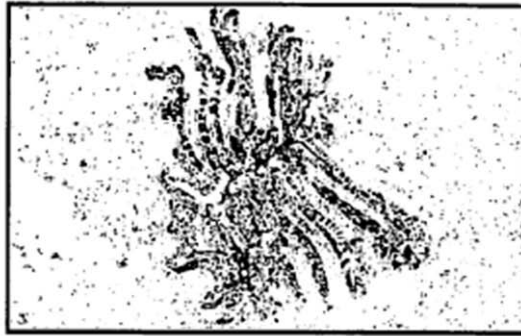


FIG. 5A



FIG. 5B

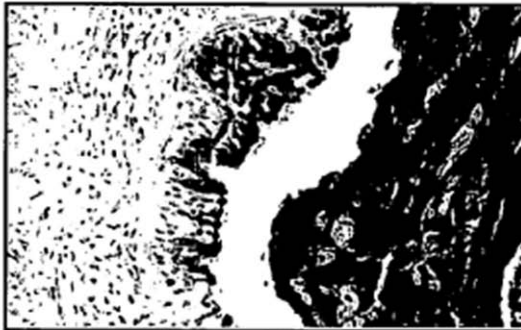


FIG. 5C