

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 993 857**

51 Int. Cl.:

C12Q 1/6883 (2008.01)

G16B 30/00 (2009.01)

G16B 20/00 (2009.01)

G16B 40/20 (2009.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **09.11.2020 PCT/EP2020/081458**

87 Fecha y número de publicación internacional: **14.05.2021 WO21089865**

96 Fecha de presentación y número de la solicitud europea: **09.11.2020 E 20800687 (4)**

97 Fecha y número de publicación de la concesión europea: **03.07.2024 EP 4055190**

54 Título: **Diagnóstico de trastornos del estado de ánimo usando biomarcadores de edición de ARN en sangre**

30 Prioridad:

08.11.2019 EP 19208124

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

10.01.2025

73 Titular/es:

**ALCEDIAG (100.00%)
Route des Michels, La Corneirelle
13790 Peynier, FR**

72 Inventor/es:

**WEISSMANN, DINAH;
SALVETAT, NICOLAS y
CHIMIENTI, FABRICE**

74 Agente/Representante:

DEL VALLE VALIENTE, Sonia

Observaciones:

Véase nota informativa (Remarks, Remarques o Bemerkungen) en el folleto original publicado por la Oficina Europea de Patentes

ES 2 993 857 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

Diagnóstico de trastornos del estado de ánimo usando biomarcadores de edición de ARN en sangre

5 La presente invención se refiere a un método para el diagnóstico *in vitro* de trastornos de depresión, más preferiblemente episodio depresivo mayor (DEP), en un paciente humano a partir de una muestra biológica de dicho paciente, usando al menos dos biomarcadores de ARN de edición de A a I particulares asociados a una puntuación de depresión y algoritmo específicos. La presente invención también se refiere a un método para monitorizar el tratamiento para depresión en un sujeto humano, implementando dicho método el método para el diagnóstico de trastornos del estado de ánimo de la presente invención. La presente invención también comprende un kit para determinar si un paciente presenta un trastorno de depresión.

15 El episodio depresivo mayor (DEP) es un trastorno mental muy común y heterogéneo relacionado con el estrés y el más prevalente, afectando a ~10 % de los hombres y a ~20 % de las mujeres en todo el mundo. El DEP está asociado con un marcado aumento de la mortalidad, principalmente debido a la conducta suicida, pero también a un mayor riesgo de desarrollar otras afecciones médicas, por ejemplo cardiopatía, diabetes y accidente cerebrovascular. El transcurso crónico y recurrente del DEP, asociado a la resistencia al tratamiento que se observa con frecuencia (la tasa de remisión tras el primer ciclo de tratamiento es del 37 %), aumenta aún más el uso del sistema de atención sanitaria. En la actualidad, el DEP se diagnostica mediante una evaluación clínica según criterios establecidos en la 5ª edición del Manual de diagnóstico y estadística de los trastornos mentales¹. El DEP se caracteriza por episodios de al menos 2 semanas durante los cuales una persona experimenta una combinación de cinco o más síntomas diferentes, por ejemplo, estado de ánimo deprimido, pérdida de interés en las actividades diarias, desregulación del sueño, fatiga o indecisión. Por lo tanto, el principal problema es que el diagnóstico de DEP se basa en criterios descriptivos y, por lo tanto, de naturaleza subjetiva. El diagnóstico preciso se complica aún más por la naturaleza heterogénea de DEP, para la que las escalas de depresión más comunes pueden no superponerse por completo, y por su similitud sintomática con otros trastornos psiquiátricos y somáticos.

25 Numerosos estudios han investigado supuestos biomarcadores de DEP², que podrían ayudar a diagnosticar la enfermedad o proporcionar sustitutos fiables para la respuesta clínica al tratamiento. La definición de biomarcador es “una característica que se mide y evalúa objetivamente como indicador de procesos biológicos normales, procesos patógenos o respuestas farmacológicas a una intervención terapéutica”. Como tal, un biomarcador puede usarse como herramienta de diagnóstico, como herramienta para estadificar la enfermedad y/o para predecir y monitorizar la respuesta clínica a una intervención. El desarrollo de biomarcador(es) que pueda(n) ser adecuado(s) para una prueba de diagnóstico en el campo de los trastornos del estado de ánimo es un desafío complejo y difícil. Se han sugerido varios biomarcadores para DEP, incluyendo citocinas proinflamatorias circulantes (por ejemplo, IL6, TNF α , IL-1 β) o el marcador de inflamación CRP, factores neurotróficos (por ejemplo, BDNF) y hormonas, por ejemplo ACTH o cortisol (para una revisión, véase ³). Sin embargo, cuando se usan por separado, ninguno de estos biomarcadores cumple los criterios de sensibilidad y especificidad para una aplicación clínica. Durante la última década, numerosos estudios se han centrado en biomarcadores cerebrales, incluyendo estudios de neuroimagen funcional, o han explorado todo el campo de la ómica, incluyendo la genómica, la transcriptómica, la proteómica o la metabolómica.

40 Los métodos más prometedores parecen radicar en un panel de biomarcadores de múltiples parámetros, para lo cual existe una explicación biológica. Estos biomarcadores deberían cuantificarse usando una muestra de sangre rutinaria y una metodología de laboratorio estándar; a menudo se desarrolla en paralelo un algoritmo de diagnóstico específico⁴. Un problema importante en la identificación de biomarcadores adecuados para un ensayo de diagnóstico en DEP es que podría requerir el uso de una combinación de varios biomarcadores que reflejen los cambios en diferentes mecanismos biológicos.

50 La evidencia acumulada ha indicado que la edición de ARN, que se realiza principalmente por las enzimas ADAR (adenosina desaminasa que actúa sobre el ARN), puede estar críticamente implicada en enfermedades neurológicas o inmunitarias, entre otras patologías. La inflamación se conoce como un factor patogénico, lo que explica posiblemente una parte de la fisiopatología de la depresión. Recientes evidencias apoyan la idea de que la edición del ARN desempeña un papel crucial en procesos de inflamación inmunológica. Se ha demostrado que ADAR1 regula las respuestas inmunitarias cruciales a través de la edición del ARN. Un estudio reciente notificó un aumento general de la edición del ARN en el lupus eritematoso sistémico (LES), un trastorno autoinmunitario crónico, que bien podría aumentar la carga de autoantígenos y promover la progresión de LES. Al editar miARN primarios específicos que se dirigen a la subunidad gp130 del receptor de IL-6, ADAR2 controla la señalización de IL-6 dependiente de gp130. Anteriormente, se demostró la alteración de la edición del ARN en pacientes psiquiátricos con virus de la hepatitis C (VHC) crónica sometidos a una terapia combinada antiviral con IFN- α y ribavirina, en quienes la depresión que surge con el tratamiento es una complicación común. En el SNC, la edición del ARN en la secuencia codificante de proteínas clave para la regulación del estado de ánimo, incluyendo los subtipos AMPA y kainato (KA) del receptor de glutamato y el receptor de serotonina, puede modular las propiedades funcionales de los productos proteicos codificados. La disminución de la edición de los sitios R/G de los receptores AMPA debido a la regulación por disminución de ADAR2 tiene un posible papel en la fisiopatología de los trastornos mentales. La edición del ARN en el sitio Q/R de AMPAR, un receptor transmembrana ionotrópico para el glutamato, puede estar implicada en la patogénesis de los trastornos del estado de ánimo y en la acción de los antidepresivos. Los presentes inventores y otros han identificado una

alteración del receptor de serotonina en el cerebro de pacientes con depresión mayor o que intentan suicidarse deprimidos. En un trabajo reciente, se han identificado en el cerebro de los fallecidos suicidas modificaciones en la edición del ARN de PDE8A como marcador cerebral relacionado con la respuesta inmunológica para el suicidio⁵.

- 5 El documento US2019/185937A1 se refiere a la edición del ARN como biomarcadores para la prueba de trastornos del estado de ánimo.

En el estudio actual, se centra en la modificación de edición de ARN a nivel de todo el transcriptoma mediante Sec-ARN en pacientes con DEP para identificar variantes de edición que pueden usarse como biomarcadores para establecer un ensayo de diagnóstico clínicamente útil para DEP. Mediante la medición de las modificaciones de edición del ARN en todo el transcriptoma, se identificó una tasa de edición diferencial significativa en varios genes relacionados con el sistema inmunitario y la función del SNC. Posteriormente, se seleccionaron los 15 mejores biomarcadores para su validación adicional en una cohorte de priorización. Estas variantes de edición del ARN estaban implicadas en diferentes rutas relevantes para los trastornos del estado de ánimo y DEP, por ejemplo, los receptores del sistema inmunitario o de los neurotransmisores y la transmisión postsináptica. Se aplicó un enfoque de aprendizaje automático para combinar las variantes de edición del ARN y optimizar la predicción, antes de validar el panel de biomarcadores en una gran cohorte de 411 pacientes deprimidos y sujetos de control de la misma edad, raza y sexo. El panel de variantes de edición de ARN discriminó a los pacientes con DEP de los controles sanos con una alta especificidad y sensibilidad. Este estudio es el primero en evaluar múltiples variantes de edición de ARN en muestras de sangre de pacientes con DEP. Nuestros hallazgos contribuirán a una mejor comprensión de la fisiopatología molecular del DEP y allanarán el camino para el desarrollo de un ensayo diagnóstico para el DEP con aplicación clínica.

La presente invención se define en las reivindicaciones adjuntas.

- 25 Se describe, pero no forma parte de la invención reivindicada, un método para seleccionar al menos un biomarcador, preferiblemente una combinación de al menos dos biomarcadores, que puede usarse para diagnosticar trastornos del estado de ánimo, particularmente trastornos de depresión, más preferiblemente episodio depresivo mayor, en un paciente humano a partir de una muestra de sangre de dicho paciente, comprendiendo dicho método la etapa de:

- 30 a) analizar el conjunto de datos de Sec-ARN usando una canalización de análisis del editoma e identificar posiciones editadas diferencialmente de A a I con una cobertura mínima de 30x, lo que garantiza un alto grado de confianza en posiciones de bases particulares;

- 35 b) realizar un análisis diferencial para identificar sitios cuya edición puede ser específicamente diferente entre trastornos del estado de ánimo, particularmente trastornos de depresión, más preferiblemente pacientes con DEP y controles sanos,

- 40 c) aplicar los siguientes criterios de calidad preespecificados del grupo que consiste en: cobertura >30; AUC >0,6; 0,95 >Múltiplo de variación>1,05, $p < 0,05$ y exclusión de sitios intergénicos;

- d) opcionalmente, comprobar que no se observaron diferencias en la edición de ARN global entre pacientes y controles, preferiblemente calculando el índice de edición de Alu (AEI) global, que es una medida de la tasa global de edición de ARN en las repeticiones de Alu;

- 45 e) mediante GSEA (análisis de enriquecimiento sobre conjuntos de genes) sobre los biomarcadores seleccionados en la etapa c) o d), preferiblemente usando herramientas de ontología genética, identificar y seleccionar biomarcadores que reflejen cambios en diferentes procesos biológicos, incluyendo funciones inmunitarias y del SNC (sistema nervioso central); y

- 50 f) aplicar los siguientes criterios de calidad especificados del grupo que consiste en: cobertura >30; AUC >0,8; 0,8 >múltiplo de variación>1,20 y $p < 0,05$ a una combinación de varios biomarcadores seleccionados en la etapa e) que representan diferentes mecanismos biológicos, y se seleccionaron aquellos que discriminaban claramente a los controles sanos y a los pacientes que presentaban trastorno del estado de ánimo, particularmente trastornos de depresión, más preferiblemente episodio depresivo mayor, en dos grupos separados.

- 55 En una realización preferida, en la etapa f), se seleccionan las combinaciones de al menos dos biomarcadores si dichas combinaciones son estadísticamente significativas entre control y casos con un valor de $p < 10^{-4}$.

- 60 La presente invención se refiere a un método *in vitro* para el diagnóstico de trastornos de depresión, más preferiblemente DEP, en un paciente humano a partir de una muestra biológica de dicho paciente, comprendiendo dicho método:

- 65 a) determinar, para al menos dos biomarcadores de ARN de edición de A a I, la proporción relativa de edición de ARN en un sitio de edición dado para al menos uno o una combinación de sitios que se pueden editar en el transcrito de ARN de dicho biomarcador, y/o

- el porcentaje relativo de una isoforma (con sitio(s) editado(s)) o la isoforma no editada) o de una combinación de isoformas del transcrito de ARN de dicho biomarcador;

en donde dichos al menos dos biomarcadores de ARN de edición de A a I se seleccionan del grupo que consiste en los biomarcadores PRKCB, MDM2, PIAS1, IFNAR1, IFNAR2, LYN, AHR, RASSF1, GAB2, CAMK1D, KCNJ15, IL17RA y PTPRC;

b) determinar un valor resultante del/de los porcentaje(s) obtenido(s) en la etapa a);

c) determinar si dicho valor de resultado o puntuación de depresión obtenido en la etapa b) es mayor o no es mayor que un valor de control obtenido para un sujeto sano de control, en donde el valor de control se determinó de una manera comparable a la del valor de resultado, obtenido en la etapa b);

en donde:

si dicha puntuación de depresión obtenida para el paciente es superior a un umbral (punto de corte), clasificar a dicho paciente como positivo o con trastornos de depresión, más preferiblemente DEP, o si dicha puntuación de depresión no es superior a dicho umbral, clasificar a dicho paciente como negativo o sin trastornos de depresión, más preferiblemente DEP.

Los síntomas que sirven como criterio para los trastornos del estado de ánimo, los trastornos de depresión y el DEP son bien conocidos por el experto y se describen, por ejemplo, en el libro DSM-5™ de la Asociación Estadounidense de Psiquiatría (2013, páginas 155-188).

En la presente invención, los trastornos de depresión incluyen el trastorno de desregulación disruptiva del estado de ánimo, el trastorno depresivo mayor (incluyendo el episodio depresivo mayor), el trastorno depresivo persistente (distimia), el trastorno disfórico premenstrual, el trastorno depresivo inducido por sustancias o medicamentos, el trastorno depresivo debido a otra afección médica, otro trastorno depresivo especificado y depresión por trastorno depresivo no especificado.

En una realización preferida, la presente invención se refiere a un método *in vitro* para el diagnóstico de trastornos de depresión, más preferiblemente DEP, en un paciente humano a partir de una muestra biológica de dicho paciente, según la presente invención, en donde dicho método comprende

a) determinar, para una combinación de al menos dos biomarcadores de ARN de edición de A a I:

- para cada biomarcador seleccionado, la proporción relativa de edición de ARN en un sitio de edición dado para al menos uno o una combinación de sitios que se pueden editar en el transcrito de ARN de dichos al menos dos biomarcadores, y/o

- el porcentaje relativo de una isoforma o de una combinación de isoformas del transcrito de ARN de cada uno de dichos dos biomarcadores;

en donde dichos al menos dos biomarcadores de ARN de edición de A a I se seleccionan del grupo que consiste en los biomarcadores PRKCB, MDM2, PIAS1, IFNAR1, IFNAR2, LYN, AHR, RASSF1, GAB2, CAMK1D, KCNJ15, IL17RA, PTPRC y PDE8A;

b) determinar un valor de resultado obtenido para cada uno de dichos al menos dos biomarcadores y un valor de resultado final basándose en un algoritmo o ecuación que incluye el valor de resultado obtenido para cada biomarcador seleccionado;

c) determinar el valor de resultado final o la puntuación de depresión obtenida para el paciente y un valor de resultado de control obtenido para un sujeto de control sano, en donde el valor de resultado de control se determinó de una manera comparable a la del valor de resultado final; y

d) si dicho resultado final o valor de puntuación de depresión es superior a un umbral/punto de corte, clasificar a dicho paciente como positivo o con trastornos de depresión, más preferiblemente DEP, o, si dicho valor de resultado final no es superior a dicho umbral, clasificar a dicho paciente como negativo o sin trastornos de depresión, más preferiblemente MDD.

Por ejemplo, pero sin limitación, en el método de diagnóstico según la presente invención, dicho umbral, punto de corte o puntuación de depresión puede ser:

- el valor Z calculado para un paciente de control sano cuando se usa el método de mROC, o

- la probabilidad de que un paciente se considere como que padece trastornos del estado de ánimo, trastornos de depresión o DEP.

En la presente descripción, el término biomarcador o diana tienen el mismo significado y pueden usarse indistintamente.

En una realización preferida, en el método para diagnosticar la depresión según la presente invención, se seleccionan las combinaciones de al menos dos biomarcadores si dichas combinaciones son estadísticamente significativas entre sujetos de control sanos y casos (pacientes que tienen trastornos del estado de ánimo, trastornos de depresión o DEP) con un valor de $p < 10^{-4}$.

En una realización más preferida, en la etapa a), dicho biomarcador de ARN de edición A a I se selecciona del grupo que consiste en PRKCB, MDM2, IFNAR1, KCNJ15, LYN, CAMK1D y GAB2.

En una realización más preferida, cuando, al menos en la etapa a), se usa una combinación de al menos dos biomarcadores de ARN de edición A a I,

dichos al menos dos biomarcadores de ARN de edición A a I se seleccionan del grupo que consiste en los biomarcadores PRKCB, MDM2, IFNAR1, KCNJ15, LYN, CAMK1D, GAB2 y PDE8A.

En una realización también más preferida, la presente invención se refiere a un método *in vitro* para diagnosticar trastornos depresivos, más preferiblemente episodio depresivo mayor

según la invención, en donde, en la etapa b), el valor de resultado o la puntuación de depresión se calcula mediante un algoritmo que implementa un método multivariante que incluye:

- algoritmo de mROC, en particular para identificar la combinación lineal, que maximiza la curva ROC de AUC (área bajo la curva) y en donde se proporciona la ecuación para la combinación respectiva y se puede usar como un nuevo marcador virtual Z, de la siguiente manera:

$$Z = a.(\text{biomarcador 1}) + b.(\text{biomarcador 2}) + \dots i.(\text{biomarcador i}) + \dots n.(\text{biomarcador n})$$

donde a, b, i, n son coeficientes calculados y (biomarcador 1, 2, i, n) son el nivel del biomarcador considerado (es decir, el nivel del sitio de edición de ARN o de isoformas para una diana/biomarcador dado); y/o

- un algoritmo de bosques aleatorios (RF) aplicado para evaluar el/los sitio(s) de edición de ARN y/o combinaciones de isoformas, en particular para clasificar la importancia del/de los sitio(s) de edición de ARN y/o isoforma(s), y para combinar el/los mejor(es) sitio(s) de edición de ARN y/o isoforma(s), y/u opcionalmente

- un análisis multivariante aplicado para evaluar el/los sitio(s) de edición de ARN y/o combinaciones de isoformas para el diagnóstico, seleccionándose dicho análisis multivariante, por ejemplo, del grupo que consiste en:

- modelo de regresión logística aplicado para análisis univariante y multivariante para estimar el riesgo relativo de un paciente a diferentes valores de nivel de sitio de edición de ARN o de isoformas;

- enfoque CART (árboles de clasificación y regresión) aplicado para evaluar combinaciones de sitio(s) de edición de ARN y/o isoformas;

- enfoque de máquina de vectores de soporte (SVM);

- enfoque de red neuronal artificial (ANN);

- enfoque de red bayesiana;

- enfoque de WKNN (k vecinos más cercanos ponderados);

- mínimos cuadrados parciales - análisis discriminante (PLS-DA);

- análisis discriminante lineal y cuadrático (LDA/ QDA), y

- cualquier otro método matemático que combina biomarcadores.

En una realización preferida particular, dicho trastorno de depresión que se desea diagnosticar es el episodio depresivo mayor (DEP).

En otra realización preferida, la presente invención se refiere a un método para diagnosticar trastornos de depresión según la presente invención, en donde dichos trastornos de depresión incluyen trastornos leves, moderados e graves.

En otra realización preferida, la presente invención se refiere a un método para diagnosticar trastornos de depresión según la presente invención, en donde dicha depresión puede clasificarse como depresión leve, moderada o grave dependiendo del nivel de la diferencia entre el valor de puntuación de depresión obtenido en la etapa b) y el valor de resultado de control obtenido para el sujeto de control.

En una realización preferida, en el método para el diagnóstico de trastornos de depresión según la presente invención, dicha muestra biológica es sangre completa, suero, plasma, orina o líquido cefalorraquídeo, siendo la sangre completa la muestra biológica más preferida.

Se prefieren biomarcadores seleccionados en donde el valor de resultado o la puntuación de depresión obtenida en la etapa c) es estadística/específicamente diferente de dicho valor de resultado de control en $p < 0,05$.

También se prefieren los métodos de la presente invención en donde han de satisfacerse los siguientes criterios para el biomarcador o para la combinación de biomarcadores seleccionados en la etapa a):

- la cobertura >30 ;
- $AUC >0,6$, más preferido $AUC >0,8$;
- $0,95 > \text{múltiplo de variación} > 1,05$, y
- $p < 0,05$.

También se prefieren los métodos de la presente invención en donde dicho(s) biomarcador(es) de ARN de edición de A a I seleccionado(s) comprendido(s) en la etapa a) se selecciona(n) del grupo que consiste en:

a) una combinación que comprende al menos 3, 4, 5, 6, 7 u 8 de los siguientes biomarcadores: PRKCB, MDM2, IFNAR1, KCNJ15, LYN, CAMK1D, GAB2 y PDE8A, preferiblemente la combinación de los 8 biomarcadores citados.

También se prefieren los métodos de la presente invención, en donde dicho(s) biomarcador(es) de ARN de edición de A a I seleccionado(s) comprendido(s) en la etapa a) la combinación de los siguientes 8 biomarcadores GAB2, IFNAR1, KCNJ15, LYN, MDM2, PRKCB, CAMK1D y PDE8A cuando el modelo se ajusta por edad, sexo, tratamiento psiquiátrico y adicción, o cuando el modelo se ajusta por edad y sexo, preferiblemente cuando se usa un algoritmo de bosques aleatorios (RF) aplicado para evaluar el/los sitio(s) de edición de ARN y/o combinaciones de isoformas, en particular para clasificar la importancia del/de los sitio(s) de edición de ARN y/o isoforma(s), y para combinar el/los mejor(es) sitio(s) de edición de ARN y/o isoforma(s).

También se prefieren los métodos de la presente invención, en donde dicho(s) biomarcador(es) de ARN de edición de A a I seleccionado(s) comprendido(s) en la etapa a) la combinación de los siguientes 8 biomarcadores GAB2, IFNAR1, KCNJ15, LYN, MDM2, PRKCB, CAMK1D y PDE8A cuando el modelo se ajusta por edad, sexo, tratamiento psiquiátrico y adicción, o cuando el modelo se ajusta por edad y sexo, usando un algoritmo de bosques aleatorios (RF) aplicado para evaluar el/los sitio(s) de edición de ARN y/o combinaciones de isoformas, en particular para clasificar la importancia del/de los sitio(s) de edición de ARN y/o isoforma(s), y para combinar el/los mejor(es) sitio(s) de edición de ARN y/o isoforma(s), y preferiblemente en donde el cálculo del porcentaje relativo de al menos uno del sitio de edición de ARN o isoforma se basa en el sitio de edición de ARN o isoforma enumerados en las tablas 5.1 y 5.2.

También se prefieren los métodos de la presente invención en donde dicho biomarcador de ARN de edición de A a I seleccionado comprende en la etapa a) uno o una combinación de al menos 2, 3, 4, 5, 6, 7 u 8 biomarcadores seleccionados y en donde el cálculo del porcentaje relativo de al menos uno del sitio de edición de ARN o isoforma se basa en el sitio de edición de ARN o isoforma enumerado en la tabla 2A (sitios de edición), 2B (isoformas) y 3 (sitios de edición) para cada uno de los biomarcadores seleccionados.

También se prefieren los métodos de la presente invención en donde dicha combinación de al menos 2, 3, 4, 5, 6, 7 u 8 biomarcadores seleccionados es una combinación seleccionada de las combinaciones de biomarcadores dadas en la tabla 4 y en la tabla 5, preferiblemente las combinaciones que comprenden los siguientes 8 biomarcadores PRKCB, MDM2, IFNAR1, KCNJ15, LYN, CAMK1D, GAB2 y PDE8A.

También se prefieren las combinaciones que comprenden los 8 biomarcadores PRKCB, MDM2, IFNAR1, KCNJ15, LYN, CAMK1D, GAB2 y PDE8A y se seleccionan entre las combinaciones enumeradas en la tabla 6.

También se prefieren los métodos de la presente invención en donde el algoritmo o la ecuación que permite el cálculo del valor de resultado o puntuación de depresión Z se seleccionan de las ecuaciones de Z enumeradas en los

ejemplos, preferiblemente la ecuación implementa una combinación de los siguientes 8 biomarcadores PRKCB, MDM2, IFNAR1, KCNJ15, LYN, CAMK1D, GAB2 y PDE8A.

En otro aspecto particular, la presente invención se refiere a un método *in vitro* para diagnosticar trastornos de depresión, más preferiblemente MDD, en un paciente humano según la invención, en donde si el paciente a analizar es una mujer, en la etapa a) dicho al menos uno o una combinación de al menos dos biomarcadores de ARN de edición de A a I se selecciona del grupo de biomarcadores que consiste en PRKCB, MDM2, PIAS1, IFNAR1, IFNAR1, IFNAR1, 2, LYN, AHR, RASSF1, GAB2, CAMK1D, KCNJ15, IL17RA y PTPRC, si se selecciona un biomarcador, o dicho grupo comprende además el ARN de edición de A a I PDE8A si se selecciona una combinación de al menos dos biomarcadores (o dianas).

También se prefieren los métodos de la presente invención en donde dicha combinación de al menos 2, 3, 4, 5, 6, 7 u 8 biomarcadores seleccionados es una combinación seleccionada del grupo que consiste en los biomarcadores (o dianas) PRKCB, MDM2, IFNAR1, KCNJ15, LYN, CAMK1D, GAB2 y PDE8A, particularmente las combinaciones de biomarcadores o la ecuación Z que figuran en los ejemplos o figuras con pacientes de la población femenina, particularmente las combinaciones que comprenden los 8 biomarcadores PRKCB, MDM2, IFNAR1, KCNJ15, LYN, CAMK1D, GAB2 y PDE8A, más preferiblemente las combinaciones enumeradas en la tabla 7.

En otro aspecto particular, la presente invención se refiere a un método *in vitro* para diagnosticar depresión en un paciente humano según la invención, en donde si el paciente a analizar es un hombre, en la etapa a) dicho al menos uno o una combinación de al menos dos biomarcadores de ARN de edición de A a I se selecciona del grupo de biomarcadores que consiste en PRKCB, MDM2, PIAS1, IFNAR1, IFNAR1, IFNAR1, 2, LYN, AHR, RASSF1, GAB2, CAMK1D, KCNJ15, IL17RA y PTPRC, si se selecciona un biomarcador, o dicho grupo comprende además el ARN de edición de A a I PDE8A si se selecciona una combinación de al menos dos biomarcadores (o dianas).

También se prefieren los métodos de la presente invención en donde dicha combinación de al menos 2, 3, 4, 5, 6, 7 u 8 biomarcadores seleccionados es una combinación seleccionada del grupo que consiste en los biomarcadores (o dianas) PRKCB, MDM2, IFNAR1, KCNJ15, LYN, CAMK1D, GAB2 y PDE8A, particularmente las combinaciones de biomarcadores o la ecuación Z que figuran en los ejemplos o figuras con pacientes de la población masculina, particularmente las combinaciones que comprenden los 8 biomarcadores PRKCB, MDM2, IFNAR1, KCNJ15, LYN, CAMK1D, GAB2 y PDE8A, más preferiblemente las combinaciones enumeradas en la tabla 9.

En una realización más preferida, la presente invención se refiere a un método *in vitro* para el diagnóstico de depresión en un paciente humano, según la presente invención, en donde la combinación de biomarcadores y/o la ecuación de Z o el algoritmo asociado a dicha combinación usada para el diagnóstico de trastornos del estado de ánimo, particularmente trastornos de depresión, es diferente si el paciente a analizar es una mujer o un hombre.

Preferiblemente, la combinación de biomarcadores y/o la ecuación de Z de si el paciente a analizar es mujer u hombre se selecciona de la combinación y/o la ecuación de Z representada en los ejemplos o en las figuras para este caso particular.

También se prefieren los métodos de la presente invención, en donde en la etapa a) se miden la proporción relativa de edición de ARN en una edición dada y/o el porcentaje de una isoforma mediante NGS en dicha muestra biológica.

También se prefieren los métodos de la presente invención, en donde, en la etapa a), el amplicón/secuencia nucleica usada para la detección de los sitios de edición de ARN y/o las isoformas del transcrito de ARN de dicho biomarcador se obtiene o se puede obtener con el conjunto de cebadores enumerados en la tabla 1 para cada uno de los biomarcadores seleccionados (Id. de sec. n.º 1 a 36). También se prefieren los métodos de la presente invención, en donde en la etapa a), la secuencia de amplicón/nucleica usada para la detección de los sitios de edición de ARN y/o las isoformas del transcrito de ARN de dicho biomarcador se obtiene o se puede obtener con un conjunto o una combinación de conjuntos de cebadores que permiten obtener un amplicón(es) o secuencia(s) de ácido nucleico que incluyen, o son idénticos a, el(los) amplicón(es) o secuencia(s) de ácido nucleico que se pueden obtener mediante el conjunto de cebadores enumerados en la Tabla 1 (Id. de sec. n.º 1 a Id. de sec. n.º 36). Dicha secuencia de amplicón o ácido nucleico usada para la detección de los sitios de edición de ARN y/o las isoformas del transcrito de ARN de dicho biomarcador podría ser la propia secuencia completa del transcrito de ARN.

En otro aspecto, la presente invención se refiere a un método para monitorizar el tratamiento de los trastornos depresivos, en un sujeto humano a partir de una muestra de sangre de dicho sujeto, preferiblemente para monitorizar el DEP, comprendiendo dicho método:

A) diagnosticar trastornos de depresión en dicho ser humano mediante el método para diagnosticar trastornos de depresión en un paciente humano según la presente invención antes del inicio del tratamiento que se desea monitorizar;

B) repetir las etapas (a) a c) del método para diagnosticar trastornos de depresión en un paciente humano según la presente invención, después de un período de tiempo durante el cual dicho paciente recibe tratamiento para dicha depresión, para obtener una puntuación de depresión postratamiento;

5 C) comparar la puntuación de depresión postratamiento de la etapa (c) con:

- la puntuación de depresión (es decir, el valor de Z, valor de probabilidad) obtenida antes del período de tiempo durante el cual dicho paciente recibe tratamiento para dicha depresión, y

10 - con la puntuación de depresión (valor final de control) para sujetos normales/sanos, y clasificar dicho tratamiento como eficaz si la puntuación de depresión postratamiento obtenida en la etapa B) está más cerca que la puntuación de depresión obtenida antes del período de tiempo durante el cual dicho paciente recibe tratamiento de la puntuación de depresión obtenida para sujetos normales/sanos.

15 En otro aspecto, la presente invención se refiere a un kit para determinar si un paciente presenta una depresión, comprendiendo dicho kit:

1) opcionalmente, instrucciones para aplicar el método según la invención tal como se define en las reivindicaciones, con el fin de obtener el valor de resultado o la puntuación de depresión cuyo análisis determina si
20 dicho paciente presenta un trastorno de estado del ánimo, preferiblemente trastorno de depresión; y

- una combinación de al menos dos conjuntos de cebadores del grupo de pares de cebadores que consiste en: Id. de sec. n.º 1 a Id. de sec. n.º 36, cuya selección depende de la combinación de biomarcadores usada para el diagnóstico del trastorno del estado de ánimo, preferiblemente del trastorno de depresión;

25 o

- una combinación de al menos dos conjuntos de cebadores que permite obtener un(os) amplicón/amplicones idéntico(s) al/a los amplicón/amplicones obtenido(s) mediante los al menos dos conjuntos de cebadores del grupo de pares de cebadores que consiste en la Id. de sec. n.º: 1 a la Id. de sec. n.º: 36.

Los siguientes ejemplos y las figuras y leyendas en lo sucesivo en la presente memoria se han escogido para proporcionar a los expertos en la materia una descripción completa para poder implementar y usar la presente invención. Estos ejemplos no pretenden limitar el alcance de lo que el inventor considera que es su invención, ni pretenden mostrar que solo se llevaron a cabo los experimentos en lo sucesivo en la presente memoria.

Otras características y ventajas de la invención se irán poniendo de manifiesto en el resto de la descripción con los Ejemplos, las Figuras y las Tablas, las leyendas para las cuales se dan a continuación en la presente memoria.

40 Breve descripción de los dibujos y las tablas

Figuras 1A-1D: Identificación y validación de sitios de edición de ARN de A a I diferencialmente entre controles sanos y pacientes con DEP usando datos de Sec-ARN de sangre humana.

45 A: Localización genómica de los sitios de edición detectados

B: Repartición de eventos de edición de A a I o genes editados por cromosoma.

50 Gris oscuro (eje y izquierdo): Repartición de eventos de edición de A a I identificados con una cobertura mínima de 30 en el estudio de Sec-ARN.

Gris claro (eje y derecho): Repartición de genes editados identificados con una cobertura mínima de 30 en el estudio de Sec-ARN

55 C: Gráfico de volcán de sitios editados diferencialmente entre controles y pacientes con DEP. El gráfico de volcán muestra los sitios de genes editados diferencialmente regulados al alza y a la baja entre el grupo de pacientes deprimidos (DEP) y el grupo de control. Para cada gráfico, el eje x representa el log (2) (múltiplo de variación) (FC), y el eje y representa -log 10 (valores de p). Los sitios de edición con un valor de p ajustado de menos de 0,05 Y FC >5 % se asignaron como editados diferencialmente y se indican en verde.

60 D: Índice de edición de Alu entre controles y pacientes deprimidos. Distribución de valores de índice de edición de Alu (AEI) en controles y pacientes con DEP. Los datos son la media ± EEM. El valor de p del índice de edición se calculó usando la prueba de suma de rangos de Wilcoxon.

65 **Figura 2:** Análisis de componentes principales de las 646 variantes en 366 genes editados diferencialmente entre pacientes deprimidos y controles sanos identificados en el estudio de Sec-ARN.

El gráfico de dispersión visualiza la primera, la segunda y la tercera componentes principales y los porcentajes de varianza respectivos en los ejes x, y y z de la relación (múltiplo de variación) de los biomarcadores. círculos grises: Pacientes con DEP; círculo negro: Controles sanos.

Figura 3: Asociaciones proteína-proteína en los 14 genes editados diferencialmente usados en las cohortes de priorización.

El diagrama muestra las interacciones proteína-proteína en el grupo de 14 genes seleccionados en la cohorte de priorización. Los genes fueron anotados y coloreados según su implicación en procesos biológicos.

-  : Proceso de regulación del sistema inmunitario (GO:0002682);
-  : Ruta de señalización del receptor a través de JAK-STAT (GO:0007259);
-  : Regulación de la respuesta inmunitaria (GO:0050776)

Figuras 4A-4D: Ejemplos de rendimiento diagnóstico de la edición de ARNm de la primera cohorte de priorización (n = 76) analizado mediante monoplex.

- A: Ejemplo de rendimiento diagnóstico obtenido mediante el uso de 4 sitios de edición de ARN en 3 dianas diferentes.
- B: Ejemplo de rendimiento diagnóstico obtenido mediante el uso de 4 sitios de edición de ARN en 2 dianas diferentes.
- C: Ejemplo de rendimiento diagnóstico obtenido mediante el uso de 6 sitios de edición de ARN en 4 dianas diferentes.
- D: Ejemplo de rendimiento diagnóstico obtenido mediante el uso de 29 sitios de edición de ARN presentes en 5 dianas diferentes.

Figuras 5A-4F: Rendimiento diagnóstico de la edición de ARNm de la segunda cohorte de priorización (n = 86) analizado mediante secuenciación multiplexada.

- A: Ejemplo de rendimiento diagnóstico obtenido mediante el uso de 2 sitios de edición de ARN en 3 dianas diferentes.
- B: Ejemplo de rendimiento diagnóstico obtenido mediante el uso de 3 sitios de edición de ARN en 3 dianas diferentes.
- C: Ejemplo de rendimiento diagnóstico obtenido mediante el uso de 4 sitios de edición de ARN en 4 dianas diferentes.
- D: Ejemplo de rendimiento diagnóstico obtenido mediante el uso de 5 sitios de edición de ARN presentes en 4 dianas diferentes.
- E: Ejemplo de rendimiento diagnóstico obtenido mediante el uso de 19 sitios de edición de ARN en 8 dianas diferentes.
- F: Ejemplo de rendimiento diagnóstico obtenido mediante el uso de 48 sitios de edición de ARN presentes en 9 dianas diferentes.

Figuras 6A-6C: Rendimiento diagnóstico de la edición de ARNm de la cohorte de validación (n = 411) analizado mediante secuenciación multiplexada de 8 dianas.

- A: Ejemplo de rendimiento diagnóstico obtenido mediante el uso de 17 sitios de edición de ARN en 5 dianas diferentes.
- B: Ejemplo de rendimiento diagnóstico obtenido mediante el uso de 53 sitios de edición de ARN en 8 dianas diferentes.
- C: Ejemplo de rendimiento diagnóstico obtenido mediante el uso de 79 sitios de edición de ARN en 8 dianas diferentes.

Figuras 7A-7E: Rendimiento diagnóstico de la edición de ARNm de la población femenina de la cohorte de validación (n = 277) analizado mediante secuenciación multiplexada de 8 dianas.

- A: Ejemplo de rendimiento diagnóstico obtenido mediante el uso de 11 sitios de edición de ARN en 4 dianas diferentes.
- B: Ejemplo de rendimiento diagnóstico obtenido mediante el uso de 30 sitios de edición de ARN en 6 dianas diferentes.
- C: Ejemplo de rendimiento diagnóstico obtenido mediante el uso de 49 sitios de edición de ARN en 8 dianas diferentes.
- D: Curva ROC obtenida tras aprendizaje automático con el método de mROC.

E: Curva ROC obtenida tras aprendizaje automático con el método de Bosque aleatorio

Figuras 8A-8E: Rendimiento diagnóstico de la edición de ARNm de la población masculina de la cohorte de validación (n = 134) analizado mediante secuenciación multiplexada de 8 dianas.

A: Ejemplo de rendimiento diagnóstico obtenido mediante el uso de 8 sitios de edición de ARN en 6 dianas diferentes.

B: Ejemplo de rendimiento diagnóstico obtenido mediante el uso de 39 sitios de edición de ARN en 7 dianas diferentes.

C: Ejemplo de rendimiento diagnóstico obtenido mediante el uso de 68 sitios de edición de ARN en 8 dianas diferentes.

D: Curva ROC obtenida tras aprendizaje automático con el método de mROC.

E: Curva ROC obtenida tras aprendizaje automático con el método de Bosque aleatorio

Figura 9: Correlación entre las puntuaciones clínicas MADRS e IDS-C30 para los sujetos incluidos en el estudio de Sec-ARN.

El gráfico muestra una confianza de intervalos del 95 %. Se indican el coeficiente de correlación de Pearson y el valor de *p*.

Figura 10: Visión general de la canalización de análisis del editoma

Figura 11: Características demográficas y diagnóstico psiquiátrico de la población de estudio incluida en el estudio de Sec-ARN.

Los datos son la media $\pm$ EEM. Los valores *p* de las características principales se visualizan con la prueba *t* de Student. CTRL: Controles sanos; DEP: episodio depresivo mayor; IMC: Índice de masa corporal

Figura 12: Categorización funcional de los 366 genes editados diferencialmente entre DEP y controles basándose en anotaciones de ontología genética (GO).

Se muestran los resultados de las 20 pruebas de sobrerrepresentación de GO principales para el enriquecimiento de términos GO de procesos biológicos con genes editados diferencialmente. La entrada a las herramientas de análisis de enriquecimiento de GO fue la lista de 366 genes editados diferencialmente entre DEP y los controles. Se presentan la relación entre el número de genes/número total de genes incluidos en el término GO y los valores *p* corregidos usando el método de Benjamini-Hochberg para múltiples pruebas. TDF, tasa de descubrimientos falsos; GO, ontología genética.

Figura 13: Análisis de enriquecimiento de la ruta de señalización de los 366 genes editados diferencialmente entre DEP y los controles.

La tabla 3 muestra las rutas significativas obtenidas realizando un análisis de sobrerrepresentación en Reactome. Se muestran la relación entre el número de genes/número total de genes incluidos en la ruta y los valores *p* corregidos usando el método de Benjamini-Hochberg para múltiples pruebas.

Figura 14: Variantes de edición de genes y ARN analizadas en la cohorte de priorización. Se muestran el gen y la posición (GRCh38) de la variante de edición, la cobertura después de Sec-ARN, el cambio entre controles y pacientes con DEP, el valor de *p* y el área bajo la curva calculada a partir de datos de Sec-ARN. Los valores *p* se calcularon usando la prueba de suma de rangos de Wilcoxon.

Figura 15: Categorización funcional de los 14 genes editados diferencialmente usados en las cohortes de priorización basándose en anotaciones de ontología genética (GO).

Se muestran los resultados de las 20 pruebas de sobrerrepresentación de GO principales para el enriquecimiento de términos GO de procesos biológicos con genes editados diferencialmente. La entrada a las herramientas de análisis de enriquecimiento de GO fue la lista de 14 genes editados diferencialmente entre DEP y controles analizados en las cohortes de priorización. Se presentan valores *p* corregidos usando el método de Benjamini-Hochberg para múltiples pruebas. TDF, tasa de descubrimientos falsos; GO, ontología genética.

Figura 16: Análisis de enriquecimiento de la ruta de señalización de los 14 genes analizados en las cohortes de priorización.

La tabla muestra las rutas significativas obtenidas al realizar el análisis de sobrerrepresentación en Reactome. Se muestran los valores *p* corregidos usando el método de Benjamini-Hochberg para pruebas múltiples.

Figura 17: Las características demográficas y el diagnóstico psiquiátrico de la población de estudio incluyeron el primer estudio de priorización analizado en monoplex.

- 5 Los datos son la media $\pm$ EEM. Los valores p de las características principales se visualizan con la prueba t de Student. CTRL: Controles sanos; DEP: episodio depresivo mayor; IMC: Índice de masa corporal

Se muestran las pruebas.

- 10 **Figura 18:** Las características demográficas y el diagnóstico psiquiátrico de la población de estudio incluyeron el segundo estudio de priorización analizado mediante secuenciación multiplexada.

Los datos son la media $\pm$ EEM. Los valores p de las características principales se visualizan con la prueba t de Student. CTRL: Controles sanos; DEP: episodio depresivo mayor; IMC: Índice de masa corporal.

- 15 **Figura 19:** Características demográficas y diagnóstico psiquiátrico de la población de estudio incluida en la cohorte de validación.

- 20 Los datos son la media $\pm$ EEM. Los valores p de las características principales se visualizan con la prueba t de Student. CTRL: Controles sanos; DEP: episodio depresivo mayor; IMC: Índice de masa corporal

Figura 20: Categorización funcional de los 366 genes editados diferencialmente entre DEP y controles basándose en anotaciones de g:Profiler.

- 25 Se muestran los resultados de las 20 pruebas de sobrerepresentación de GO principales para el enriquecimiento de términos GO de procesos biológicos con genes editados diferencialmente. La entrada al g:Profiler fue la lista de 366 genes editados diferencialmente entre DEP y los controles. Se presentan la relación entre el número de genes/número total de genes incluidos en el término GO y los valores p corregidos usando el método de Benjamini-Hochberg para múltiples pruebas. TDF, tasa de descubrimientos falsos; GO, ontología genética.

- 30 **Figura 21:** Categorización funcional de los 366 genes editados diferencialmente entre DEP y controles basándose en anotaciones de GOnet.

- 35 Se muestran los resultados de las 20 pruebas de sobrerepresentación de GO principales para el enriquecimiento de términos GO de procesos biológicos con genes editados diferencialmente. La entrada al GOnet fue la lista de 366 genes editados diferencialmente entre DEP y los controles. Se presentan la relación entre el número de genes/número total de genes incluidos en el término GO y los valores p corregidos usando el método de Benjamini-Hochberg para múltiples pruebas. TDF, tasa de descubrimientos falsos; GO, ontología genética.

- 40 **Figuras 22A y 22B:** Rendimiento diagnóstico de la edición de ARNm de la cohorte de validación (n = 411) analizado mediante secuenciación multiplexada de 8 dianas usando bosques aleatorios.

- 45 A. Curva característica operativa del receptor y rendimientos diagnósticos para la depresión con el primer modelo ajustado por edad, sexo, tratamiento psiquiátrico y adicción. Se representan gráficamente las probabilidades de la respuesta correcta de los conjuntos sometidos a prueba

- B. Curva característica operativa del receptor y rendimientos diagnósticos para la depresión con el segundo modelo ajustado por edad y sexo. Se representan gráficamente las probabilidades de la respuesta correcta de los conjuntos sometidos a prueba

- 50 Ejemplo 1: Materiales y métodos

Sujetos y evaluación clínica

- 55 Se reclutaron pacientes con depresión (n= 268) entre los pacientes ambulatorios del Departamento de Psiquiatría de Emergencia y Cuidados Intermedios (CHRU de Montpellier) según los principios de la Declaración de Helsinki de 1975 y sus actualizaciones sucesivas. Este estudio fue aprobado por el Comité Ético local francés (CPP Sud-Méditerranée IV en Montpellier, CPP n.º A01978-41). Todos los participantes, de entre 18 y 65 años, entendieron y firmaron un consentimiento informado por escrito antes de ingresar al estudio. El estudio fue aprobado por la Junta de Revisión Institucional local, según los requisitos de aprobación y las buenas prácticas clínicas. Todos los pacientes cumplieron los criterios de DEP del Manual diagnóstico y estadístico de los trastornos mentales IV (DSM-IV) usando la Entrevista Neuropsiquiátrica Internacional MINI. La presencia o ausencia de diagnósticos psiquiátricos y estados de ánimo al momento de ingresar al estudio fueron confirmados por psiquiatras capacitados. Durante la entrevista estandarizada, los psiquiatras emplearon la versión francesa de la Escala de Clasificación de la Depresión de Montgomery-Asberg (MADRS) y el Inventario de Sintomatología Depresiva de 30 puntos, calificado por médicos (IDS-C30), para puntuar la depresión. Los pacientes con DEP, todos de raza blanca, se reclutaron entre los pacientes ambulatorios del

Departamento de Psiquiatría de Emergencia y Cuidados Intermedios (CHRU de Montpellier). Los sujetos de control emparejados por edad, raza y sexo (n = 143) fueron reclutados de una lista de voluntarios del Centro de Investigación Clínica (CHRU de Montpellier).

5 Extracción y cualificación de ARN a partir de sangre completa

Se recuperó un volumen de 4 ml de sangre completa de cada paciente en tubos de ARN sanguíneo PAXgene™, que están optimizados para la estabilización del ARN y se almacenan a -20 °C antes de transferirse a -80 °C. Las muestras se distribuyeron aleatoriamente en los diferentes conjuntos de extracciones. Las muestras de sangre se descongelaron a 4 °C durante la noche y se centrifugaron durante 10 min a 3.000 g. El sedimento se lavó con 4 ml de DPBS 1X y se resuspendió en 400 µl de DPBS 1X. Las muestras de ARN totales se aislaron usando el kit MagNA Pure 96 Cellular RNA Large Volume (LifeScience), según el protocolo automatizado del fabricante. Los ARN se eluyeron en 100 µl de tampón de elución MagNA Pure 96. Durante la preparación de la muestra y la extracción de ARN, se tomaron precauciones estándar para evitar la degradación del ARN por las ARNasas. Las concentraciones de ARN total y los niveles de calidad se determinaron con los instrumentos Qubit Fluorometer (Invitrogen) y LabChip (Perkin-Elmer, HT RNA Reagent Kit), respectivamente.

Preparación de bibliotecas para secuenciación de ARN

Para la preparación de bibliotecas de ARN a partir de ARN derivados de la sangre, elegimos una estrategia específica destinada a agotar en un solo paso tanto el ARN ribosómico como los ARN de globina, que son dos formas de ARN abundante, para enriquecer la biblioteca con ARN totales (ARNm). Se prefirieron los ARN totales a los ARN mensajeros, ya que se ha mostrado que la edición de ARN en las células sanguíneas se enriqueció en la región intrónica de elementos repetitivos, que se eliminan durante el empalme para producir ARN mensajeros. Por ello, usamos un kit de biblioteca TruSeq Stranded Total RNA (Illumina) con Ribo-Zero Globin, que está diseñado específicamente para muestras de sangre, según las instrucciones del fabricante. Brevemente, se agotaron 300 ng de ARN totales en ARN ribosómico y ARN de globina usando sondas de agotamiento de ARNr y globina en combinación con perlas magnéticas. Después de la purificación, los ARN se fragmentaron en fragmentos de 250 pb en promedio usando cationes divalentes a temperatura elevada. A continuación, se logró la síntesis de ADNc de primera cadena usando la transcriptasa inversa Superscript II (Thermo Fisher Scientific) y cebadores aleatorios. A continuación, se realiza la síntesis de ADNc de segunda cadena usando la ADN polimerasa I y la ARNasa H. De manera importante, durante la síntesis de ADNc de segunda cadena, se incorpora dUTP en lugar de dTTP para generar una biblioteca específica de cadena. Después de la síntesis de ADNc de segunda cadena, los extremos 3' de los fragmentos romos se adenilan en 3' antes de la ligadura del adaptador. Es importante destacar que cada índice se ha atribuido a todos los grupos para evitar un sesgo putativo debido al efecto inespecífico del índice sobre la profundidad de la secuenciación. A continuación, fragmentos de ADN que tienen moléculas adaptadoras en ambos extremos se amplificaron mediante PCR para aumentar la cantidad de ADN en la biblioteca. A continuación, se realizaron controles de calidad usando instrumentos tanto Qubit Fluorometer (Invitrogen) como LabChip (Perkin-Elmer, HT DNA 5K Reagent Kit) para la cuantificación de las plantillas de la biblioteca de ADN y el análisis de control de calidad, respectivamente. A continuación, las muestras se agruparon en una biblioteca y se purificaron usando el sistema de limpieza de PCR Magbio. La biblioteca se desnaturalizó usando NaOH 0,1 M y se secuenció en un experimento con Illumina NextSeq 500/550 de alto rendimiento usando química de extremos emparejados con una longitud de lectura de 75 pb.

45 Procesamiento de datos Sec-ARN, análisis de expresión diferencial y análisis del editoma

Las lecturas de extremos emparejados se generaron usando un Illumina NextSeq 500 y se desmultiplexaron usando bcl2fastq (versión 2;17.1.14, Illumina). La calidad de la secuenciación se realizó usando el software FastQC (versión 0.11.7, <https://github.com/s-andrews/FastQC>). Las alineaciones de lectura se realizaron usando la versión del genoma humano hg38 con el alineador STAR⁶, seguido de una realineación con Genome Analysis Tool Kit (GATK, versión 4.1.4.0)⁷, eliminación de duplicados de PCR y denominación de variantes de un solo nucleótido (SNV) finales. La identificación y cuantificación de los eventos de edición de A a I se realizó usando RNAEditor (versión 1.0)⁸. A continuación, se realizó un filtrado adicional y una anotación funcional de los eventos editados con ANNOVAR⁹, RepeatMasker¹⁰ y BiomaRt¹¹.

55 Preparación y secuenciación de bibliotecas de secuenciación de próxima generación dirigida

Para la preparación de bibliotecas de secuenciación de nueva generación (NGS), elegimos un enfoque dirigido multiplexado para secuenciar selectivamente la región de interés dentro de cada diana relevante. Se usaron cebadores de PCR validados para amplificar la región de interés mediante PCR (tabla 20). Para la amplificación por PCR, se usó la enzima Q5 Hot Start High Fidelity (New England Biolabs) según las pautas del fabricante. La reacción de PCR se realizó en un termociclador Pqstar 96x usando un protocolo de PCR optimizado. Se evaluaron tanto la cantidad como la calidad del producto de PCR. La pureza del amplicón se determinó con un analizador de ácidos nucleicos (LabChipGx, Perkin Elmer) y la cuantificación se realizó usando el método Qubit basado en fluorescencia. Después del control de calidad, las reacciones de PCR se purificaron usando perlas magnéticas (sistema High Prep PCR MAGbio, Mokascience). A continuación, se cuantificó el ADN usando el sistema Qubit y se calculó el rendimiento de

purificación. A continuación, las muestras se indexaron individualmente mediante amplificación por PCR usando la enzima PCR de alta fidelidad Q5 Hot start y el kit Illumina 96 Indexes (kit Nextera XT index; Illumina). A continuación, las muestras se agruparon en una biblioteca y se purificaron usando el sistema de limpieza de PCR Magbio. La biblioteca se desnaturalizó usando NaOH 0,1 M y se cargó en un cartucho de secuenciación (Illumina MiSeq Reagent Kit V3 o Illumina NextSeq 500/550 Mid-Output) según las pautas de Illumina. Se incorporó a las bibliotecas una agrupación de ARN total comercial de leucocitos periféricos de sangre humana (Clontech, n.º 636592) para determinar la variabilidad entre diferentes celdas de flujo de secuenciación durante el transcurso del experimento. Se agregaron bibliotecas de NGS para introducir diversidad en las bibliotecas usando PhiX Control V3 (Illumina) y se secuenciaron (secuenciación de lectura única, longitud de lectura de 150 pb) en concentraciones estándar usando secuenciación profunda (> 50.000 secuencias por muestra).

Análisis bioinformático de datos de secuenciación dirigida

Los datos de secuenciación se descargaron de los secuenciadores NextSeq o MiSeq (Illumina). La calidad de la secuenciación se realizó usando el software FastQC (versión 0.11.7, <https://github.com/s-andrews/FastQC/>). Se consideró una profundidad de secuenciación mínima de 20.000 lecturas para cada muestra para un análisis adicional. Se realizó una etapa de pretratamiento que consistió en eliminar las secuencias adaptadoras y filtrar las secuencias según su tamaño y puntuación de calidad. Se eliminaron las lecturas cortas (< 100 nucleótidos) y las lecturas con un QC promedio < 20. Para mejorar la calidad de alineación de las secuencias, se usaron herramientas de filtrado y recorte de lectura flexibles para datos NGS de Illumina (fastq_toolkit v0.0.14 y prinseq, versión 0.20.4). Después de realizar las etapas de preprocesamiento, se realizó un control de calidad adicional de cada archivo fastq limpio antes de realizar un análisis adicional.

La alineación de las lecturas procesadas se realizó usando bowtie2¹² (versión 2.2.9) con modo sensible de extremo a extremo. La alineación se realizó con la secuencia de referencia más reciente del genoma humano (GRCh38). Las alineaciones no únicas, las lecturas no alineadas o las lecturas que contienen inserción/eliminación (INDEL) se eliminaron del análisis en sentido 3' mediante el software SAMtools¹³ (versión 1.7). Se usó SAMtools mpileup¹⁴ para la denominación de SVN. La búsqueda de la posición editada en la alineación se realizó mediante secuencias de comandos internas para contar el número de nucleótidos diferentes en cada ubicación genómica. Para cada posición, se calcula el porcentaje de lecturas que apoyan el nucleótido "G" en lugar del nucleótido "A" [Número de lecturas "G"/((Número de lecturas "G" + Número de lecturas "A")*100)]. La secuencia de comandos detecta automáticamente una tasa de edición >0,1 y la considera un "sitio de edición de A a I". La última etapa es calcular el porcentaje de todas las isoformas posibles de cada transcrito. Por definición, la proporción relativa de edición de ARN en un "sitio" de edición determinado representa la suma de las modificaciones de edición medidas en esta coordenada genómica única. Por el contrario, una isoforma de ARNm es una molécula única que puede contener o no múltiples modificaciones de edición en el mismo transcrito. Por ejemplo, para un transcrito dado, la isoforma BC del ARNm contiene una modificación de A a I tanto en el sitio B como en el sitio C dentro del mismo transcrito. Se estableció una proporción relativa de al menos el 0,1 % como umbral para ser incluido en el análisis.

Análisis estadístico de datos

Todas las estadísticas y cifras se calcularon con el software estadístico de código abierto "R/Bioconductor"¹⁵ o el software GraphPad Prism (versión 7.0). Los valores de los biomarcadores (es decir, los sitios de edición de ARN y las isoformas de los genes diana y la expresión de ARNm de los ADAR) se presentan habitualmente como media ± error estándar de la media (EEM). Para garantizar una distribución normal de los datos, todos los biomarcadores podrían transformarse usando el paquete R bestNormalize (versión 1.4.2). Se realizó un análisis diferencial usando la prueba más adecuada entre la prueba de suma de rangos de Mann-Whitney, la prueba t de Student o la prueba t de Welch según la normalidad y la distribución de la varianza de la muestra. Un valor de *p* inferior a 0,05 se consideró estadísticamente significativo.

El rendimiento diagnóstico del biomarcador podría caracterizarse por: sensibilidad, que representa su capacidad para detectar el grupo "deprimido", y especificidad, que representa su capacidad para detectar el grupo de "control". Los resultados de la evaluación de una prueba diagnóstica se pueden resumir en una tabla de contingencia 2x2 comparando estos dos grupos bien definidos. Fijando un punto de corte, los dos grupos podrían clasificarse en categorías según los resultados de la prueba, categorizados o bien como positivos o bien como negativos. Dado un biomarcador particular, se puede identificar un número de pacientes A con un resultado positivo en la prueba dentro del grupo deprimido (el "verdadero positivo": TP) y pacientes B con un resultado negativo en la prueba dentro del grupo de "control" (el "verdadero negativo": TN). De la misma manera, se observan pacientes C con un resultado de prueba negativo dentro del grupo "deprimido" (el "falso negativo": FN) y pacientes D con un resultado de prueba positivo dentro del grupo de "control" (el "falso positivo": FP). La sensibilidad se define como TP/(TP + FN); que se denomina en la presente memoria "tasa de positivos verdaderos". La especificidad se define como TN/(TN + FP); que se denomina en la presente memoria "tasa de negativos verdaderos".

La precisión de cada biomarcador y su poder discriminatorio se evaluaron usando un análisis de características operativas del receptor (ROC)¹⁶. Las curvas ROC son la visualización gráfica de la relación recíproca entre la sensibilidad (Se) y la especificidad (Sp) de una prueba para varios valores. Además, todos los biomarcadores se

combinaron entre sí para evaluar el aumento potencial de la sensibilidad y especificidad usando múltiples enfoques como, por ejemplo, el programa mROC¹⁷ o bosque aleatorio¹⁸.

- 5 mROC es un programa dedicado para identificar la combinación lineal, que maximiza la curva ROC de AUC (área bajo la curva). Se proporciona la ecuación para la combinación respectiva y puede usarse como un nuevo marcador virtual Z, de la siguiente manera:

$$Z = a \times \text{biomarcador 1} + b \times \text{biomarcador 2} + c \times \text{biomarcador 3}$$

- 10 donde a, b y c son coeficientes calculados y los biomarcadores 1, 2 y 3 son el nivel del biomarcador considerado.

Se aplicó Bosque Aleatorio (RF) como anteriormente para evaluar las combinaciones de isoformas/sitios de edición de ARN. Este método combina la idea de “embolsado” de Breiman y la selección aleatoria de características para construir una colección de árboles de decisión con varianza controlada. Pueden usarse bosques aleatorios para clasificar la importancia de la edición de isoformas/sitios y combinar las mejores isoformas. Se implementó el embolsado asimétrico con reducción de tamaño múltiple (MultDS) con clasificador de variables integrado. Este método intenta aprovechar toda la información de las clases mayoritarias mediante un submuestreo múltiple pero manteniendo fija la clase minoritaria. Nuestro algoritmo generó 100 bosques aleatorios a partir del conjunto de entrenamiento, cada uno con el mismo número fijo de la clase minoritaria y un número igual o ajustado de lo muestreado aleatoriamente, generando árboles casi equilibrados. Para la clasificación de nuevos datos, los resultados se combinaron posteriormente mediante votación mayoritaria de los 100 bosques aleatorios entrenados. Para estimar los parámetros RF para cada 100 bosques aleatorios entrenados, se aplicó una validación cruzada de 10 veces. La implementación se realizó usando el paquete R de bosque aleatorio (versión 4.6-14) y el paquete R caret (versión 6.0-84).

25 Ejemplo 2: Resultados

Características de las muestras del estudio Discovery Los pacientes y los sujetos de control incluidos en las muestras de Sec-ARN (n = 32 para cada estado) fueron emparejados con respecto a todas las variables, por ejemplo, sexo, edad o IMC, que no fueron estadísticamente diferentes entre los grupos (figura 11). Se eligieron dos evaluaciones clínicas diferentes para confirmar la presencia o ausencia de diagnósticos psiquiátricos, en concreto, MADRS e IDSC-30, ambas designadas por el Manual Diagnóstico y Estadístico de los Trastornos Mentales de la Asociación Estadounidense de Psiquiatría. Los controles sanos no mostraron ningún signo de estado del ánimo deprimido en el momento de la recogida de muestras (MADRS < 7 e IDSC-30 < 10). Los pacientes deprimidos (DEP) tenían puntuaciones significativamente más altas tanto en el MADRS (valor p de 5,93E-09) como en el IDSC-30 (valor p de 2,59E-09) (tabla 1). Además, se observó una asociación significativa ($r^2=0,84$, $p<0,0001$) entre las dos evaluaciones clínicas (figura 9), sugiriendo una asignación correcta de pacientes y una homogeneidad relativa en la puntuación de la depresión.

40 Análisis del editoma

Para obtener un panorama global de la modificación de los eventos de edición de ARN en pacientes deprimidos, se analizó el conjunto de datos de Sec-ARN usando una canalización de análisis del editoma interna (véanse los métodos y la figura 10). Identificamos 37.599 posiciones editadas diferencialmente de A a I con una cobertura mínima de 30x, lo que garantiza un alto grado de confianza en posiciones de bases particulares. La cobertura describe el número promedio de lecturas que se alinean con o “cubren” bases de referencia conocidas. La mayoría de las posiciones editadas (sitios 32, 183 sitios, un 85,6 % del total) estaban presentes en intrones (figura 1A), en consonancia con otros estudios en tejidos humanos. El porcentaje de posiciones editadas en regiones 3' no traducidas (3'UTR) de ARNm fue del 6,2 % (2344 sitios), en línea con informes anteriores que usaron Sec-ARN. El número de sitios de edición tuvo una distribución homogénea por todo el genoma, aunque los genes editados en los cromosomas 4, 9, 18, 19, 20, 21 y 22 estuvieron ligeramente sobrerrepresentados (Figura 1B). A continuación, realizamos un análisis diferencial para identificar sitios cuya edición podría ser específicamente diferente entre DEP y controles sanos. Después de aplicar criterios de calidad preestablecidos (exclusión de sitios intergénicos, cobertura >30; AUC >0,6; 0,95 >Múltiplo de variación >1,05 y $p < 0,05$), se identificaron 646 variantes editadas diferencialmente entre pacientes deprimidos y controles sanos (Figura 1C), representando 366 genes. Los criterios de calidad se establecieron con el fin de proporcionar biomarcadores (BMK) de diagnóstico potenciales para trastornos depresivos con utilidad clínica. Es importante destacar que no se observaron diferencias en la edición global del ARN entre pacientes y controles. De hecho, se calculó el índice de edición de Alu (AEI) global, que es una medida de la tasa global de edición de ARN en repeticiones de Alu. Las repeticiones de Alu son, con diferencia, la región editada más común en el genoma humano y, por tanto, son indicativas de la tasa de edición global. No se hallaron diferencias significativas en el AEI entre los controles y los pacientes deprimidos (media de control de AEI = $0,300 \pm 0,002$, media de DEP de AEI = $0,299 \pm 0,003$, valor de $p = 0,57$, figura 1D). Como el índice de edición de Alu caracteriza el nivel de edición promedio ponderado en todas las secuencias de Alu expresadas, estos datos sugieren que la edición diferencial del ARN era específica de la diana en lugar de depender del estado de la enfermedad.

65 GSEA (análisis de enriquecimiento en conjuntos de genes) sobre estos 366 genes usando herramientas de ontología genética¹⁹ mostró un fuerte enriquecimiento de términos (tasa de falsos descubrimientos < 0,05) para los procesos

biológicos, incluyendo múltiples categorías inmunitarias (respuesta, activación y regulación), activación celular y procesos metabólicos (figura 12). Este análisis se repitió con otras herramientas, como G:profiler²⁰ (figura 9) o GOnet²¹ (figura 10), e indicó consistentemente un fuerte enriquecimiento en los procesos biológicos relacionados con el sistema inmunitario. Además, el análisis de las rutas de Reactome de esos 366 genes reveló un fuerte enriquecimiento en rutas relacionadas con el sistema inmunitario, es decir, la señalización de TCR, la translocación de ZAP-70 a la sinapsis inmunitaria o la señalización del interferón gamma (figura 13). Cabe destacar que el análisis de componentes principales (PCA) mostró que este subconjunto de las 646 variantes de edición que representan 366 genes discriminaban claramente entre controles y pacientes deprimidos en 2 grupos separados (Figura 2), lo que sugiere que las variantes de edición en esos genes podrían usarse en combinación para discriminar selectivamente entre pacientes deprimidos y controles sanos.

Para acotar adicionalmente la lista de genes, se usaron criterios de calidad mucho más estrictos (cobertura >30; AUC >0,8; 0,8 >Múltiplo de variación>1,20 y $p < 0,05$). A continuación, la base de datos de estos marcadores se seleccionó manualmente para alcanzar una combinación de varios biomarcadores que representan diferentes mecanismos biológicos. Usando estos criterios secundarios, se seleccionó una lista de 15 variantes, que representan 13 genes (figura 14), tanto por su rendimiento diagnóstico como por su papel potencial en el sistema inmunitario o la biología de la enfermedad psiquiátrica y para estudiarse en una cohorte de priorización posterior. Además, se nominó otra diana, PDE8A, para el cual los datos previos indicaron una edición alterada en el cerebro suicida y en la sangre de la depresión inducida por fármacos en pacientes con VHC^{5,22}. El análisis de redes de asociación proteína-proteína con la base de datos STRING²³ reveló una red con fuerte enriquecimiento en la interacción proteína-proteína (valor de p de PPI <4,5 · 10⁻⁶). Las interacciones dentro de STRING se basan en evidencia y consisten en interacciones directas (físicas) e indirectas (funcionales); en el presente caso nos centramos en una puntuación de interacción requerida mínima de 0,7, lo que es indicativo una confianza alta en las interacciones proteína-proteína reales. El análisis de red de asociación de proteínas funcionales de los 14 genes seleccionados usando herramientas de ontología genética¹⁹ también mostró un fuerte enriquecimiento en genes relacionados con: regulación de proceso de sistema inmunitario (8 genes, GO:0002682; $p(\text{TDF}) = 4 \cdot 10^{-4}$), regulación de la respuesta inmunitaria, (6 genes, (GO:0050776, $p(\text{TDF}) = 2,2 \cdot 10^{-3}$) o ruta de señalización del receptor a través de JAK-STAT (4 genes, GO:0007259), que generalmente se activa después de la estimulación con interferón-alfa (figura 3 y figura 15). Además, el análisis de las rutas de Reactome indicó que esos 14 genes estaban enriquecidos en rutas relacionadas con el sistema inmunitario (por ejemplo, regulación de la señalización de IFNA (HSA-912694, $p(\text{TDF}) < 0,025$) o señalización de citocinas (HSA-1280215, $p(\text{TDF}) < 0,026$)) y regulación de neurotransmisores (por ejemplo, tráfico de receptores AMPA (HSA-399719, $p(\text{TDF}) < 0,025$), unión de glutamato y plasticidad sináptica (HAS-399721, $p(\text{TDF}) < 0,025$) o receptores de neurotransmisores y transmisión de señales postsinápticas (HSA-112314, $p(\text{TDF}) < 0,029$)) (figura 16). Resulta interesante observar que la ruta Top identificada en el análisis de Reactome ($p(\text{TDF}) < 0,012$) está vinculada al factor de transcripción Runx1, que recientemente se ha mostrado que media el efecto del estrés crónico por derrota social en un modelo de depresión en ratones a través de la regulación de la familia miR-30 de miARN.

Descripción de la diana

PRKCB: Se sabe bien que la proteína quinasa C beta (PRKCB), que pertenece a la familia de las proteínas quinasas específicas de serina y treonina, desempeña un papel crucial en el sistema inmunitario, ya que los ratones con genes inactivados para PKCB desarrollan una inmunodeficiencia caracterizada por una activación deficiente de las células B, una respuesta del receptor de las células B y defectos en las respuestas inmunitarias independientes de las células T. PRKCB también actúa como un regulador de la respuesta del eje HPA al estrés²⁴. Un estudio de asociación de 44 genes candidatos con síntomas depresivos y de ansiedad en mujeres posparto reveló una asociación entre PRKCB y la depresión mayor. De hecho, PRKCB interactúa con CREB1 y lo fosforila, lo que regula directamente la expresión de varios genes implicados en la respuesta al estrés en el cerebro, tales como BDNF, el receptor de BDNF Trk-b y el receptor de glucocorticoides²⁵.

MDM2: MDM2 es un inhibidor celular primario de p53 y podría ser una diana para el tratamiento contra el cáncer. Sin embargo, MDM2 también cataliza la ubiquitinación de la β -arrestina, que es una diana de los antidepresivos. En el SNC, MDM2 está implicado en la expresión de superficie de AMPAR (receptor de glutamato) durante la plasticidad sináptica²⁶. Resulta interesante observar que se ha mostrado que la edición de ARN en MDM2 regula los niveles de proteínas. De hecho, el aumento de la edición de ARN en el UTR 3' de MDM2 elimina la represión mediada por microARN, lo que puede aumentar sus niveles de ARNm.

PIAS1: PIAS1 (inhibidor proteico de STAT 1 activado) funciona como una SUMO ligasa que bloquea la actividad de unión al ADN de Stat1 e inhibe la activación genética mediada por Stat1 en respuesta al interferón. La sumoilación de factores de transcripción (HSA-3232118) es una de las principales rutas de Reactome en los 14 genes seleccionados para el análisis en la cohorte de priorización (figura 16). PIAS1 inhibe selectivamente los genes inducibles por interferón y es un actor importante en la inmunidad innata²⁷. Además, Pias1 interactúa con el receptor de glutamato metabotrópico presináptico 8 (mGluR8) y agrega una modificación postranscripcional por sumoilación, lo que sugiere que PIAS1 podría tener funciones reguladoras en la orientación y modulación de la señalización del receptor mGluR8.

IFNAR1: Numerosos estudios han vinculado los interferones con cambios inducidos por la inflamación en la función cerebral y la depresión, al menos en parte a través del efecto de IL-6, CXCL10 o por la inducción de la indolamina 2,3-

dioxigenasa1 (IDO1)²⁸. Además, se ha mostrado que los polimorfismos en la región promotora del receptor IFN-alfa/beta 1 (IFNAR1) pueden influir en el riesgo de desarrollar depresión. Los interferones se usan para el tratamiento de la hepatitis viral, trastornos hematoproliferativos, trastornos autoinmunitarios y enfermedades malignas. Sin embargo, independientemente de la indicación de la terapia con IFN, los IFN se asocian con un riesgo del 30 al 70 % de depresión emergente con el tratamiento. En un trabajo previo hemos mostrado que los biomarcadores de edición de ARN permitieron discriminar entre pacientes que desarrollaron una depresión emergente con el tratamiento y aquellos que no.

IFNAR2: Los interferones actúan a través de sus receptores; el receptor de interferón tipo I (IFNAR) está compuesto por dos subunidades, IFNAR1 e IFNAR2. Un polimorfismo en el gen IFNAR2 (SNP 11876) mostró una asociación con la susceptibilidad a la esclerosis múltiple ($p = 0,035$). Resulta interesante observar que el mismo estudio también identificó un polimorfismo en el gen IFNAR1 (SNP 18417), que es miembro del mismo receptor.

LYN: LYN codifica una proteína tirosina quinasa no receptora que transmite señales desde los receptores de la superficie celular. Lyn desempeña un papel importante en la regulación de las respuestas inmunitarias innatas y adaptativas²⁹. En el sistema nervioso central de los mamíferos, Lyn se expresa en gran medida en el núcleo *accumbens*, la corteza cerebral y el tálamo. Además, Lyn participa en el control de la ruta del receptor N-metil-D-aspartato (NMDAR, un receptor de glutamato ionotrópico). De hecho, las pruebas de comportamiento mostraron que la actividad motora espontánea se suprime en los ratones Lyn^{-/-}, debido a la potenciación de la señalización del receptor NMDA. Además, Lyn interactúa físicamente con el receptor AMPA, un miembro de la familia de receptores de glutamato³⁰. Tras la estimulación del receptor, Lyn activa la ruta de señalización de la proteína quinasa activada por mitógeno (MAPK), lo que a su vez aumenta la expresión del factor neurotrófico derivado del cerebro (BDNF). Es importante destacar que la activación de receptores AMPA junto con la liberación de BDNF y la activación de las rutas de señalización sinaptogénica en sentido 3' median los efectos antidepresivos de la ketamina⁷⁹. Las interacciones entre el receptor de glutamato y la Lyn cinasa experimentan adaptaciones drásticas en trastornos del estado de ánimo tales como el trastorno depresivo mayor.

AHR: El receptor de hidrocarburos arílicos (AHR) es un factor de transcripción dependiente de ligando altamente conservado que detecta toxinas ambientales y ligandos endógenos. El AHR también modula la diferenciación y la respuesta de las células inmunitarias³¹. AHR también controla la diferenciación de las células tanto T(reg) como Th17 de manera específica para cada ligando. La activación de AHR induce IDO y, por tanto, aumenta los niveles de quinurenina. El papel de la quinureína está bien documentado en el SNC, donde desempeña un papel importante en la fisiopatología de la depresión inducida por inflamación. El bloqueo farmacológico de AHR rescató el tráfico de monocitos inducido por la inflamación sistémica, el trastorno neuroinmunitario y el comportamiento de tipo depresivo en ratones. La edición de ARN afecta a la 3'UTR del ARNm de AHR humano, creando de ese modo una secuencia de reconocimiento de microARN que modula la expresión de AHR.

RASSF1: El miembro 1 de la familia del dominio de asociación Ras (RASSF1) RASSF1A es un gen supresor de tumores cuya inactivación se ha relacionado con una amplia variedad de cánceres humanos esporádicos y se inactiva epigenéticamente con alta frecuencia en una variedad de tumores sólidos. RASSF1A se asocia con MDM2 y potencia la actividad de la autoubiquitina ligasa de MDM2³². Resulta interesante observar que un análisis de micromatrices identificó recientemente RASSF1A como un gen significativamente regulado a la baja en mujeres con antecedentes de depresión posparto.

GAB2: Las proteínas Gab funcionan en sentido 3' de múltiples receptores, tal como los receptores de EGF, el factor de crecimiento de hepatocitos y los FGF. GAB2 es un componente importante de la diferenciación neuronal inducida por ácido retinoico, en parte al actuar en sentido 3' del bFGF para mediar la supervivencia a través de la PI3 quinasa-AKT. GAB2 aumentó positivamente la expresión de IL-4 e IL-13 en células T activadas, lo que sugiere que GAB2 podría ser un regulador importante de la respuesta inmunitaria Th2 humana³³.

CAMK1D: CAMK1D es un miembro de la familia de la proteína quinasa 1 dependiente de calcio/calmodulina y está implicado en la ruta de transducción de señales de quimiocinas que regula la función de los granulocitos³⁴. En el SNC, CAMK1D se expresa principalmente en la región CA1 del hipocampo, donde se induce después de la potenciación a largo plazo (PLP), una potenciación duradera de la eficacia sináptica en el sistema nervioso central. Podría translocarse al núcleo en respuesta a estímulos que desencadenan la afluencia de Ca²⁺ intracelular, por ejemplo, glutamato, y activar la transcripción de genes dependiente de CREB. Cabe destacar que SNP rs2724812 en CAMK1D se ha asociado con la remisión después del tratamiento con inhibidores selectivos de la recaptación de serotonina mediante un enfoque de aprendizaje profundo³⁵.

KCNJ15: El miembro 15 de la subfamilia J del canal dependiente de voltaje de potasio (KCNJ15) codifica un canal de potasio de tipo rectificador interno, Kir4.2. En el SNC, el canal KIR4.2 se expresa tardíamente en el desarrollo tanto en ratones como en el cerebro humano. Los canales de potasio de rectificación interna (canales Kir) son reguladores importantes de la señalización neuronal y la excitabilidad de la membrana. La activación de las subunidades del canal Kir desempeña un papel crucial en las enfermedades poligénicas del SNC. Además, se halló que rs928771 en KCNJ15 estaba asociado con alteraciones en los niveles de transcrito de KCNJ15 en sangre completa; el análisis de redes de

conjuntos de datos del transcriptoma de la sangre y del hipocampo sugirió que las variantes de riesgo en el locus KCNJ15 podrían actuar a través de sus efectos reguladores sobre las rutas relacionadas con el sistema inmunitario³⁶.

IL17RA: El receptor A de interleucina 17 (IL17RA) es una glicoproteína de membrana tipo I, que tiene una estructura similar a IFNAR1, que se une con baja afinidad a la interleucina 17A. La interleucina 17A (IL17A) es una citocina proinflamatoria secretada principalmente por linfocitos Th17 activados. Al unirse a su receptor cognado, IL17 envía señales a través de las rutas ERK, MAPK y NFκB, la totalidad de las cuales se han implicado en el trastorno depresivo. Las células Th17 son potentes reguladores de la función cerebral, ya que se dirigen a varias células residentes del cerebro, incluyendo neuronas, astrocitos y microglía, y participan en la neuroinflamación³⁷. IL17RA se expresa en gran medida en las células endoteliales de la barrera hematoencefálica (BHE) en las lesiones de esclerosis múltiple, donde IL17 aumenta la permeabilidad de la BHE al alterar uniones estrechas, promoviendo de ese modo la inflamación del SNC. La IL17 proinflamatoria también es secretada por microglía y astrocitos activados y tiene un efecto perjudicial sobre las neuronas. El análisis de citometría de flujo de las PBMC en sujetos deprimidos no medicados reveló un aumento significativo en el número de células Th17 periféricas y una concentración sérica elevada de IL-17, lo que sugiere un papel potencial de las células Th17 en pacientes con TDM. De manera similar, las células Th17 se aumentan en el cerebro de los ratones después de una situación de indefensión aprendida o estrés crónico por restricción, que son dos modelos comunes similares a la depresión. La inhibición de la producción o función de las células Th17 reduce la vulnerabilidad a un comportamiento similar a la depresión en los mismos modelos de ratón

PTPRC: La proteína tirosina fosfatasa, receptor tipo C (PTPRC), también conocida como antígeno leucocitario común o CD45, es una proteína tirosina fosfatasa de membrana integral, que contiene un segmento externo y es capaz de iniciar la señalización transmembrana en respuesta a ligandos externos³⁸. En el sistema inmunitario, PTPRC es esencial para la activación de células T y B al mediar los contactos entre células y regular las proteínas tirosina quinasas implicadas en la transducción de señales. PTPRC suprime y regula negativamente la señalización del receptor de citocinas; la interrupción selectiva de la expresión de PTPRC conduce a una activación mediada por el receptor de interferón potenciada de las proteínas quinasas JAK y STAT³⁹. Cabe destacar que una mutación puntual en PTPRC (transición de nucleótido C a G en la posición 77 del exón 4) se ha asociado con el desarrollo de la esclerosis múltiple, la enfermedad desmielinizante más común del SNC. En el cerebro, el comportamiento similar a la ansiedad inducido por la derrota social repetida, un modelo de depresión en ratones, se correspondió con un aumento en los macrófagos cerebrales que sobreexpresaban PTPRC. En el mismo modelo de depresión en ratones, se mostró que las benzodiazepinas clásicas revierten los síntomas y bloquean la acumulación inducida por el estrés de macrófagos que expresan niveles altos de PTPRC en el SNC. Además, en pacientes diagnosticados con trastorno depresivo mayor, el PTPRC fue mayor en los pacientes que en los sujetos sanos. La relación de PTPRC disminuyó con tratamiento antidepresivo tal como venlafaxina o fluoxetina, y se correlacionó positivamente con la gravedad de la depresión⁴⁰, lo que sugiere que el tratamiento antidepresivo contribuye a la regulación inmunológica en pacientes con trastorno depresivo mayor.

PDE8A: Los receptores de glutamato o serotonina son receptores acoplados a proteína G (GPCR) implicados en una variedad de trastornos psiquiátricos, para los cuales la regulación espacial y temporal de la concentración de segundos mensajeros es crucial para la señalización posterior. Las fosfodiesterasas (PDE) son moduladores clave de la transducción de señales y están implicadas en la activación de células inflamatorias, el comportamiento, la memoria y la cognición. PDE8A es una fosfodiesterasa específica de AMPc que se expresa en el cerebro y la sangre. Se ha identificado una edición alterada del ARN en las transcripciones de PDE8A en células T de pacientes que padecen lupus eritematoso sistémico, un trastorno autoinmunitario crónico, al igual que en células T normales con IFN-α. Hay una disminución del doble en la expresión de la fosfodiesterasa 8A (PDE8A) en la corteza temporal de pacientes con trastorno depresivo mayor (MDD)⁴¹. Recientemente se han mostrado alteraciones específicas de la región cerebral en la edición de ARN en el ARNm de PDE8A en fallecidos por suicidio, lo que proporcionó el primer marcador cerebral relacionado con la respuesta inmunitaria para el suicidio⁵. La edición del ARN de PDE8A también aumenta en las células de neuroblastoma SHSY-5Y tratadas con interferón alfa, como en los pacientes con VHC sometidos a terapia de combinación antiviral con IFN-α y ribavirina. Es importante destacar que la medición de la edición de ARN en PDE8A permitió discriminar entre el grupo de pacientes que desarrollaron una depresión emergente con el tratamiento y los que no, lo que sugiere que los biomarcadores de edición de ARN A-I podrían ser útiles para monitorizar la depresión que surge con el tratamiento²².

Con el fin de priorizar un subconjunto de variantes clínicamente útiles para un ensayo de diagnóstico, cada uno de los genes diana descritos anteriormente se sometió a prueba individualmente en dos cohortes de “priorización” separadas (n = 76, figura 17 y n = 86, figura 18) midiendo la tasa de edición en controles y en pacientes deprimidos mediante NGS. En la primera cohorte, cada diana se secuenció individualmente, mientras que en la segunda cohorte se midieron las tasas de edición de biomarcadores multiplexando 4 BMK. Se realizó un método estadístico de análisis multivariante, evaluando para cada combinación de sitios editados de ARN y/o isoformas, para hallar la mejor combinación de biomarcadores. Para potenciar el rendimiento diagnóstico, minimizando el número de biomarcadores, se calculó la mejor combinación con diferentes números de biomarcadores.

La lista de cebadores usados para la detección de sitios de edición de ARN se indica en la tabla 1.

Tabla 1: Cebadores usados para la detección de sitios de edición de ARN

NOMBRE DEL CEBADOR	Nombre del gen	Tamaño del producto de PCR (pb)	Secuencia de cebador (5' -> 3')	Longitud de cebador (pb)	Id. de sec. n.º
IFNAR1 DIRECTO_1	IFNAR1	259	ATACGGGCAAGCTCTTAACTATT	23	1
IFNAR1 INVERSO_1			AACCTCCACCTCCTGTTTTTC	20	2
IFNAR1_MiSeqF	IFNAR1	472	AAAATACGGGCAAGCTCTTAACT	23	3
IFNAR1_MiSeqR			CTGCACTTAGGCTCACAGGA	20	4
PDE8A Directo	PDE8A	225	ATGCAAGTTGTGGACATGGAG	21	5
PDE8A Inverso			TTCTGAAAACAATGGGCACCA	21	6
CAMK1DPCR0_F	CAMK1D	577	TGTGCAAAGGGCTCCATAC	19	7
CAMK1DPCR0_R			CCTTGCAAGTGTCCAGTAAAC	21	8
CAMK1D MPX_F1		195	CTAGGGTGAGGTGGTGCAAT	20	9
CAMK1D MPx R1			CAAGGCCAGGAGTTCAAGAC	20	10
LYN_F_PCR0	LYN	358	TTTAAAAGCTTCTCCAGTCAGC	22	11
LYN_R_PCR0			CAAGACTGAGCATACATTATATTCACC	27	12
LYN MPX_F1		193	TTTAAAAGCTTCTCCAGTCAGC	22	13
LYN MPX_R1			TCCTGAGTAGCTGGGACCAT	20	14
PRKCB_MPx_F1	PRKCB	176	CAGCACATTCTGTTGCCTTC	20	15
PRKCB MPx R2			GTGACTTTGCCCTCATTTG	20	16
KCNJ15_MPx_F1	KCNJ15	160	GGCCACTCTGTACTATGCAAATC	23	17
KCNJ15 MPx R2			TGAATTTAGATGGACAGCAA	21	18
MDM2_F5 MPx-Rev	MDM2	302	AAGCCCATCCTCGAACTACAG	21	19
MDM2_R5 MPx-Rev			TTGCTCTGTCACCTGGACTG	20	20
GAB2 MPx_F1	GAB2	154	GAACCAGGCTAGGTGCAATG	20	21
GAB2 MPx_R2			CCATGCCCTGGTAATTTTTTG	20	22
AHR-F1	AHR	409	TCTGATGGAATCCCAAAGTGAAGA	23	23
AHR_F2			TCAGAGATATTTTGTCCGAAGCCA	24	24
IFNAR2 MPx_F1_rev	IFNAR2	344	CCCACAGTTTCAGAGGTGGT	18	25

NOMBRE DEL CEBADOR	Nombre del gen	Tamaño del producto de PCR (pb)	Secuencia de cebador (5' -> 3')	Longitud de cebador (pb)	Id. de sec. n.º
IFNAR2 MPx_R1_rev			GGCTCACCCCTGTAATCC	20	26
IL17RA_Directo1	IL17RA	172	GGGATGTAAACGCTGAATGG	20	27
IL17RA_Inverso1			ATGTGCCACCATGACTGACT	20	28
RASSF1_F1		480	CCCTTCCGGCAAGAATCCAG	20	29
RASSF1_R1			ACCCAAGCTGCATAAAAGGC	20	30
PIAS1_F1	PIAS1	407	GCAACCTTAGATCTCAATG GCTAA	24	31
PIAS1_R2			AGCAGAGTCACAGTCTACAGC	21	32
PTPRC-MiSeqF	PTPTC	505	CCGGGGCTCATTTATAGGAT	20	33
PTPRC-MiSeqR			TGCTAGATACCGCCACATTG	20	34
PTPRC F rd2-1	PTPRC	215	TAAGGACGGAGTTCGAGACC	20	35
PTPRC R rd2-1			TGAGAGAGTTTTGCTCTGTGG	21	36

Como era importante validar los resultados obtenidos en el estudio de Sec-ARN, en primer lugar realizamos un análisis diferencial de cada diana en las cohortes de priorización. Como se muestra en las tablas siguientes (tablas 2A-2B y 3), se halló que todos los sitios e isoformas identificados en el estudio de Sec-ARN eran significativos en un análisis posterior sobre una población más grande.

Tablas 2A y 2B: Análisis diferencial de sitios de edición (2A) e isoformas (2B) identificados en el estudio de Sec-ARN en los pacientes de la primera cohorte de priorización analizados en monoplex.

Se muestran ejemplos de datos de Sec-ARN de diferentes dianas con nombre de diana, sitio o isoforma editado diferencialmente, valor de p y múltiplo de variación del sitio o isoforma editado, curva ROC de AUC, el valor medio de la edición en el sitio considerado y el número de población.

Tabla 2A

ID	Diana	Cr.:Pos (hg38)	Sitios	Valor de p	Múltiplo de variación	Curva ROC de AUC	% de edición (media)	Población
1	AHR	7:17345070	X118	0,0318	1,80	0,635	0,52	N= 76 [41 (0); 35 (1)]
3	IFNAR1	21:33355789	X111	0,0239	1,07	0,651	19,16	N= 76 [41 (0); 35 (1)]
4	IFNAR1	21:33355793	X115	0,0098	1,11	0,678	10,89	N= 76 [41 (0); 35 (1)]
5	IFNAR1	21:33355807	X129	0,0337	1,06	0,647	20,53	N= 76 [41 (0); 35 (1)]
6	IFNAR1	21:33355838	X160	0,0030	1,13	0,682	7,60	N= 76 [41 (0); 35 (1)]
7	IFNAR1	21:33355840	X162	0,0069	1,10	0,681	10,05	N= 76 [41 (0); 35 (1)]
8	IFNAR1	21:33355854	X 176	0,0004	1,18	0,745	0,54	N= 76 [41 (0); 35 (1)]

ES 2 993 857 T3

ID	Diana	Cr.:Pos (hg38)	Sitios	Valor de <i>p</i>	Múltiplo de variación	Curva ROC de AUC	% de edición (media)	Población
9	IFNAR1	21:33355860	X182	0,0142	1,13	0,652	4,30	N= 76 [41 (0); 35 (1)]
10	IFNAR1	21:33355865	X187	0,0019	1,15	0,698	1,96	N= 76 [41 (0); 35 (1)]
11	IFNAR1	21:33355869	X191	0,0059	1,14	0,688	7,47	N= 76 [41 (0); 35 (1)]
12	IFNAR1	21:33355905	X227	0,0066	1,10	0,681	5,45	N= 76 [41 (0); 35 (1)]
13	IFNAR1	21:33355909	X231	0,0121	1,10	0,661	10,39	N= 76 [41 (0); 35 (1)]
14	IFNAR1	21:33355762	X84	0,0056	1,12	0,678	10,56	N= 76 [41 (0); 35 (1)]
15	IFNAR1	21:33355763	X85	0,0067	1,12	0,681	6,25	N= 76 [41 (0); 35 (1)]
16	IL17RA	22:17110647	X151	0,0278	1,05	0,637	8,40	N= 76 [41 (0); 35 (1)]
17	IL17RA	22:17110680	X184	0,0049	1,13	0,674	0,83	N= 76 [41 (0); 35 (1)]
18	IL17RA	22:17110551	X55	0,0164	1,14	0,655	1,41	N= 76 [41 (0); 35 (1)]
20	IL17RA	22:17110574	X78	0,0049	1,09	0,675	1,95	N= 76 [41 (0); 35 (1)]
21	IL17RA	22:17110591	X95	0,0411	1,05	0,629	15,97	N= 60 [41 (0); 19 (1)]
22	LYN	8:56012521	X116	0,0266	1,12	0,652	0,97	N= 76 [41 (0); 35 (1)]
23	LYN	8:56012553	X148	0,0238	1,06	0,676	18,79	N= 60 [41 (0); 19 (1)]
24	LYN	8:56012558	X153	0,0074	1,10	0,661	3,81	N= 76 [41 (0); 35 (1)]
25	LYN	8:56012689	X284	0,0140	1,10	0,633	4,34	N= 76 [41 (0); 35 (1)]
26	LYN	8:56012705	X300	0,0161	1,10	0,652	3,03	N= 76 [41 (0); 35 (1)]
27	LYN	8:56012490	X85	0,0049	1,13	0,679	0,75	N= 76 [41 (0); 35 (1)]
28	MDM2	12:68821545	X166	0,0142	0,91	0,661	2,91	N= 76 [41 (0); 35 (1)]
29	MDM2	12:68821629	X250	0,0215	0,89	0,649	4,42	N= 76 [41 (0); 35 (1)]
30	PIAS1	15:68138001	X201	0,0446	1,09	0,664	0,64	N= 60 [41 (0); 19 (1)]
31	PTPRC	1:198680346	X240	0,0216	1,06	0,689	25,83	N= 60 [41 (0); 19 (1)]
32	PTPRC	1:198680405	X299	0,0263	1,08	0,675	4,30	N= 60 [41 (0); 19 (1)]
33	RASSF1	3:50333653	X192	0,0204	1,09	0,654	1,32	N= 60 [41 (0); 19 (1)]
35	PDE8A	15:85212876	X116246	0,0055	0,92	0,668	19,81	N= 76 [41 (0); 35 (1)]

Tabla 2B

Diana	Isoforma	Valor de p	Múltiplo de variación	Curva ROC de AUC	% de edición (media)	Población
IFNAR1	A	0,0002	1,13	0,735	0,53	N= 76 [41 (0); 35 (1)]
IFNAR1	ABCDEFGH	0,0042	1,19	0,686	0,99	N= 76 [41 (0); 35 (1)]
IFNAR1	GH	0,0002	1,16	0,742	0,73	N= 76 [41 (0); 35 (1)]
LYN	DO	0,0275	0,94	0,660	1,45	N= 76 [41 (0); 35 (1)]
MDM2	GVA	0,0023	1,09	0,710	0,94	N= 76 [41 (0); 35 (1)]
MDM2	W	0,0154	0,93	0,669	0,45	N= 76 [41 (0); 35 (1)]
PDE8A	B	0,0059	0,92	0,660	19,00	N= 76 [41 (0); 35 (1)]
PIAS1	F	0,0359	1,04	0,641	1,10	N= 76 [41 (0); 35 (1)]
PTPRC	HI	0,0331	1,07	0,623	1,61	N= 76 [41 (0); 35 (1)]
RASSF1	ABE	0,0224	1,04	0,637	1,87	N= 76 [41 (0); 35 (1)]
RASSF1	ABEH	0,0290	1,07	0,695	0,69	N= 76 [41 (0); 35 (1)]
RASSF1	HJ	0,0260	0,93	0,624	0,83	N= 76 [41 (0); 35 (1)]

Tabla 3: Análisis diferencial de sitios de edición e isoformas identificados en el estudio de Sec-ARN en los pacientes de la segunda cohorte de priorización analizados con secuenciación multiplexada.

Se muestran ejemplos de datos de Sec-ARN de diferentes dianas con nombre de diana, sitio editado diferencialmente, valor de p y múltiplo de variación del sitio editado, curva ROC de AUC, el valor medio de la edición en el sitio considerado y el número de población.

Diana	Cr.:Pos (hg38)	Sitios	Valor de p	Múltiplo de variación	Curva ROC de AUC	% de edición (media)	Población
CAMK1D	10:12378423	X131	0,0136	0,90	0,659	1,33	N= 86 [28 (0);58 (1)]
CAMK1D	10:12378341	X49	0,0303	0,92	0,639	5,59	N= 86 [28 (0);58 (1)]
CAMK1D	10:12378415	X123	0,0387	0,93	0,698	25,23	N= 30 [15 (0);15 (1)]
GAB2	11:78330952	X88	0,0004	0,88	0,705	1,5	N= 86 [28 (0);58 (1)]
GAB2	11:78330914	X126	0,0260	0,96	0,610	17,18	N= 86 [28 (0);58 (1)]
IFNAR1	21:33355779	X101	0,0033	0,87	0,717	0,87	N= 86 [28 (0);58 (1)]
IFNAR1	21:33355820	X142	0,0159	0,89	0,663	1,3	N= 86 [28 (0);58 (1)]
IFNAR1	21:33355754	X76	0,0330	0,91	0,643	4,4	N= 86 [28 (0);58 (1)]

Diana	Cr.:Pos (hg38)	Sitios	Valor de p	Múltiplo de variación	Curva ROC de AUC	% de edición (media)	Población
KCNJ15	21:38287590	X134	0,0099	0,89	0,673	3,6	N= 86 [28 (0);58 (1)]
KCNJ15	21:38287586	X130	0,0308	0,93	0,650	25,2	N= 86 [28 (0);58 (1)]
KCNJ15	21:38287585	X129	0,0237	0,91	0,647	3,2	N= 86 [28 (0);58 (1)]
KCNJ15	21:38287589	X133	0,0197	0,90	0,644	14,53	N= 86 [28 (0);58 (1)]
KCNJ15	21:38287587	X131	0,0488	0,91	0,634	0,9	N= 86 [28 (0);58 (1)]

LYN	8:56012495	X90	0,0040	0,89	0,696	1,2	N= 86 [28 (0);58 (1)]
LYN	8:56012560	X155	0,0443	1,09	0,635	0,6	N= 86 [28 (0);58 (1)]
LYN	8:56012553	X148	0,0435	0,99	0,532	15,7	N= 57 [28 (0);29 (1)]
MDM2	12:68821586	X207	0,0004	0,89	0,746	12,58	N= 86 [28 (0);58 (1)]
MDM2	12:68821522	X143	0,0018	0,79	0,699	0,76	N= 86 [28 (0);58 (1)]
MDM2	12:68821570	X191	0,0296	0,92	0,652	1,09	N= 86 [28 (0);58 (1)]
PDE8A	15:85212876	X116248	0,0032	0,90	0,680	20,3	N= 86 [28 (0);58 (1)]
PRKCB	16:23920909	X103	0,0236	0,94	0,649	5,7	N= 86 [28 (0);58 (1)]
PRKCB	16:23920932	X126	0,0235	0,94	0,649	1,5	N= 86 [28 (0);58 (1)]
PRKCB	16:23920938	X132	0,0360	0,96	0,637	17,3	N= 86 [28 (0);58 (1)]
PTPRC	1:198680365	X259	0,0466	0,94	0,625	0,8	N= 86 [28 (0);58 (1)]
PTPRC	1:198680332	X226	0,0295	0,95	0,664	13,0	N= 57 [28 (0);29 (1)]
MDM2	12:68848833	X76	0,0035	0,83	0,704	0,5	N= 86 [28 (0);58 (1)]
MDM2	12:68848843	X86	0,0410	0,89	0,664	1,0	N= 86 [28 (0);58 (1)]
MDM2	12:68848844	X87	0,0388	0,95	0,639	4,8	N= 86 [28 (0);58 (1)]

A continuación, se analizaron combinaciones de biomarcadores para identificar el mejor rendimiento diagnóstico con diferentes números de biomarcadores. En la primera cohorte, cuando sólo se usaron 4 sitios de 3 dianas diferentes, se obtuvo un AUC de 0,811 (Sp 78,1 %; Se 81,3 %, figura 4A). Otra combinación de 4 sitios, usando 2 BMK, dio resultados aún mejores, con un AUC de 0,866 (Sp 92,7 %; Se 79,0 %, figura 4B). La ecuación usada para esta combinación es la siguiente:

$$Z = + (0,361161202728045) * \text{IFNAR1_X227_H} + (-0,0763586204473731) * \text{IFNAR1_X162_E} + (1,35409123139041) * \text{LYN_X284_P} + (-2,2905526156995) * \text{LYN_X207_M}.$$

Si bien la medición de las variantes de edición del ARN arrojó buenos resultados en todas las combinaciones sometidas a prueba, el rendimiento no se reflejó necesariamente en el número de biomarcadores: cuando el número de biomarcadores aumentó a 6, con 4 dianas diferentes, el AUC fue de 0,794 (Sp 80,5 %; Se 77,1 %, figura 4C). Sin embargo, cuando se combinaron 5 BMK y se analizaron todos los sitios editados de manera diferencial en esas dianas, se obtuvo un rendimiento sobresaliente, con un AUC de 0,977, con una excelente especificidad y sensibilidad (92,7 % y 94,7 %, respectivamente, figura 4D). La ecuación usada para esta combinación es la siguiente:

$$Z = + (1,242350672584) * \text{IFNAR1_X227_H} + (1,36013909308819) * \text{IFNAR1_X115_G} + (-0,739944712498771) * \text{IFNAR1_X84_F} + (1,42700898962888) * \text{IFNAR1_X162_E} + (-1,66838875199491) * \text{IFNAR1_X231_D} + (-0,887010245248778) * \text{IFNAR1_X129_B} + (0,0876014897108424) * \text{IFNAR1_X111_A} + (0,985129912485808) * \text{LYN_X153_G} + (4,33197064304577) * \text{LYN_X284_P} + (-0,200203331256123) * \text{IFNAR1_X152_C} + (-15,708397165937) * \text{IL17RA_X78_D} + (2,50160465829557) * \text{LYN_X116_T} + (5,53594778946524) * \text{IL17RA_X77_C} + (-3,04072407838489) * \text{LYN_X47_Q} + (1,82611854035731) * \text{IL17RA_X184_H} + (1,14130391422593) * \text{LYN_X85_R} + (-1,36538923404951) * \text{LYN_X63_C} + (-1,52552008654041) * \text{LYN_X172_J} + (3,47820883385469) * \text{LYN_X203_L} + (-1,80085106045942) * \text{IL17RA_X55_B} + (-1,25213023530841) * \text{LYN_X163_Z} + (-1,01733883175651) * \text{PTPRC_X240_I}$$

+ (122,471673425459)* IL17RA_X151_G + (1,09079061375429)* LYN_X64_D + (2,32804478266105)* LYN_X255_N + (-3,86751458009041)* LYN_X167_I + (-12,6254508356666)* LYN_X148_A + (3,00274639389995)* PTPRC_X299_O - (1,0889988893022)* PDE8A_SiteB.

- 5 La lista de todas las combinaciones sometidas a prueba en esta cohorte se presenta en la tabla 4. Cabe destacar que, en la primera cohorte, algunas muestras no se analizaron y la población total de la combinación puede ser inferior a la población total de la cohorte. Todas las combinaciones fueron estadísticamente significativas entre el control y los casos con un valor de $p < 10^{-4}$.

- 10 Tabla 4: Rendimientos diagnósticos obtenidos con pacientes de la primera cohorte de priorización analizados en monoplex.

Núm.	mROC de combinación	Combi.	ID	Curva ROC de AUC	IC 95 %	Umbral	Prec. (%)	Sp (%)	Se (%)	VPP (%)	VPN (%)	Población
1	RASSF1_X263 + AHR_X156	C2	2	0,782	[0,628; 0,936]	-1,04	80,7	82,9	75,0	63,2	89,5	N= 57 [41 (1); 16 (2)]
2	AHR_X118 + AHR_X156	C2	4	0,745	[0,599; 0,892]	0,12	80,7	92,7	50,0	72,7	82,6	N= 57 [41 (1); 16 (2)]
3	AHR_X156 + PDE8A_X116246	C2	6	0,773	[0,633; 0,913]	-6,50	77,2	78,1	75,0	57,1	88,9	N= 57 [41 (1); 16 (2)]
4	IFNAR1_X227 + PTPRC_X299	C2	28	0,815	[0,698; 0,932]	3,53	76,7	68,3	94,7	58,1	96,6	N= 60 [41 (1); 19 (2)]
5	IFNAR1_X84 + PTPRC_X299	C2	83	0,810	[0,694; 0,926]	3,95	76,7	68,3	94,7	58,1	96,6	N= 60 [41 (1); 19 (2)]
6	RASSF1_X263 + AHR_X118	C2	1	0,767	[0,626; 0,907]	-1,42	75,4	73,2	81,3	54,2	90,9	N= 57 [41 (1); 16 (2)]
7	IFNAR1_X227 + PDE8A_X116246	C2	36	0,764	[0,655; 0,873]	-3,28	75,0	73,2	77,1	71,1	79,0	N= 76 [41 (1); 35 (2)]
8	IFNAR1_X84 + PDE8A_X116246	C2	55	0,748	[0,638; 0,859]	-2,70	73,7	73,2	74,3	70,3	76,9	N= 76 [41 (1); 35 (2)]
9	IFNAR1_X227 + PTPRC_X240	C2	22	0,813	[0,692; 0,933]	2,05	73,3	65,9	89,5	54,8	93,1	N= 60 [41 (1); 19 (2)]
10	IL17RA_X77 + IFNAR1_X162	C2	1	0,703	[0,584; 0,823]	-0,36	72,4	73,2	71,4	69,4	75,0	N= 76 [41 (1); 35 (2)]
11	LYN_X153 + PTPRC_X299	C2	203	0,793	[0,679; 0,908]	4,08	71,7	61,0	94,7	52,9	96,2	N= 60 [41 (1); 19 (2)]
12	PTPRC_X299 + PDE8A_X116246	C2	435	0,791	[0,679; 0,903]	-1,73	71,7	61,0	94,7	52,9	96,2	N= 60 [41 (1); 19 (2)]
13	LYN_X85 + PDE8A_X116246	C2	46	0,749	[0,639; 0,859]	-5,33	71,1	58,5	85,7	63,8	82,8	N= 76 [41 (1); 35 (2)]
14	IFNAR1_X162 + PDE8A_X116246	C2	25	0,743	[0,632; 0,854]	-2,21	71,1	73,2	68,6	68,6	73,2	N= 76 [41 (1); 35 (2)]
15	IFNAR1_X231 + PDE8A_X116246	C2	85	0,734	[0,622; 0,847]	-2,80	71,1	61,0	82,9	64,4	80,7	N= 76 [41 (1); 35 (2)]
16	IL17RA_X77 + IFNAR1_X227	C2	2	0,713	[0,595; 0,831]	-0,74	71,1	70,7	71,4	67,6	74,4	N= 76 [41 (1); 35 (2)]
17	IFNAR1_X84 + IFNAR1_X231	C2	51	0,702	[0,582; 0,821]	1,68	71,1	75,6	65,7	69,7	72,1	N= 76 [41 (1); 35 (2)]
18	AHR_X118 + PDE8A_X116246	C2	5	0,755	[0,626; 0,883]	-7,44	68,4	61,0	87,5	46,7	92,6	N= 57 [41 (1); 16 (2)]
19	RASSF1_X263 + PDE8A_X116246	C2	3	0,753	[0,621; 0,885]	-6,76	64,9	53,7	93,8	44,1	95,7	N= 57 [41 (1); 16 (2)]

Núm.	mROC de combinación	Combi.	ID	Curva ROC de AUC	IC 95 %	Umbral	Prec. (%)	Sp (%)	Se (%)	VPP (%)	VPN (%)	Población
20	RASSF1_X263 + AHR_X156 + PDE8A_X116246	C3	3	0,793	[0,65; 0,936]	-6,42	84,2	87,8	75,0	70,6	90,0	N= 57 [41 (1); 16 (2)]
21	RASSF1_X263 + AHR_X118 + AHR_X156	C3	1	0,788	[0,638; 0,938]	-1,10	84,2	90,2	68,8	73,3	88,1	N= 57 [41 (1); 16 (2)]
22	IFNAR1_X227 + LYN_X284 + LYN_X63	C3	184	0,861	[0,754; 0,969]	2,16	83,3	87,8	73,7	73,7	87,8	N= 60 [41 (1); 19 (2)]
23	IFNAR1_X227 + LYN_X284 + LYN_X172	C3	185	0,874	[0,787; 0,962]	3,81	80,0	80,5	79,0	65,2	89,2	N= 60 [41 (1); 19 (2)]
24	IFNAR1_X227 + PTPRC_X299 + PDE8A_X116246	C3	406	0,859	[0,763; 0,955]	-0,10	80,0	75,6	89,5	63,0	93,9	N= 60 [41 (1); 19 (2)]
25	IFNAR1_X227 + LYN_X284 + LYN_X64	C3	191	0,852	[0,748; 0,957]	2,13	80,0	80,5	79,0	65,2	89,2	N= 60 [41 (1); 19 (2)]
26	IFNAR1_X84 + LYN_X284 + LYN_X172	C3	914	0,850	[0,754; 0,945]	4,30	78,3	78,1	79,0	62,5	88,9	N= 60 [41 (1); 19 (2)]
27	IFNAR1_X227 + IFNAR1_X111 + PDE8A_X116246	C3	199	0,769	[0,66; 0,878]	-3,55	77,6	73,2	82,9	72,5	83,3	N= 76 [41 (1); 35 (2)]
28	AHR_X118 + AHR_X156 + PDE8A_X116246	C3	4	0,777	[0,649; 0,905]	-6,99	77,2	82,9	62,5	58,8	85,0	N= 57 [41 (1); 16 (2)]
29	IFNAR1_X84 + IFNAR1_X231 + PDE8A_X116246	C3	274	0,771	[0,665; 0,878]	-3,58	76,3	78,1	74,3	74,3	78,1	N= 76 [41 (1); 35 (2)]
30	RASSF1_X263 + AHR_X118 + PDE8A_X116246	C3	2	0,793	[0,664; 0,921]	-7,18	75,4	73,2	81,3	54,2	90,9	N= 57 [41 (1); 16 (2)]
31	IFNAR1_X227 + IL17RA_X55 + PDE8A_X116246	C3	196	0,771	[0,663; 0,88]	-3,21	75,0	75,6	74,3	72,2	77,5	N= 76 [41 (1); 35 (2)]
32	IFNAR1_X162 + IFNAR1_X227 + PDE8A_X116246	C3	89	0,766	[0,657; 0,875]	-3,40	75,0	73,2	77,1	71,1	79,0	N= 76 [41 (1); 35 (2)]
33	IFNAR1_X84 + PTPRC_X299 + PDE8A_X116246	C3	1135	0,851	[0,757; 0,945]	-0,03	73,3	63,4	94,7	54,6	96,3	N= 60 [41 (1); 19 (2)]
34	IFNAR1_X227 + IFNAR1_X162 + LYN_X284 + LYN_X207	C4	831	0,866	[0,761; 0,972]	-0,44	88,3	92,7	79,0	83,3	90,5	N= 60 [41 (1); 19 (2)]
35	IFNAR1_X227 + IFNAR1_X129 + LYN_X284 + LYN_X207	C4	1407	0,864	[0,752; 0,976]	-3,03	88,3	92,7	79,0	83,3	90,5	N= 60 [41 (1); 19 (2)]
36	IFNAR1_X227 + IFNAR1_X115 + LYN_X284 + LYN_X207	C4	155	0,863	[0,756; 0,97]	-0,20	88,3	92,7	79,0	83,3	90,5	N= 60 [41 (1); 19 (2)]
37	IFNAR1_X227 + LYN_X284 + LYN_X207 + PTPRC_X240	C4	2239	0,858	[0,748; 0,967]	-1,45	88,3	92,7	79,0	83,3	90,5	N= 60 [41 (1); 19 (2)]
38	IFNAR1_X227 + LYN_X255 + PTPRC_X299 + PDE8A_X116246	C4	3650	0,860	[0,763; 0,957]	-0,25	83,3	80,5	89,5	68,0	94,3	N= 60 [41 (1); 19 (2)]

Núm.	mROC de combinación	Combi.	ID	Curva ROC de AUC	IC 95 %	Umbral	Prec. (%)	Sp (%)	Se (%)	VPP (%)	VPN (%)	Población
39	IFNAR1_X227 + IFNAR1_X231 + PTPRC_X299 + PDE8A_X116246	C4	1354	0,861	[0,762; 0,961]	-1,14	81,7	75,6	94,7	64,3	96,9	N= 60 [41 (1); 19 (2)]
40	IFNAR1_X227 + LYN_X255 + LYN_X148 + PTPRC_X299	C4	3648	0,845	[0,731; 0,958]	40,28	81,7	75,6	94,7	64,3	96,9	N= 60 [41 (1); 19 (2)]
41	RASSF1_X263 + AHR_X118 + AHR_X156 + PDE8A_X116246	C4	1	0,811	[0,677; 0,945]	-7,18	79,0	78,1	81,3	59,1	91,4	N= 57 [41 (1); 16 (2)]
42	IFNAR1_X84 + IFNAR1_X231 + IFNAR1_X111 + PDE8A_X116246	C4	871	0,775	[0,669; 0,881]	-3,43	77,6	80,5	74,3	76,5	78,6	N= 76 [41 (1); 35 (2)]
43	IFNAR1_X227 + IL17RA_X55 + IFNAR1_X111 + PDE8A_X116246	C4	670	0,774	[0,666; 0,882]	-3,60	76,3	68,3	85,7	69,8	84,9	N= 76 [41 (1); 35 (2)]
44	IFNAR1_X227 + IFNAR1_X84 + IFNAR1_X111 + PDE8A_X116246	C4	587	0,772	[0,666; 0,878]	-3,00	76,3	73,2	80,0	71,8	81,1	N= 76 [41 (1); 35 (2)]
45	IFNAR1_X84 + IL17RA_X184 + IFNAR1_X231 + PDE8A_X116246	C4	849	0,772	[0,665; 0,879]	-4,12	76,3	80,5	71,4	75,8	76,7	N= 76 [41 (1); 35 (2)]
46	IFNAR1_X227 + LYN_X85 + IFNAR1_X231 + PDE8A_X116246	C4	545	0,782	[0,678; 0,886]	-5,41	75,0	73,2	77,1	71,1	79,0	N= 76 [41 (1); 35 (2)]
47	IFNAR1_X84 + LYN_X153 + IFNAR1_X231 + PDE8A_X116246	C4	859	0,776	[0,671; 0,881]	-2,93	75,0	78,1	71,4	73,5	76,2	N= 76 [41 (1); 35 (2)]
48	IL17RA_X77 + IFNAR1_X227 + IFNAR1_X84 + IFNAR1_X231 + PDE8A_X116246	C5	295	0,803	[0,704; 0,903]	-6,86	77,6	75,6	80,0	73,7	81,6	N= 76 [41 (1); 35 (2)]
49	IL17RA_X77 + IFNAR1_X227 + IFNAR1_X115 + IFNAR1_X231 + PDE8A_X116246	C5	323	0,792	[0,689; 0,896]	-6,82	77,6	78,1	77,1	75,0	80,0	N= 76 [41 (1); 35 (2)]
50	IFNAR1_X227 + IFNAR1_X115 + IFNAR1_X231 + IL17RA_X55 + PDE8A_X116246	C5	1463	0,794	[0,691; 0,898]	-4,96	76,3	73,2	80,0	71,8	81,1	N= 76 [41 (1); 35 (2)]
51	IL17RA_X77 + IFNAR1_X227 + LYN_X85 + IFNAR1_X231 + PDE8A_X116246	C5	259	0,790	[0,687; 0,893]	-6,62	76,3	75,6	77,1	73,0	79,5	N= 76 [41 (1); 35 (2)]
52	IFNAR1_X227 + LYN_X85 + IL17RA_X184 +	C5	1304	0,788	[0,685; 0,891]	-5,73	76,3	75,6	77,1	73,0	79,5	N= 76 [41 (1); 35 (2)]

Núm.	mROC de combinación	Combi.	ID	Curva ROC de AUC	IC 95 %	Umbral	Prec. (%)	Sp (%)	Se (%)	VPP (%)	VPN (%)	Población
	IFNAR1_X231 + PDE8A_X116246											
53	IFNAR1_X227 + IFNAR1_X84 + IL17RA_X184 + IFNAR1_X231 + PDE8A_X116246	C5	1388	0,801	[0,702; 0,899]	-6,10	75,0	70,7	80,0	70,0	80,6	N= 76 [41 (1); 35 (2)]
54	IFNAR1_X227 + LYN_X85 + IFNAR1_X115 + IFNAR1_X231 + PDE8A_X116246	C5	1268	0,796	[0,696; 0,896]	-6,07	75,0	73,2	77,1	71,1	79,0	N= 76 [41 (1); 35 (2)]
55	IFNAR1_X227 + IFNAR1_X84 + LYN_X153 + IFNAR1_X231 + PDE8A_X116246	C5	1398	0,798	[0,699; 0,897]	-4,70	73,7	68,3	80,0	68,3	80,0	N= 76 [41 (1); 35 (2)]
56	IFNAR1_X227 + LYN_X85 + IFNAR1_X84 + IFNAR1_X231 + PDE8A_X116246	C5	1240	0,794	[0,694; 0,894]	-6,17	73,7	75,6	71,4	71,4	75,6	N= 76 [41 (1); 35 (2)]
57	IFNAR1_X227 + IFNAR1_X84 + IL17RA_X78 + IFNAR1_X231 + PDE8A_X116246	C5	1373	0,792	[0,691; 0,893]	-5,16	73,7	70,7	77,1	69,2	78,4	N= 76 [41 (1); 35 (2)]
58	IFNAR1_X227 + IFNAR1_X84 + IFNAR1_X231 + IL17RA_X55 + PDE8A_X116246	C5	1407	0,790	[0,688; 0,892]	-5,34	73,7	68,3	80,0	68,3	80,0	N= 76 [41 (1); 35 (2)]
59	IL17RA_X77 + LYN_X85 + IFNAR1_X84 + IFNAR1_X231 + PDE8A_X116246	C5	415	0,790	[0,69; 0,89]	-6,12	73,7	73,2	74,3	70,3	76,9	N= 76 [41 (1); 35 (2)]
60	IFNAR1_X227 + IFNAR1_X84 + IFNAR1_X231 + IFNAR1_X111 + PDE8A_X116246	C5	1410	0,790	[0,688; 0,891]	-5,30	72,4	61,0	85,7	65,2	83,3	N= 76 [41 (1); 35 (2)]
61	IFNAR1_X227 + IFNAR1_X84 + IFNAR1_X115 + IFNAR1_X231 + PDE8A_X116246	C5	1352	0,789	[0,688; 0,89]	-5,43	72,4	61,0	85,7	65,2	83,3	N= 76 [41 (1); 35 (2)]
62	IL17RA_X77 + IFNAR1_X227 + LYN_X85 + IFNAR1_X115 + IFNAR1_X231 + PDE8A_X116246	C6	553	0,803	[0,702; 0,903]	-7,26	80,3	92,7	65,7	88,5	76,0	N= 76 [41 (1); 35 (2)]
63	IL17RA_X77 + IFNAR1_X227 + LYN_X85 + IL17RA_X78 + IFNAR1_X231 + PDE8A_X116246	C6	574	0,794	[0,692; 0,897]	-7,59	79,0	80,5	77,1	77,1	80,5	N= 76 [41 (1); 35 (2)]
64	IL17RA_X77 + IFNAR1_X227 +	C6	714	0,797	[0,694; 0,899]	-7,06	77,6	78,1	77,1	75,0	80,0	N= 76 [41 (1); 35 (2)]

Núm.	mROC de combinación	Combi.	ID	Curva ROC de AUC	IC 95 %	Umbral	Prec. (%)	Sp (%)	Se (%)	VPP (%)	VPN (%)	Población
	IFNAR1_X115 + IL17RA_X78 + IFNAR1_X231 + PDE8A_X116246											
65	IL17RA_X77 + IFNAR1_X162 + IFNAR1_X227 + IFNAR1_X84 + IFNAR1_X231 + PDE8A_X116246	C6	75	0,805	[0,707; 0,903]	-6,63	76,3	73,2	80,0	71,8	81,1	N= 76 [41 (1); 35 (2)]
66	IL17RA_X77 + IFNAR1_X227 + IFNAR1_X115 + IFNAR1_X231 + IFNAR1_X111 + PDE8A_X116246	C6	751	0,803	[0,702; 0,904]	-6,68	76,3	75,6	77,1	73,0	79,5	N= 76 [41 (1); 35 (2)]
67	IL17RA_X77 + IFNAR1_X227 + IFNAR1_X84 + IFNAR1_X231 + IFNAR1_X111 + PDE8A_X116246	C6	695	0,802	[0,703; 0,902]	-6,61	76,3	78,1	74,3	74,3	78,1	N= 76 [41 (1); 35 (2)]
68	IL17RA_X77 + IFNAR1_X227 + IFNAR1_X84 + IL17RA_X78 + IFNAR1_X231 + PDE8A_X116246	C6	658	0,801	[0,702; 0,901]	-7,18	76,3	75,6	77,1	73,0	79,5	N= 76 [41 (1); 35 (2)]
69	IFNAR1_X227 + IFNAR1_X84 + IL17RA_X78 + IL17RA_X184 + IFNAR1_X231 + PDE8A_X116246	C6	2354	0,798	[0,698; 0,897]	-6,44	76,3	70,7	82,9	70,7	82,9	N= 76 [41 (1); 35 (2)]
70	IFNAR1_X227 + LYN_X85 + IFNAR1_X84 + IL17RA_X184 + IFNAR1_X231 + PDE8A_X116246	C6	2137	0,808	[0,711; 0,904]	-6,43	75,0	82,9	65,7	76,7	73,9	N= 76 [41 (1); 35 (2)]
71	IFNAR1_X227 + LYN_X85 + IFNAR1_X84 + IFNAR1_X231 + IFNAR1_X111 + PDE8A_X116246	C6	2159	0,803	[0,706; 0,901]	-5,94	75,0	82,9	65,7	76,7	73,9	N= 76 [41 (1); 35 (2)]
72	IL17RA_X77 + IFNAR1_X227 + IFNAR1_X84 + IFNAR1_X231 + LYN_X116_T + PDE8A_X116246	C6	694	0,803	[0,705; 0,901]	-6,96	75,0	80,5	68,6	75,0	75,0	N= 76 [41 (1); 35 (2)]
73	IFNAR1_X227 + LYN_X85 + IFNAR1_X115 + IFNAR1_X231 + IL17RA_X55 + PDE8A_X116246	C6	2212	0,802	[0,702; 0,902]	-5,94	75,0	78,1	71,4	73,5	76,2	N= 76 [41 (1); 35 (2)]

Núm.	mROC de combinación	Combi.	ID	Curva ROC de AUC	IC 95 %	Umbral	Prec. (%)	Sp (%)	Se (%)	VPP (%)	VPN (%)	Población
74	IFNAR1_X227 + IFNAR1_X84 + IFNAR1_X115 + IL17RA_X184 + IFNAR1_X231 + PDE8A_X116246	C6	2319	0,802	[0,704; 0,901]	-6,11	75,0	70,7	80,0	70,0	80,6	N= 76 [41 (1); 35 (2)]
75	IFNAR1_X227 + LYN_X85 + IFNAR1_X84 + LYN_X153 + IFNAR1_X231 + PDE8A_X116246	C6	2147	0,801	[0,704; 0,899]	-5,36	75,0	87,8	60,0	80,8	72,0	N= 76 [41 (1); 35 (2)]
76	IL17RA_X77 + IFNAR1_X227 + IFNAR1_X115 + IFNAR1_X231 + IL17RA_X55 + PDE8A_X116246	C6	748	0,801	[0,701; 0,9]	-8,23	75,0	70,7	80,0	70,0	80,6	N= 76 [41 (1); 35 (2)]
77	IFNAR1_X227 + IFNAR1_X84 + IL17RA_X184 + IFNAR1_X231 + IL17RA_X55 + PDE8A_X116246	C6	2393	0,801	[0,702; 0,899]	-6,14	75,0	70,7	80,0	70,0	80,6	N= 76 [41 (1); 35 (2)]
78	IFNAR1_X227 + IFNAR1_X84 + LYN_X153 + IFNAR1_X231 + IL17RA_X55 + PDE8A_X116246	C6	2403	0,799	[0,7; 0,897]	-4,86	73,7	58,5	91,4	65,3	88,9	N= 76 [41 (1); 35 (2)]
79	IFNAR1_X227 + IFNAR1_X84 + IL17RA_X78 + LYN_X153 + IFNAR1_X231 + PDE8A_X116246	C6	2364	0,802	[0,704; 0,9]	-4,86	72,4	56,1	91,4	64,0	88,5	N= 76 [41 (1); 35 (2)]
80	IFNAR1_X227 + IFNAR1_X115 + IFNAR1_X84 + IFNAR1_X162 + IFNAR1_X231 + IFNAR1_X129 + IFNAR1_X111 + LYN_X153 + LYN_X284 + IFNAR1_X152 + IL17RA_X78 + LYN_X116_T + IL17RA_X77 + LYN_X47 + IL17RA_X184 + LYN_X85 + LYN_X63 + LYN_X172 + LYN_X203 + IL17RA_X55 + LYN_X163 + PTPRC_X240 + IL17RA_X151 + LYN_X64 + LYN_X255 + LYN_X167 + LYN_X148 + PTPRC_X299 + PDE8A_X116246	C29	14	0,977	[0,948; 1]	68,06	93,3	92,7	94,7	85,7	97,4	N= 60 [41 (1); 19 (2)]

En la segunda cohorte, usando secuenciación multiplexada, cuando sólo se eligieron 2 sitios, cada uno para un BMK, el AUC fue 0,786 (Sp 78,6 %; Se 81,0 %, figura 5A). En esta cohorte, el aumento del número de biomarcadores

aumentó sistemáticamente el rendimiento diagnóstico global del ensayo. Cuando se combinaron 3 sitios en 3 BMK, el AUC aumentó a 0,839 (Sp 89,3 %; Se 74,1 %, figura 5B). Con 4 y 5 sitios, el AUC aumentó a 0,852 y 0,860, con especificidades de 82,1 % y 92,9 % y sensibilidades de 81 % y 77,6 %, respectivamente (figura 5C y D). Cuando se aumentó la combinación para incluir 19 sitios, el AUC fue superior a 0,9 (AUC 0,909, Sp 92,9 %, Se 86,2 %, figura 5E). Además, cuando se combinaron 48 sitios editados de manera diferencial presentes en 9 BMK diferentes, se obtuvo un AUC excepcional de 0,993, con una especificidad del 100 % y una sensibilidad del 96,6 % (figura 5F). Para esta combinación, la ecuación usada es la siguiente:

$$Z = -(0,179462180023875) * \text{MDM2_Site207_W} - (-1,04404354492012) * \text{IFNAR1_Site101_nan} - (2,64479093229565) * \text{GAB2_Site88} - (4,11125599566128) * \text{v1MDM2_Site76} - (1,61229526347695) * \text{MDM2_Site143} - (5,57184346092994) * \text{LYN_Site90_S} - (5,25866173918605) * \text{PDE8A_Site116248_B} - (18,995426153835) * \text{KCNJ15_Site134_D} - (-1,04699821617189) * \text{v1MDM2_Site86} - (0,480007114248953) * \text{CAMK1D_Site131} - (3,56662634744781) * \text{KCNJ15_Site130_A} - (2,4602823172239) * \text{KCNJ15_Site129_B} - (-9,43134276848121) * \text{KCNJ15_Site133_C} - (1,42715048538683) * \text{IFNAR1_Site76_nan} - (1,00606326866477) * \text{CAMK1D_Site49} - (0,0542132983688757) * \text{v1MDM2_Site87_E} + (5,91151340839856) * \text{LYN_Site155_X} - (-2,20437543623734) * \text{KCNJ15_Site131} - (1,30800292033623) * \text{PTPRC_Site169_E} - (4,82052635778379) * \text{MDM2_Site277} - (-10,1173014238565) * \text{KCNJ15_Site190} - (-3,7454636873821) * \text{PTPRC_Site170_F} - (1,03362539706716) * \text{PTPRC_Site259} - (-4,87935759821363) * \text{PTPRC_Site163_D} - (0,628308423459796) * \text{PRKCB_Site119_E} - (-0,901688238257015) * \text{KCNJ15_Site193_E} - (-0,425246156880022) * \text{MDM2_Site264} - (-0,533167398964742) * \text{IFNAR1_Site123_nan} - (-5,64307592868436) * \text{IFNAR1_Site131_nan} - (47,981969789844) * \text{GAB2_Site131_H} - (0,0402950393688472) * \text{MDM2_Site237_X} - (-3,33278344933062) * \text{MDM2_Site192} - (0,928449146700883) * \text{CAMK1D_Site64_D} - (-1,08069775385396) * \text{GAB2_Site126_A} - (-4,23184777312141) * \text{CAMK1D_Site113_G} - (1,99562009099246) * \text{GAB2_Site52_B} - (2,3505776086063) * \text{IFNAR1_Site142_nan} - (3,97251790792544) * \text{PRKCB_Site103_C} - (0,570977372558063) * \text{PRKCB_Site126} - (-0,410508492870566) * \text{PRKCB_Site132_H} - (0,975605316457069) * \text{IFNAR1_Site128_nan} - (-1,09288061572838) * \text{PDE8A_Site116278_C} + (12,6078350770629) * \text{CAMK1D_Site125} - (1,48817555399057) * \text{PTPRC_Site253} - (-0,716028094006562) * \text{PRKCB_Site99} - (-0,722585574440859) * \text{IFNAR1_Site151_nan} - (0,00161227142613181) * \text{IFNAR1_Site152_C} - (-7,23169188321118) * \text{PRKCB_Site183}.$$

La lista de todas las combinaciones sometidas a prueba en la segunda cohorte de priorización, la totalidad de las cuales incluyeron los 86 pacientes, se presenta en la tabla 5. Todas las combinaciones fueron estadísticamente significativas entre el control y los casos con un valor de $p < 10^{-4}$.

Tabla 5: Rendimientos diagnósticos obtenidos con pacientes de la segunda cohorte de priorización analizados con secuenciación multiplexada

Núm.	mROC de combinación	Combi.	ID	Curva ROC de AUC	IC 95 %	Umbral	Prec. (%)	Sp (%)	Se (%)	VPP (%)	VPN (%)	Población
1	GAB2_X88 + KCNJ15_X133	C2	44	0,786	[0,667; 0,905]	-3,1179	80,2	78,6	81,0	88,7	66,7	N= 86 [28 (1); 58 (2)]
2	GAB2_X88 + KCNJ15_X134	C2	39	0,800	[0,685; 0,915]	-2,4900	79,1	78,6	79,3	88,5	64,7	N= 86 [28 (1); 58 (2)]
3	LYN_X90 + KCNJ15_X134	C2	73	0,786	[0,676; 0,896]	-1,9946	79,1	60,7	87,9	82,3	70,8	N= 86 [28 (1); 58 (2)]
4	GAB2_X88 + KCNJ15_X130	C2	42	0,772	[0,651; 0,892]	-3,7777	79,1	71,4	82,8	85,7	66,7	N= 86 [28 (1); 58 (2)]
5	GAB2_X88 + KCNJ15_X193	C2	51	0,778	[0,662; 0,895]	-2,5714	77,9	78,6	77,6	88,2	62,9	N= 86 [28 (1); 58 (2)]
6	GAB2_X88 + KCNJ15_X129	C2	43	0,782	[0,663; 0,901]	-2,0607	76,7	75,0	77,6	86,5	61,8	N= 86 [28 (1); 58 (2)]
7	GAB2_X88 + KCNJ15_X190	C2	48	0,774	[0,656; 0,892]	-1,3815	75,6	75,0	75,9	86,3	60,0	N= 86 [28 (1); 58 (2)]
8	MDM2_X207 + CAMK1D_X125	C2	31	0,779	[0,676; 0,881]	-1,2096	74,4	78,6	72,4	87,5	57,9	N= 86 [28 (1); 58 (2)]
9	LYN_X90 + KCNJ15_X129	C2	77	0,770	[0,655; 0,884]	-1,3684	74,4	71,4	75,9	84,6	58,8	N= 86 [28 (1); 58 (2)]
10	MDM2_X207 + LYN_X90	C2	2	0,812	[0,712; 0,911]	-2,5068	73,3	85,7	67,2	90,7	55,8	N= 86 [28 (1); 58 (2)]

Núm.	mROC de combinación	Combi.	ID	Curva ROC de AUC	IC 95 %	Umbral	Prec. (%)	Sp (%)	Se (%)	VPP (%)	VPN (%)	Población
11	LYN_X90 + CAMK1D_X131	C2	74	0,772	[0,664; 0,879]	-0,8314	73,3	78,6	70,7	87,2	56,4	N= 86 [28 (1); 58 (2)]
12	MDM2_X207 + GAB2_X88	C2	1	0,785	[0,685; 0,885]	-2,7339	72,1	82,1	67,2	88,6	54,8	N= 86 [28 (1); 58 (2)]
13	MDM2_X207 + IFNAR1_X142	C2	26	0,782	[0,683; 0,881]	-2,2274	72,1	89,3	63,8	92,5	54,4	N= 86 [28 (1); 58 (2)]
14	LYN_X90 + PDE8A_X116248	C2	72	0,772	[0,665; 0,878]	-7,3860	70,9	82,1	65,5	88,4	53,5	N= 86 [28 (1); 58 (2)]
15	GAB2_X88 + LYN_X90 + KCNJ15_X130	C3	635	0,823	[0,714; 0,933]	-3,9804	83,7	67,9	91,4	85,5	79,2	N= 86 [28 (1); 58 (2)]
16	GAB2_X88 + LYN_X90 + KCNJ15_X190	C3	641	0,816	[0,708; 0,924]	-1,7582	82,6	71,4	87,9	86,4	74,1	N= 86 [28 (1); 58 (2)]
17	GAB2_X88 + KCNJ15_X134 + CAMK1D_X125	C3	724	0,817	[0,709; 0,924]	-1,7720	82,6	71,4	87,9	86,4	74,1	N= 86 [28 (1); 58 (2)]
18	GAB2_X88 + KCNJ15_X134 + MDM2_X191	C3	699	0,804	[0,69; 0,919]	-2,4499	82,6	75,0	86,2	87,7	72,4	N= 86 [28 (1); 58 (2)]
19	GAB2_X88 + KCNJ15_X134 + KCNJ15_X193	C3	709	0,803	[0,688; 0,918]	-1,5841	81,4	85,7	79,3	92,0	66,7	N= 86 [28 (1); 58 (2)]
20	GAB2_X88 + KCNJ15_X134 + KCNJ15_X190	C3	706	0,808	[0,692; 0,924]	-3,0436	81,4	82,1	81,0	90,4	67,7	N= 86 [28 (1); 58 (2)]
21	GAB2_X88 + KCNJ15_X133 + IFNAR1_X131	C3	856	0,789	[0,672; 0,906]	-3,0585	81,4	78,6	82,8	88,9	68,8	N= 86 [28 (1); 58 (2)]
22	GAB2_X88 + KCNJ15_X133 + GAB2_X52	C3	863	0,789	[0,67; 0,909]	-2,3545	81,4	78,6	82,8	88,9	68,8	N= 86 [28 (1); 58 (2)]
23	GAB2_X88 + KCNJ15_X133 + GAB2_X126	C3	861	0,787	[0,668; 0,906]	-3,0734	81,4	78,6	82,8	88,9	68,8	N= 86 [28 (1); 58 (2)]
24	GAB2_X88 + KCNJ15_X134 + MDM2_X237	C3	713	0,808	[0,696; 0,921]	-3,3208	80,2	82,1	79,3	90,2	65,7	N= 86 [28 (1); 58 (2)]
25	GAB2_X88 + KCNJ15_X134 + CAMK1D_X64	C3	715	0,797	[0,681; 0,912]	-2,8067	80,2	82,1	79,3	90,2	65,7	N= 86 [28 (1); 58 (2)]
26	GAB2_X88 + LYN_X90 + KCNJ15_X133	C3	637	0,829	[0,72; 0,938]	-3,0116	79,1	89,3	74,1	93,5	62,5	N= 86 [28 (1); 58 (2)]
27	GAB2_X88 + LYN_X90 + KCNJ15_X134	C3	632	0,839	[0,733; 0,945]	-2,5595	79,1	89,3	74,1	93,5	62,5	N= 86 [28 (1); 58 (2)]
28	GAB2_X88 + LYN_X90 + KCNJ15_X134 + MDM2_X191	C4	7175	0,840	[0,734; 0,945]	-2,6812	86,1	78,6	89,7	89,7	78,6	N= 86 [28 (1); 58 (2)]
29	MDM2_X207 + GAB2_X88 + LYN_X90 + MDM2_X192	C4	19	0,841	[0,742; 0,939]	-2,9468	84,9	78,6	87,9	89,5	75,9	N= 86 [28 (1); 58 (2)]

Núm.	mROC de combinación	Combi.	ID	Curva ROC de AUC	IC 95 %	Umbral	Prec. (%)	Sp (%)	Se (%)	VPP (%)	VPN (%)	Población
30	MDM2_X207 + GAB2_X88 + LYN_X90 + CAMK1D_X131	C4	3	0,854	[0,76; 0,948]	-3,3309	83,7	78,6	86,2	89,3	73,3	N= 86 [28 (1); 58 (2)]
31	GAB2_X88 + KCNJ15_X134 + MDM2_X237 + GAB2_X126	C4	8592	0,820	[0,709; 0,932]	-2,6335	82,6	82,1	82,8	90,6	69,7	N= 86 [28 (1); 58 (2)]
32	GAB2_X88 + KCNJ15_X134 + KCNJ15_X129 + KCNJ15_X190	C4	8324	0,813	[0,696; 0,93]	-3,7791	82,6	85,7	81,0	92,2	68,6	N= 86 [28 (1); 58 (2)]
33	MDM2_X207 + KCNJ15_X134 + IFNAR1_X142 + IFNAR1_X128	C4	2129	0,816	[0,718; 0,914]	-3,0729	81,4	89,3	77,6	93,8	65,8	N= 86 [28 (1); 58 (2)]
34	GAB2_X88 + LYN_X90 + KCNJ15_X134 + CAMK1D_X125	C4	7200	0,852	[0,752; 0,953]	-1,6869	81,4	82,1	81,0	90,4	67,7	N= 86 [28 (1); 58 (2)]
35	MDM2_X207 + GAB2_X88 + LYN_X90 + KCNJ15_X134 + IFNAR1_X142	C5	27	0,850	[0,751; 0,949]	-3,8492	88,4	82,1	91,4	91,4	82,1	N= 86 [28 (1); 58 (2)]
36	GAB2_X88 + LYN_X90 + KCNJ15_X134 + CAMK1D_X131 + PTPRC_X170	C5	4005	0,854	[0,753; 0,955]	-3,2924	87,2	71,4	94,8	87,3	87,0	N= 86 [28 (1); 58 (2)]
37	GAB2_X88 + LYN_X90 + KCNJ15_X134 + PRKCB_X103 + PRKCB_X132	C5	4097	0,853	[0,751; 0,955]	-2,7727	86,1	78,6	89,7	89,7	78,6	N= 86 [28 (1); 58 (2)]
38	GAB2_X88 + LYN_X90 + MDM2_X191 + KCNJ15_X129 + PRKCB_X132	C5	4214	0,828	[0,722; 0,935]	-2,1014	86,1	75,0	91,4	88,3	80,8	N= 86 [28 (1); 58 (2)]
39	GAB2_X88 + LYN_X90 + CAMK1D_X131 + KCNJ15_X130 + CAMK1D_X49	C5	4118	0,845	[0,744; 0,945]	-3,6631	84,9	82,1	86,2	90,9	74,2	N= 86 [28 (1); 58 (2)]
40	GAB2_X88 + LYN_X90 + KCNJ15_X134 + MDM2_X191 + KCNJ15_X190	C5	4017	0,841	[0,736; 0,946]	-3,3395	84,9	82,1	86,2	90,9	74,2	N= 86 [28 (1); 58 (2)]
41	GAB2_X88 + LYN_X90 + PDE8A_X116248 + CAMK1D_X131 + KCNJ15_X129	C5	3894	0,840	[0,74; 0,94]	-5,7536	84,9	82,1	86,2	90,9	74,2	N= 86 [28 (1); 58 (2)]
42	MDM2_X207 + LYN_X90 + KCNJ15_X134 + CAMK1D_X131 + KCNJ15_X133	C5	940	0,860	[0,77; 0,951]	-0,7551	82,6	92,9	77,6	95,7	66,7	N= 86 [28 (1); 58 (2)]
43	MDM2_X207 + GAB2_X88 + LYN_X90 + KCNJ15_X134 +	C6	185	0,849	[0,75; 0,947]	-3,6239	88,4	85,7	89,7	92,9	80,0	N= 86 [28 (1); 58 (2)]

Núm.	mROC de combinación	Combi.	ID	Curva ROC de AUC	IC 95 %	Umbral	Prec. (%)	Sp (%)	Se (%)	VPP (%)	VPN (%)	Población
	IFNAR1_X76 + IFNAR1_X142											
44	GAB2_X88 + LYN_X90 + KCNJ15_X134 + MDM2_X191 + PRKCB_X103 + PRKCB_X132	C6	12353	0,858	[0,756; 0,959]	-2,4262	88,4	78,6	93,1	90,0	84,6	N= 86 [28 (1); 58 (2)]
45	MDM2_X207 + GAB2_X88 + LYN_X90 + KCNJ15_X134 + IFNAR1_X142 + IFNAR1_X128	C6	219	0,853	[0,756; 0,95]	-3,9043	87,2	85,7	87,9	92,7	77,4	N= 86 [28 (1); 58 (2)]
46	GAB2_X88 + LYN_X90 + CAMK1D_X131 + KCNJ15_X129 + CAMK1D_X49 + PTPRC_X170	C6	12809	0,855	[0,755; 0,954]	-3,3713	87,2	82,1	89,7	91,2	79,3	N= 86 [28 (1); 58 (2)]
47	GAB2_X88 + LYN_X90 + KCNJ15_X134 + IFNAR1_X142 + PRKCB_X103 + PRKCB_X132	C6	12635	0,854	[0,754; 0,954]	-2,5817	87,2	82,1	89,7	91,2	79,3	N= 86 [28 (1); 58 (2)]
48	GAB2_X88 + LYN_X90 + KCNJ15_X133 + PRKCB_X103 + PRKCB_X126 + PRKCB_X132	C6	13795	0,840	[0,734; 0,946]	-2,9446	87,2	82,1	89,7	91,2	79,3	N= 86 [28 (1); 58 (2)]
49	MDM2_X207 + GAB2_X88 + LYN_X90 + KCNJ15_X134 + CAMK1D_X131 + KCNJ15_X190	C6	128	0,877	[0,786; 0,968]	-5,1449	86,1	85,7	86,2	92,6	75,0	N= 86 [28 (1); 58 (2)]
50	MDM2_X207 + GAB2_X88 + LYN_X90 + PDE8A_X116248 + CAMK1D_X49 + IFNAR1_X142	C6	88	0,861	[0,773; 0,949]	-7,4108	86,1	85,7	86,2	92,6	75,0	N= 86 [28 (1); 58 (2)]
51	MDM2_X207 + GAB2_X88 + LYN_X90 + PDE8A_X116248 + KCNJ15_X134 + IFNAR1_X142	C6	11	0,858	[0,766; 0,95]	-7,0457	86,1	85,7	86,2	92,6	75,0	N= 86 [28 (1); 58 (2)]
52	MDM2_X207 + GAB2_X88 + LYN_X90 + KCNJ15_X134 + KCNJ15_X190 + IFNAR1_X142	C6	206	0,852	[0,755; 0,948]	-5,4371	86,1	85,7	86,2	92,6	75,0	N= 86 [28 (1); 58 (2)]
53	MDM2_X207 + KCNJ15_X134 + CAMK1D_X131 + MDM2_X191 + IFNAR1_X142 + IFNAR1_X128	C6	7332	0,819	[0,724; 0,914]	-3,3357	83,7	89,3	81,0	94,0	69,4	N= 86 [28 (1); 58 (2)]
54	MDM2_X207 + LYN_X90 +	C6	3658	0,861	[0,77; 0,951]	-0,8548	82,6	92,9	77,6	95,7	66,7	N= 86 [28 (1); 58 (2)]

Núm.	mROC de combinación	Combi.	ID	Curva ROC de AUC	IC 95 %	Umbral	Prec. (%)	Sp (%)	Se (%)	VPP (%)	VPN (%)	Población
	KCNJ15_X134 + CAMK1D_X131 + KCNJ15_X133 + CAMK1D_X49											
55	MDM2_X207 + GAB2_X88 + LYN_X90 + PDE8A_X116248 + KCNJ15_X134 + CAMK1D_X131 + MDM2_X191 + KCNJ15_X130 + KCNJ15_X129 + KCNJ15_X133 + IFNAR1_X76 + CAMK1D_X49 + MDM2_X87 + KCNJ15_X190 + IFNAR1_X142 +	C19	6	0,909	[0,833; 0,985]	-6,0428	88,4	92,9	86,2	96,2	76,5	N= 86 [28 (1); 58 (2)]
	PRKCB_X103 + PRKCB_X126 + PRKCB_X132 + IFNAR1_X128											
56	MDM2_X207 + GAB2_X88 + LYN_X90 + PDE8A_X116248 + KCNJ15_X134 + CAMK1D_X131 + MDM2_X191 + KCNJ15_X130 + KCNJ15_X129 + KCNJ15_X133 + IFNAR1_X76 + CAMK1D_X49 + MDM2_X87 + KCNJ15_X190 + PTPRC_X170 + PRKCB_X103 + PRKCB_X126 + PRKCB_X132 + IFNAR1_X128	C19	5	0,908	[0,829; 0,986]	-5,2338	88,4	89,3	87,9	94,4	78,1	N= 86 [28 (1); 58 (2)]
57	MDM2_X207 + GAB2_X88 + LYN_X90 + PDE8A_X116248 + KCNJ15_X134 + CAMK1D_X131 + MDM2_X191 + KCNJ15_X130 + KCNJ15_X129 + KCNJ15_X133 + IFNAR1_X76 + CAMK1D_X49 + MDM2_X87 + KCNJ15_X190 + PTPRC_X170 + PRKCB_X119 + KCNJ15_X193 + IFNAR1_X123 + IFNAR1_X131 + GAB2_X131 +	C36	11	0,969	[0,938; 0,999]	36,3081	94,2	92,9	94,8	96,5	89,7	N= 86 [28 (1); 58 (2)]

Núm.	mROC de combinación	Combi.	ID	Curva ROC de AUC	IC 95 %	Umbral	Prec. (%)	Sp (%)	Se (%)	VPP (%)	VPN (%)	Población
	MDM2_X237 + MDM2_X192 + CAMK1D_X64 + GAB2_X126 + CAMK1D_X113 + GAB2_X52 + PRKCB_X103 + PRKCB_X126 + PRKCB_X132 + IFNAR1_X128 + CAMK1D_X125 + MDM2_X82 + PTPRC_X253 + PRKCB_X99 + IFNAR1_X151 + IFNAR1_X152											
58	MDM2_X207 + GAB2_X88 + LYN_X90 + KCNJ15_X134 + CAMK1D_X131 + MDM2_X191 + KCNJ15_X130 + KCNJ15_X129 + KCNJ15_X133 + IFNAR1_X76 + CAMK1D_X49 + MDM2_X87 + KCNJ15_X190 + PTPRC_X170 + PRKCB_X119 + KCNJ15_X193 + IFNAR1_X123 + IFNAR1_X131 + GAB2_X131 + MDM2_X237 + MDM2_X192 + CAMK1D_X64 + GAB2_X126 + CAMK1D_X113 + GAB2_X52 + IFNAR1_X142 + PRKCB_X103 + PRKCB_X126 + PRKCB_X132 + IFNAR1_X128 + CAMK1D_X125 + MDM2_X82 + PTPRC_X253 + PRKCB_X99 + IFNAR1_X151 + IFNAR1_X152	C36	34	0,963	[0,925; 1]	46,2509	94,2	96,4	93,1	98,2	87,1	N= 86 [28 (1); 58 (2)]
59	MDM2_X207 + GAB2_X88 + LYN_X90 + PDE8A_X116248 + KCNJ15_X134 + CAMK1D_X131 + MDM2_X191 + KCNJ15_X130 + KCNJ15_X129 + KCNJ15_X133 +	C36	7	0,967	[0,935; 0,999]	15,6122	89,5	100,0	84,5	100,0	75,7	N= 86 [28 (1); 58 (2)]

Núm.	mROC de combinación	Combi.	ID	Curva ROC de AUC	IC 95 %	Umbral	Prec. (%)	Sp (%)	Se (%)	VPP (%)	VPN (%)	Población
	IFNAR1_X76 + CAMK1D_X49 + MDM2_X87 + KCNJ15_X190 + PTPRC_X170 + PRKCB_X119 + KCNJ15_X193 + IFNAR1_X123 + IFNAR1_X131 + GAB2_X131 + MDM2_X237 + MDM2_X192 + CAMK1D_X64 + GAB2_X126 + CAMK1D_X113 + GAB2_X52 + IFNAR1_X142 + PRKCB_X103 + PRKCB_X126 + PRKCB_X132 + CAMK1D_X125 + MDM2_X82 + PTPRC_X253 + PRKCB_X99 + IFNAR1_X151 + IFNAR1_X152											
60	MDM2_X207 + IFNAR1_X101 + GAB2_X88 + MDM2_X76 + MDM2_X143 + LYN_X90 + PDE8A_X116248 + KCNJ15_X134 + MDM2_X86 + CAMK1D_X131 + KCNJ15_X130 + KCNJ15_X129 + KCNJ15_X133 + IFNAR1_X76 + CAMK1D_X49 + MDM2_X87 + LYN_X155_X + KCNJ15_X131 + PTPRC_X169 + MDM2_X277 + KCNJ15_X190 + PTPRC_X170 + PTPRC_X259 + PTPRC_X163 + PRKCB_X119 + KCNJ15_X193 + MDM2_X264 + IFNAR1_X123 + IFNAR1_X131 + GAB2_X131 + MDM2_X237 + MDM2_X192 + CAMK1D_X64 + GAB2_X126 +	C48	747	0,993	[0,981; 1]	- 84,5770	97,7	100,0	96,6	100,0	93,3	N= 86 [28 (1); 58 (2)]

Núm.	mROC de combinación	Combi.	ID	Curva ROC de AUC	IC 95 %	Umbral	Prec. (%)	Sp (%)	Se (%)	VPP (%)	VPN (%)	Población
	CAMK1D_X113 + GAB2_X52 + IFNAR1_X142 + PRKCB_X103 + PRKCB_X126 + PRKCB_X132 + IFNAR1_X128 + PDE8A_X116278 + CAMK1D_X125 + PTPRC_X253 + PRKCB_X99 + IFNAR1_X151 + IFNAR1_X152 + PRKCB_X183											

En conjunto, estos datos sugieren que los diferentes biomarcadores del panel de 14 genes seleccionados en las cohortes de priorización pueden usarse para el diagnóstico secuencial de DEP. Es importante destacar que todos los biomarcadores identificados incluidos en este panel de 14 genes se incluyeron en al menos una combinación, lo que sugiere que todos esos BMK pueden usarse para diagnosticar la DEP con un rendimiento diagnóstico adecuado.

Se seleccionó un panel de los 8 biomarcadores más interesantes y se secuenció en un formato múltiplex para validar adicionalmente los presentes resultados en una cohorte grande (cohorte de validación) de 411 sujetos, de los cuales 143 eran controles sanos y 268 eran pacientes con DEP (figura 19).

Todos los BMK se combinaron entre sí para evaluar el aumento potencial en la sensibilidad y especificidad usando bosque aleatorio (RF), un enfoque de aprendizaje automático. El análisis de las variantes de edición de ARN de 8 BMK (GAB2, IFNAR1, KCNJ15, LYN, MDM2, PRKCB, CAMK1D y PDE8A) dio una curva ROC de AUC de 0,929 (IC 95 %: [0,876 - 0,981]), con una sensibilidad del 82,9 % y una especificidad del 87,1 % (figura 22 A) sobre el conjunto de datos de prueba, para el modelo 1 ajustado por edad, sexo, tratamiento psiquiátrico y adicción.

Tabla 5.1: Tabla 5.1: modelo de bosque aleatorio ajustado por edad, sexo, tratamiento psiquiátrico y adicciones (modelo 1, véase la figura 22A)

Lista de dianas en el modelo 1 de bosque aleatorio	Lista de variantes de edición de ARN I-A en el modelo 1 de bosque aleatorio
IFNAR1, GAB2, KCNJ15, MDM2, LYN, PRKCB, CAMK1D,	CAMK1D_E, CAMK1D_FS, IFNAR1.1_Site33355798, IFNAR1.1_Site33355809, IFNAR1.1_Site33355817, IFNAR1.1_Site33355838_P, IFNAR1.1_Site33355840_E, GAB2_AE, GAB2_BCDE, GAB2_CE, GAB2_D, GAB2_EJ, GAB2_Site78331030, GAB2_Site78331054, KCNJ15_C, KCNJ15_CE, KCNJ15_Site38287584, LYN_AC, LYN_Site56012495, MDM2_AIKTWV, MDM2_I, MDM2_SVW, MDM2_SW, PDE8A_MDM2_SWX, MDM2_Site68821573, MDM2_Site68821584, MDM2_Site68821586_W, MDM2_Site68821602, MDM2_Site68821616_X, MDM2_Site68821643, MDM2_W, PDE8A_B, PDE8A_Site85096687_B, PRKCB_B, PRKCB_DHIJKM, PRKCB_DIJ, PRKCB_DIJK, PRKCB_IJK, PRKCB_NoEdit, PRKCB_Site23920893_G, PRKCB_Site23920902_I, PRKCB_Site23920904, PRKCB_Site23920910_J, PRKCB_Site23920911, PRKCB_Site23920919_K

Además, el modelo 2 de bosque aleatorio ajustado sólo por edad y sexo usando los mismos 8 BMK (AMK1D, IFNAR1, GAB2, MDM2, KCNJ15, PRKCB, LYN y PDE8A) dio una curva ROC de AUC de 0,911 (IC 95 %: [0,854 - 0,968]), con una sensibilidad del 80,0 % y una especificidad del 82,4 % (figura 22 B) sobre el conjunto de datos de prueba.

Ambos modelos permiten una clara separación entre pacientes de control y deprimidos.

Tabla 5.2: modelo de bosque aleatorio ajustado por edad y sexo (modelo 2, véase la figura 22B)

Lista de dianas en el modelo 2 de bosque aleatorio	Lista de variantes de edición de ARN I-A en el modelo 2 de bosque aleatorio
IFNAR1, GAB2, KCNJ15, MDM2, LYN, PRKCB, CAMK1D, PDE8A	CAMK1D_A, CAMK1D_AE, CAMK1D_E, CAMK1D_Site12378440, IFNAR1.1_Site33355789_A, IFNAR1.1_Site33355793_G, IFNAR1.1_Site33355798, IFNAR1.1_Site33355809, IFNAR1.1_Site33355817, IFNAR1.1_Site33355838_P, IFNAR1.1_Site33355839_L, IFNAR1.1_Site33355840_E, GAB2_CE, GAB2_Site78331054, KCNJ15_C, KCNJ15_CE, KCNJ15_Site38287525_B, KCNJ15_Site38287529_C, KCNJ15_Site38287530_D, KCNJ15_Site38287586, KCNJ15_Site38287589_E, KCNJ15_Site38287590, KCNJ15_Site38287591, LYN_A, LYN_AC, MDM2_ASV, MDM2_KSX, MDM2_SVW, MDM2_SW, MDM2_SWX, MDM2_Site68821573, MDM2_Site68821586_W, MDM2_Site68821602, MDM2_Site68821616_X, MDM2_Site68821643, MDM2_T, MDM2_W, MDM2_WX, PDE8A_B, PDE8A_Site85096687_B, PRKCB_B, PRKCB_DIJ, PRKCB_DIJK, PRKCB_DJ, PRKCB_I, PRKCB_IJK, PRKCB_J, PRKCB_JK, PRKCB_Site23920884_D, PRKCB_Site23920893_G, PRKCB_Site23920902_I, PRKCB_Site23920904, PRKCB_Site23920910_J, PRKCB_Site23920911, PRKCB_Site23920919_K, PRKCB_Site23920921_L

5 Cuando se analizaron los sitios de edición de toda la población y se combinaron 17 variantes de edición en 5 BMK, se obtuvo un AUC de 0,751 (Sp 70,6 %; Se 72,4 %, figura 6A). Cuando se usaron los 8 biomarcadores incluidos en el multiplex y se combinaron 53 sitios editados de manera diferencial, el AUC se elevó a 0,824, con una especificidad del 76,2 % y una sensibilidad del 81 % (figura 6B). Cuando se aumentó adicionalmente el número de sitios de edición en esos 8 BMK para alcanzar 79 sitios, todos los cuales se secuenciaron en multiplex en un solo experimento, el AUC aumentó adicionalmente a 0,899 (Sp 82,5 %; Se 81,7 %, figura 6C). En este último caso, la ecuación usada es la siguiente:

$$\begin{aligned}
 Z = & (-0,181885362537332) * \text{MDM2_Site68821586_W} - (0,45875321689627) * \text{MDM2_SWX} + (0,932293378426399) * \\
 & \text{PRKCB_Site23920910_J} - (0,235744342326864) * \text{MDM2_W} - (0,188045930057369) * \text{MDM2_Site68821643} - \\
 & (0,129971821791191) * \text{PRKCB_DIJK} - (0,112543222902224) * \text{MDM2_SVW} + (0,0714766020787205) * \text{CAMK1D_E} - \\
 & (-0,0566149561843161) * \text{MDM2_SW} + (0,343458398647021) * \text{PRKCB_B} - (0,0511754835571903) * \\
 & \text{GAB2_Site78331054} + (0,0777964634309061) * \text{PRKCB_Site23920911} + (-0,309760257245042) * \\
 & \text{PRKCB_Site23920919_K} + (0,0892196021964463) * \text{MDM2_T} + (0,0591702551145443) * \text{PRKCB_J} - \\
 & (0,163034043783367) * \text{KCNJ15_CE} - (-0,3554615992579) * \text{MDM2_Site68821616_X} + (0,194383441652219) * \\
 & \text{CAMK1D_AE} - (-0,0357242710579534) * \text{KCNJ15_C} + (-0,137804039545265) * \text{PRKCB_JK} - (0,073268779527063) * \\
 & \text{GAB2_CE} - (0,583431075372027) * \text{PDE8A_Site85096687_B} - (0,083656284732413) * \text{KCNJ15_Site38287529_C} + \\
 & (0,241597825714507) * \text{PRKCB_Site23920902_I} + (0,0897642626007787) * \text{PRKCB_DIJ} - (0,311212124874424) * \\
 & \text{PRKCB_Site23920893_G} - (-0,0694365074403336) * \text{KCNJ15_Site38287586} - (0,356961148091174) * \\
 & \text{KCNJ15_Site38287525_B} - (0,0284155839590126) * \text{MDM2_Site68821573} + (0,161274457136425) * \\
 & \text{IFNAR1_1_Site33355809} - (-0,578010271214577) * \text{PDE8A_B} + (0,00936832259402957) * \text{PRKCB_Site23920921_L} - \\
 & (-0,0220552696975651) * \text{MDM2_WX} + (0,0416981587266934) * \text{IFNAR1_1_Site33355817} - (0,056383762969179) * \\
 & \text{KCNJ15_Site38287589_E} + (0,034447202518999) * \text{IFNAR1_1_Site33355838_P} - (0,181855602464856) * \\
 & \text{KCNJ15_Site38287590} + (-0,232428154178395) * \text{PRKCB_Site23920884_D} - (0,0566507858081719) * \\
 & \text{KCNJ15_Site38287530_D} + (0,00350968357773277) * \text{PRKCB_IJK} - (0,0986696717892891) * \text{KCNJ15_Site38287591} \\
 & + (0,0416357072635886) * \text{PRKCB_Site23920889_F} + (-0,510062501581064) * \text{PRKCB_Site23920844_B} + (- \\
 & 0,201713338082801) * \text{PRKCB_DJ} - (0,219330533848547) * \text{LYN_AC} + (-0,0902753379195889) * \\
 & \text{IFNAR1_1_Site33355798} + (0,0364611306918377) * \text{PRKCB_DHIJK} + (-0,0154244319603596) * \text{GAB2_BCDE} + (- \\
 & 0,0721590982562632) * 0,130383083484889 * \text{LYN_AG} + (0,136853965680293) * \text{GAB2_Site78331030} - \\
 & (0,201599671305856) * \text{CAMK1D_Site12378440} + (-0,0494625908197476) * \text{IFNAR1_1_Site33355793_G} - \\
 & (0,112438468962481) * \text{MDM2_KSX} + (0,129569099073175) * \text{IFNAR1_1_Site33355839_L} - (0,12561875043809) * \\
 & \text{GAB2_ABDE} - (0,127875837752046) * \text{PRKCB_DJM} - (-0,421356734328941) * \text{KCNJ15_ACE} - \\
 & (0,125572361520418) * \text{MDM2_ASV} - (-0,0421188089161209) * \text{MDM2_Site68821576_V} + (0,333082982874752) * \\
 & \text{LYN_A} + (0,205974310972437) * \text{IFNAR1_1_Site33355762_F} - (0,26063881758559) * \text{MDM2_Site68821602} - \\
 & (0,16091158013623) * \text{LYN_Site56012491} - (0,585856554487505) * \text{PRKCB_Site23920927_M} - (- \\
 & 0,328951375049536) * \text{KCNJ15_Site38287526_A} + (0,0681422708577013) * \text{GAB2_EJ} + (0,0265646151191535) *
 \end{aligned}$$

PRKCB_I + (-0,114728592530574)* LYN_Site56012490 + (0,0423564487829415)* CAMK1D_Site12378387 + (-0,270495232628549)* IFNAR1,1_Site33355807_B - (0,202529212838335)* PRKCB_Site23920952 - (-0,0881700257476326)* KCNJ15_ACD + (0,0213312643325165)* GAB2_DE + (-0,00540846808767115)* MDM2_STX + (0,18510874500411)* GAB2_Site78331008_D - (0,060213548631588)* MDM2_Site68821567 - (-0,123438884158084)* MDM2_Site68821612_A

La lista de combinaciones sometidas a prueba en toda la población de la cohorte de validación (n = 411) se presenta en la tabla 6. Todas las combinaciones fueron estadísticamente significativas entre el control y los casos con un valor de $p < 10^{-4}$.

Tabla 6: Rendimientos diagnósticos obtenidos con pacientes de la población completa de la cohorte de validación (n = 411) analizados mediante secuenciación multiplexada de 8 dianas.

Núm.	mROC de combinación	Combi.	ID	Curva ROC de AUC	IC 95 %	ID de umbral	Prec. (%)	Sp (%)	Se (%)	VPP (%)	VPN (%)	Población
1	MDM2_Site68821586_W MDM2_SWX PRKCB_Site23920910_J MDM2_W MDM2_Site68821643_PRKCB_DIJK MDM2_SVW CAMK1D_E MDM2_SW PRKCB_B GAB2_Site78331054_PRKCB_Site23920911 PRKCB_Site23920919_K MDM2_T PRKCB_J KCNJ15_CE MDM2_Site68821616_X CAMK1D_AE KCNJ15_C PRKCB_JK GAB2_CE PDE8A_Site85096687_B KCNJ15_Site38287529_C PRKCB_Site23920902_I PRKCB_DIJ PRKCB_Site23920893_G KCNJ15_Site38287586 KCNJ15_Site38287525_B MDM2_Site68821573 IFNAR1.1_Site33355809 PDE8A_B PRKCB_Site23920921_L MDM2_WX	C80	49	0,889	[0,858;0,92]	-0,1901	81,02	82,52	80,22	89,58	69,01	N= 411 [143(1); 268(2)]

IFNAR1.1_Site 33355817 KCNJ15_Site3 8287589_E IFNAR1.1_Site 33355838_P KCNJ15_Site3 8287590 PRKCB_Site23 920884_D KCNJ15_Site3 8287530_D PRKCB_IJK KCNJ15_Site3 8287591 PRKCB_Site23 920889_F PRKCB_Site23 920844_B PRKCB_DJ LYN_AC IFNAR1.1_Site 33355798 PRKCB_DHIJK GAB2_BCDE IFNAR1.1_Site 33355801_N CAMK1D_A LYN_AG GAB2_Site783 31030 CAMK1D_Site 12378440 IFNAR1.1_Site 33355793_G MDM2_KSX IFNAR1.1_Site 33355839_L IFNAR1.1_Site 33355789_A GAB2_ABDE PRKCB_DJM KCNJ15_ACE MDM2_ASV MDM2_Site688 21576_V LYN_A IFNAR1.1_Site 33355762_F MDM2_Site688 21602 LYN_Site5601 2491 PRKCB_Site23 920927_M KCNJ15_Site3 8287526_A GAB2_EJ PRKCB_I LYN_Site5601 2490 CAMK1D_Site 12378387 IFNAR1.1_Site 33355807_B										
---	--	--	--	--	--	--	--	--	--	--

	PRKCB_Site23 920952 KCNJ15_ACD GAB2_DE MDM2_STX GAB2_Site783 31008_D MDM2_Site688 21567 MDM2_Sitio68 821612_A											
2	MDM2_Site688 21586_W MDM2_SWX PRKCB_Site23 920910_J MDM2_W MDM2_Site688 21643 PRKCB_DIJK MDM2_SVW CAMK1D_E MDM2_SW PRKCB_B GAB2_Site783 31054 PRKCB_Site23 920911 PRKCB_Site23 920919_K MDM2_T PRKCB_J KCNJ15_CE MDM2_Site688 21616_X CAMK1D_AE KCNJ15_C PRKCB_JK GAB2_CE PDE8A_Site85 096687_B KCNJ15_Site3 8287529_C PRKCB_Site23 920902_I PRKCB_DIJ PRKCB_Site23 920893_G KCNJ15_Site3 8287586 KCNJ15_Site3 8287525_B MDM2_Site688 21573 IFNAR1.1_Site 33355809 PDE8A_B PRKCB_Site23 920921_L IFNAR1.1_Site 33355840_E MDM2_WX IFNAR1.1_Site 33355817 KCNJ15_Site3 8287589_E	C80	8	0,889	[0,858;0,92]	-0,2879	82,97	79,02	88,14	89,58	71,52	N= 411 [143 (1); 268 (2)]

IFNAR1.1_Site 33355838_P KCNJ15_Site3 8287590 PRKCB_Site23 920884_D KCNJ15_Site3 8287530_D PRKCB_IJK KCNJ15_Site3 8287591 PRKCB_Site23 920889_F PRKCB_Site23 920844_B PRKCB_DJ LYN_AC IFNAR1.1_Site 33355798 PRKCB_DHIJK GAB2_BCDE IFNAR1.1_Site 33355801_N CAMK1D_A LYN_AG GAB2_Site783 31030 CAMK1D_Site 12378440 IFNAR1.1_Site 33355793_G MDM2_KSX IFNAR1.1_Site 33355839_L IFNAR1.1_Site 33355789_A GAB2_ABDE PRKCB_DJM KCNJ15_ACE MDM2_ASV MDM2_Site688 21576_V LYN_A IFNAR1.1_Site 33355762_F MDM2_Site688 21602 LYN_Site5601 2491 PRKCB_Site23 920927_M KCNJ15_Site3 8287526_A GAB2_EJ PRKCB_I LYN_Site5601 2490 CAMK1D_Site 12378387 PRKCB_Site23 920952 KCNJ15_ACD GAB2_DE MDM2_STX GAB2_Site783 31008_D										
--	--	--	--	--	--	--	--	--	--	--

	MDM2_Site688 21567 MDM2_Site688 21612_A											
3	MDM2_Site688 21586_W MDM2_SWX PRKCB_Site23 920910_J MDM2_W MDM2_Site688 21643 PRKCB_DIJK MDM2_SVW CAMK1D_E MDM2_SW PRKCB_B GAB2_Site783 31054 PRKCB_Site23 920911 PRKCB_Site23 920919_K MDM2_T PRKCB_J KCNJ15_CE MDM2_Site688 21616_X CAMK1D_AE KCNJ15_C PRKCB_JK GAB2_CE PDE8A_Site85 096687_B KCNJ15_Site3 8287529_C PRKCB_Site23 920902_I PRKCB_DIJ PRKCB_Site23 920893_G KCNJ15_Site3 8287586 KCNJ15_Site3 8287525_B MDM2_Site688 21573 IFNAR1.1_Site 33355809 PDE8A_B PRKCB_Site23 920921_L IFNAR1.1_Site 33355840_E MDM2_WX IFNAR1.1_Site 33355817 KCNJ15_Site3 8287589_E IFNAR1.1_Site 33355838_P KCNJ15_Site3 8287590 PRKCB_Site23 920884_D	C80	16	0,889	[0,858;0,92]	-0,2396	81,02	81,82	80,6	89,26	69,23	N= 411 [143 (1); 268 (2)]

KCNJ15_Site3 8287530_D PRKCB_IJK KCNJ15_Site3 8287591 PRKCB_Site23 920889_F PRKCB_Site23 920844_B PRKCB_DJ LYN_AC IFNAR1.1_Site 33355798 PRKCB_DHIJK GAB2_BCDE IFNAR1.1_Site 33355801_N CAMK1D_A LYN_AG GAB2_Site783 31030 CAMK1D_Site 12378440 IFNAR1.1_Site 33355793_G MDM2_KSX IFNAR1.1_Site 33355839_L IFNAR1.1_Site 33355789_A GAB2_ABDE PRKCB_DJM KCNJ15_ACE MDM2_ASV MDM2_Site688 21576_V LYN_A IFNAR1.1_Site 33355762_F LYN_Site5601 2491 PRKCB_Site23 920927_M KCNJ15_Site3 8287526_A GAB2_EJ PRKCB_I LYN_Site5601 2490 CAMK1D_Site 12378387 IFNAR1.1_Site 33355807_B PRKCB_Site23 920952 KCNJ15_ACD GAB2_DE MDM2_STX GAB2_Site783 31008_D MDM2_Site688 21567 MDM2_Site688 21612_A										
--	--	--	--	--	--	--	--	--	--	--

4	MDM2_Site688 21586_W MDM2_SWX PRKCB_Site23 920910_J MDM2_W MDM2_Site688 21643 PRKCB_DIJK MDM2_SVW CAMK1D_E MDM2_SW PRKCB_B GAB2_Site783 31054 PRKCB_Site23 920911 PRKCB_Site23 920919_K MDM2_T PRKCB_J KCNJ15_CE MDM2_Site688 21616_X CAMK1D_AE KCNJ15_C PRKCB_JK GAB2_CE PDE8A_Site85 096687_B KCNJ15_Site3 8287529_C PRKCB_Site23 920902_I PRKCB_DIJ PRKCB_Site23 920893_G KCNJ15_Site3 8287586 KCNJ15_Site3 8287525_B MDM2_Site688 21573 IFNAR1.1_Site 33355809 PDE8A_B PRKCB_Site23 920921_L IFNAR1.1_Site 33355840_E MDM2_WX IFNAR1.1_Site 33355817 KCNJ15_Site3 8287589_E IFNAR1.1_Site 33355838_P KCNJ15_Site3 8287590 PRKCB_Site23 920884_D KCNJ15_Site3 8287530_D PRKCB_IJK KCNJ15_Site3 8287591	C79	11 8	0,889	[0,857;0,92]	-0,0579	81,27	87,41	77,99	92,07	67,93	N= 411 [143 (1); 268 (2)]
---	--	-----	---------	-------	--------------	---------	-------	-------	-------	-------	-------	------------------------------

	PRKCB_Site23920889_F PRKCB_Site23920844_B PRKCB_DJ LYN_AC IFNAR1.1_Site33355798 PRKCB_DHIJK GAB2_BCDE IFNAR1.1_Site33355801_N CAMK1D_A LYN_AG GAB2_Site78331030 CAMK1D_Site12378440 IFNAR1.1_Site33355793_G MDM2_KSX IFNAR1.1_Site33355839_L IFNAR1.1_Site33355789_A GAB2_ABDE PRKCB_DJM KCNJ15_ACE MDM2_ASV MDM2_Site68821576_V LYN_A IFNAR1.1_Site33355762_F LYN_Site56012491 PRKCB_Site23920927_M GAB2_EJ PRKCB_I LYN_Site56012490 CAMK1D_Site12378387 IFNAR1.1_Site33355807_B PRKCB_Site23920952 KCNJ15_ACD GAB2_DE MDM2_STX GAB2_Site78331008_D MDM2_Site68821567 MDM2_Sitio68821612_A											
5	MDM2_Site68821586_W MDM2_SWX PRKCB_Site23920910_J MDM2_W MDM2_Site68821643	C79	218	0,887	[0,855;0,918]	-0,3557	83,7	78,32	86,57	88,21	75,68	N= 411 [143 (1); 268 (2)]

PRKCB_DIJK											
MDM2_SVW											
CAMK1D_E											
MDM2_SW											
PRKCB_B											
GAB2_Site783											
31054											
PRKCB_Site23											
20911											
PRKCB_Site23											
920919_K											
MDM2_T											
PRKCB_J											
KCNJ15_CE											
MDM2_Site688											
21616_X											
CAMK1D_AE											
KCNJ15_C											
PRKCB_JK											
GAB2_CE											
PDE8A_Site85											
096687_B											
KCNJ15_Site3											
8287529_C											
PRKCB_Site23											
920902_I											
PRKCB_DIJ											
PRKCB_Site23											
920893_G											
KCNJ15_Site3											
8287586											
KCNJ15_Site3											
8287525_B											
MDM2_Site688											
21573											
IFNAR1.1_Site											
33355809											
PDE8A_B											
PRKCB_Site23											
920921_L											
IFNAR1.1_Site											
33355840_E											
MDM2_WX											
IFNAR1.1_Site											
33355817											
KCNJ15_Site3											
8287589_E											
IFNAR1.1_Site											
33355838_P											
KCNJ15_Site3											
8287590											
PRKCB_Site23											
920884_D											
KCNJ15_Site3											
8287530_D											
PRKCB_IJK											
KCNJ15_Site3											
8287591											
PRKCB_Site23											
920889_F											
PRKCB_Site23											
920844_B											
PRKCB_DJ											
LYN_AC											
IFNAR1.1_Site											
33355798											

PRKCB_DHIJK GAB2_BCDE IFNAR1.1_Site 33355801_N CAMK1D_A LYN_AG GAB2_Site783 31030 CAMK1D_Site 12378440 IFNAR1.1_Site 33355793_G MDM2_KSX IFNAR1.1_Site 33355839_L IFNAR1.1_Site 33355789_A GAB2_ABDE KCNJ15_ACE MDM2_ASV MDM2_Site688 21576_V LYN_A IFNAR1.1_Site 33355762_F MDM2_Site688 21602 LYN_Site5601 2491 PRKCB_Site23 920927_M KCNJ15_Site3 8287526_A GAB2_EJ PRKCB_I LYN_Site5601 2490 CAMK1D_Site 12378387 PRKCB_Site23 920952 KCNJ15_ACD GAB2_DE MDM2_STX GAB2_Site783 31008_D MDM2_Site688 21567 MDM2_Site688 21612_A										
---	--	--	--	--	--	--	--	--	--	--

6	MDM2_Site688 21586_WMDM 2_SWX PRKCB_Site23 920910_J MDM2_W MDM2_Site688 21643 PRKCB_DIJK MDM2_SVW CAMK1D_E MDM2_SW PRKCB_B GAB2_Site783 31054 PRKCB_Site23 920911 PRKCB_Site23 920919_K MDM2_T PRKCB_J KCNJ15_CE MDM2_Site688 21616_XCAMK 1D_AE KCNJ15_C PRKCB_JK GAB2_CE PDE8A_Site85 096687_B KCNJ15_Site3 8287529_C PRKCB_Site23 920902_I PRKCB_DIJ PRKCB_Site23 920893_G KCNJ15_Site3 8287586 KCNJ15_Site3 8287525_B MDM2_Site688 21573 IFNAR1.1_Site 33355809 PDE8A_B PRKCB_Site23 920921_L MDM2_WX IFNAR1.1_Site 33355817 KCNJ15_Site3 8287589_E IFNAR1.1_Site 33355838_P KCNJ15_Site3 8287590 PRKCB_Site23 920884_D KCNJ15_Site3 8287530_D PRKCB_IJK KCNJ15_Site3 8287591 PRKCB_Site23 920889_F	C79	11 52	0,889	[0,858;0,92]	-0,2397	82	82,52	81,72	89,75	70,66	N= 411 [143 (1); 268 (2)]
---	---	-----	----------	-------	--------------	---------	----	-------	-------	-------	-------	------------------------------

PRKCB_Site23 920844_B PRKCB_DJ LYN_AC IFNAR1.1_Site 33355798 PRKCB_DHIJK GAB2_BCDE IFNAR1.1_Site 33355801_N CAMK1D_A LYN_AG GAB2_Site783 31030 CAMK1D_Site 12378440 IFNAR1.1_Site 33355793_G MDM2_KSX IFNAR1.1_Site 33355839_L GAB2_ABDE PRKCB_DJM KCNJ15_ACE MDM2_ASV MDM2_Site688 21576_V LYN_A IFNAR1.1_Site 33355762_F MDM2_Site688 21602 LYN_Site5601 2491 PRKCB_Site23 920927_M KCNJ15_Site3 8287526_A GAB2_EJ PRKCB_I LYN_Site5601 2490 CAMK1D_Site 12378387 IFNAR1.1_Site 33355807_B PRKCB_Site23 920952 KCNJ15_ACD GAB2_DE MDM2_STX GAB2_Site783 31008_D MDM2_Site688 21567 MDM2_Site688 21612_A										
---	--	--	--	--	--	--	--	--	--	--

7	MDM2_Site688 21586_W MDM2_SWX PRKCB_Site23 920910_J MDM2_W MDM2_Site688 21643 PRKCB_DIJK MDM2_SVW MDM2_SW PRKCB_B GAB2_Site783 31054 PRKCB_Sitio2 3920911 PRKCB_Sitio2 3920919_K MDM2_T PRKCB_J KCNJ15_CE MDM2_Site688 21616_XCAMK 1D_AE KCNJ15_C PRKCB_JK PDE8A_Site85 096687_B KCNJ15_Site3 8287529_C PRKCB_Site23 920902_I PRKCB_DIJ KCNJ15_Site3 8287586 KCNJ15_Site3 8287525_B MDM2_Site688 21573 IFNAR1.1_Site 33355809 PDE8A_B PRKCB_Site23 920921_L IFNAR1.1_Site 33355840_E MDM2_WX IFNAR1.1_Site 33355817 KCNJ15_Site3 8287589_E IFNAR1.1_Site 33355838_P KCNJ15_Site3 8287590 PRKCB Site23920884_ D KCNJ15_Site3 8287530_D PRKCB_IJK KCNJ15_Site3 8287591 PRKCB_Site23 920889_F	C53	## ##	0,825	[0,783;0,867]	-0,1737	80,29	76,92	76,87	86,19	63,95	N= 411 [143 (1); 268 (2)]
---	--	-----	----------	-------	---------------	---------	-------	-------	-------	-------	-------	------------------------------

PRKCB_Site23 920844_B PRKCB_DJ LYN_AC IFNAR1.1_Site 33355798 PRKCB_DHIJK GAB2_BCDE IFNAR1.1_Site 33355801_N CAMK1D_A LYN_AG GAB2_Site783 31030 CAMK1D_Site 12378440 IFNAR1.1_Site 33355793_G MDM2_KSX											
---	--	--	--	--	--	--	--	--	--	--	--

8	MDM2_Site688 21586_W MDM2_SWX PRKCB_Site23 920910_J MDM2_W MDM2_Site688 21643 PRKCB_DIJK MDM2_SVW CAMK1D_E MDM2_SW PRKCB_B GAB2_Site783 31054_PRKCB _Site23920911 PRKCB_Site23 920919_KMDM 2_T PRKCB_J KCNJ15_CE MDM2_Site688 21616_X CAMK1D_AE KCNJ15_C PRKCB_JK PDE8A_Site85 096687_B KCNJ15_Site3 8287529_C PRKCB_Site23 920902_I PRKCB_DIJ KCNJ15_Site3 8287586 KCNJ15_Site3 8287525_B MDM2_Site688 21573 IFNAR1.1_Site 33355809 PDE8A_B PRKCB_Site23 920921_L IFNAR1.1_Site 33355840_E MDM2_WX IFNAR1.1_Site 33355817 KCNJ15_Site3 8287589_E IFNAR1.1_Site 33355838_P KCNJ15_Site3 8287590 PRKCB_Site23 920884_D KCNJ15_Site3 8287530_D PRKCB_IJK PRKCB_Site23 920889_F PRKCB_Site23 920844_B PRKCB_DJ LYN_AC	C53	69 95	0,824	[0,782;0,865]	-0,2023	879,32	76,22	80,97	86,45	68,12	N= 411 [143 (1); 268 (2)]
---	--	-----	----------	-------	---------------	---------	--------	-------	-------	-------	-------	------------------------------

	IFNAR1.1_Site 33355798 PRKCB_DHIJK GAB2_BCDE IFNAR1.1_Site 33355801_N CAMK1D_A LYN_AG GAB2_Site783 31030 CAMK1D_Site 12378440 IFNAR1.1_Site 33355793_G MDM2_KSX											
9	MDM2_SWX PRKCB_Site23 920910_J MDM2_W PRKCB_DIJK MDM2_SVW CAMK1D_E MDM2_SW PRKCB_B GAB2_Site783 31054 PRKCB_Site23 920911 PRKCB_Site23 920919_K MDM2_T PRKCB_J KCNJ15_CE MDM2_Site688 21616_X CAMK1D_AE KCNJ15_C PRKCB_JK	C18	18 7	0,752	[0,704;0,8]	-0,0589	71,1	71,3	70,9	82,3	56,7	N= 411 [143 (1); 268 (2)]
10	MDM2_Site688 21586_W MDM2_SWX PRKCB_Site23 920910_J MDM2_Site688 21643 PRKCB_DIJK MDM2_SVW CAMK1D_E MDM2_SW GAB2_Site783 31054 PRKCB_Site23 920911 PRKCB_Site23 920919_K MDM2_T PRKCB_J KCNJ15_CE MDM2_Site688 21616_X CAMK1D_AE KCNJ15_C	C17	60 6	0,751	[0,702;0,799]	-0,0981	71,78	70,63	72,39	82,2	57,71	N= 411 [143 (1); 268 (2)]

El panel de biomarcadores analizado anteriormente dio unos rendimientos diagnósticos excelentes en el diagnóstico de DEP en toda la población de la cohorte de validación (411 sujetos), lo que sugiere que el presente panel de biomarcadores puede de hecho tener aplicaciones clínicas en el diagnóstico de DEP.

A continuación, buscamos investigar si el rendimiento diagnóstico de nuestro ensayo podría mejorarse adicionalmente teniendo en cuenta las diferencias específicas según el sexo. De hecho, se ha sugerido la existencia de firmas transcripcionales específicas del sexo en DEP. Se ha mostrado previamente que el perfil transcripcional cerebral de DEP difiere en gran medida según el sexo, lo que sugiere que los mecanismos moleculares subyacentes a DEP podrían diferir entre mujeres y hombres⁴². Por tanto, se analizaron por separado unas modificaciones en la edición del ARN en la población masculina y femenina de la cohorte de validación. Cuando se introdujo la especificidad del sexo en el análisis, se observó una mejora significativa en el rendimiento diagnóstico. Cabe destacar que la cohorte de validación femenina (n = 277) incluyó más sujetos que la cohorte de validación masculina (n = 134), ya que esto refleja la proporción de hombres y mujeres que padecen DEP en la población general en todo el mundo. De hecho, en la población femenina de la cohorte de validación (n = 277, 90 controles sanos y 187 pacientes con DEP), cuando se combinaron 11 variantes de edición del ARN, que representan 4 BMK diferentes, el AUC fue de 0,784 (Sp 71,1 %; Se 78,6 %, figura 7A). La ecuación usada en este caso es la siguiente:

$$Z = -(-0,000533736120198011) * \text{MDM2_Site68821586_W} - (0,211359639047036) * \text{MDM2_W} - (0,146264452332436) * \text{MDM2_SWX} - (0,160086235452429) * \text{GAB2_Site78331054} - (0,090792634422732) * \text{MDM2_SW} + (0,117672779615208) * \text{MDM2_T} + (0,209398709260439) * \text{CAMK1D_AE} + (0,115253600910406) * \text{PRKCB_J} - (0,193553830436994) * \text{CAMK1D_BC} + (0,0404496406797322) * \text{CAMK1D_E} + (0,0771184775348071) * \text{PRKCB_Site23920910_J}$$

Cuando se aumentó el número de biomarcadores a 6 dianas diferentes y se incluyeron 30 variantes de edición de ARN, el AUC aumentó a 0,835 (Sp 80,0 %; Se 83,4 %, figura 7B). Además, cuando se usaron los 8 biomarcadores incluidos en el multiplex y se combinaron 49 sitios editados de manera diferencial, el AUC alcanzó 0,880, tanto con la especificidad como con la sensibilidad por encima del 80 % (Sp 83,3 %; Se 88,2 %, figura 7C).

Usando un enfoque de aprendizaje automático con mROC y Bosque aleatorio (véanse el material y los métodos), separamos aleatoriamente las poblaciones específicas por sexo en un conjunto de entrenamiento y uno de prueba. Al dividir aleatoriamente la población femenina en un conjunto de entrenamiento (75 % de toda la población) y un conjunto de prueba (25 % de toda la población) y aplicar los métodos mencionados anteriormente, se obtuvo un AUC de 0,843 en el conjunto de prueba con el método mROC (Sp 88,9 %; Se 80,4 %, figura 7D) y un AUC de 0,895 con el método de bosque aleatorio MultDS (Sp 86 %; Se 80 %, figura 7E). En conjunto, los datos obtenidos mediante dos métodos diferentes de aprendizaje automático en la cohorte de validación femenina confirmaron el excelente rendimiento diagnóstico de nuestro panel de biomarcadores.

La lista de combinaciones sometidas a prueba en la cohorte de validación femenina (n=277) se presenta en la tabla 16. Todas las combinaciones fueron estadísticamente significativas entre el control y los casos con un valor de $p < 10^{-4}$.

Tabla 16: Rendimientos diagnósticos obtenidos con pacientes de la población femenina de la cohorte de validación analizados mediante secuenciación multiplexada de 8 dianas.

Núm.	mROC de combinación	Combi.	ID	Curva ROC de AUC	IC 95 %	Umbral	Prec. (%)	Sp (%)	Se (%)	VPP (%)	VPN (%)	Población
	MDM2_Site68821586_W MDM2_W MDM2_SWX GAB2_Site78331054 MDM2_SW MDM2_T CAMK1D_AE PRKCB_J CAMK1D_BC CAMK1D_E MDM2_Site68821612_A MDM2Site68821576_V MDM2_Site68821643 PRKCB_Site23920910_J MDM2_Site68821573 PRKCB_DJ MDM2_Site68821616_X CAMK1D_AF PRKCB_DHIJK MDM2_Site68821570											

Núm.	mROC de combinación	Combi.	ID	Curva ROC de AUC	IC 95 %	Umbral	Prec. (%)	Sp (%)	Se (%)	VPP (%)	VPN (%)	Población
1	MDM2_Site68821602 MDM2_WX MDM2_SVW MDM2_SVWX CAMK1D_Site12378375_G PRKCB_I CAMK1D_A PDE8A_Site85096687_B CAMK1D_Site12378321_C PRKCB_Site23920952 MDM2_X LYN_A CAMK1D_ABG PRKCB_JK CAMK1D_FS GAB2_Site78331030 MDM2_K KCNJ15_C GAB2_IJ PRKCB_Site23920921_L PRKCB_IJK IFNAR1.1_Site33355809 MDM2_Site68821572_T GAB2_BDEHJ GAB2_H PRKCB_AD PDE8A_B KCNJ15_CD GAB2_BCDJ IFNAR1.1_Site33355840_E	C50	443	0,883	[0,841; 0,925]	- 0,2642	85,6	84,4	86,1	92,0	74,5	N= 277 [90 (1); 187 (2)]
2	MDM2_Site68821586_W MDM2_W MDM2_SWX GAB2_Site78331054 MDM2_SW MDM2_T CAMK1D_AE PRKCB_J CAMK1D_BC CAMK1D_E MDM2_Site68821612_A MDM2_Site68821576_V MDM2_Site68821643 PRKCB_Site23920910_J MDM2_Site68821573 PRKCB_DJ MDM2_Site68821616_X CAMK1D_AF PRKCB_DHIJK MDM2_Site68821570 MDM2_Site68821602 MDM2_WX MDM2_SVW MDM2_SVWX CAMK1D_Site12378375_G PRKCB_I CAMK1D_A PDE8A_Site85096687_B CAMK1D_Site12378321_C PRKCB_Site23920952 MDM2_X LYN_A CAMK1D _ABG PRKCB_JK	C50	440	0,880	[0,837; 0,923]	- 0,2692	85,2	85,6	85,0	92,4	73,3	N= 277 [90 (1); 187 (2)]

Núm.	mROC de combinación	Combi.	ID	Curva ROC de AUC	IC 95 %	Umbral	Prec. (%)	Sp (%)	Se (%)	VPP (%)	VPN (%)	Población
	CAMK1D_FS GAB2_Site78331030 MDM2_K KCNJ15_C GAB2_IJ PRKCB_Site23920921_L PRKCB_IJK IFNAR1.1_Site33355809 MDM2_Site68821572_T GAB2_BEJ GAB2_BDEHJ GAB2_H PDE8A_B KCNJ15_CD GAB2_BCDJ IFNAR1.1_Site33355840_E											
3	MDM2_Site68821586_W MDM2_W MDM2_SW MDM2_T CAMK1D_AE PRKCB_J CAMK1D_BC CAMK1D_E MDM2_Site68821612_A MDM2_Site68821576_V PRKCB_Site23920910_J MDM2_Site68821573 PRKCB_DJ MDM2_Site68821616_X CAMK1D_AF PRKCB_DHIJK MDM2_Site68821570 MDM2_Site68821602 PRKCB_DIJK MDM2_WX MDM2_SVW MDM2_SVWX CAMK1D_Site12378375_G PRKCB_I CAMK1D_A PDE8A_Site85096687_B CAMK1D_Site12378321_C PRKCB_Site23920952 MDM2_X LYN_A CAMK1D_ABG PRKCB_JK CAMK1D_FS GAB2_Site78331030 MDM2_K KCNJ15_C GAB2_IJ PRKCB_Site23920921_L PRKCB_IJK IFNAR1.1_Site33355809 MDM2_Site68821572_T GAB2_BEJ GAB2_BDEHJ GAB2_H PRKCB_AD PDE8A_B	C49	19592	0,880	[0,834; 0,925]	- 0,3719	87,0	80,0	90,4	90,4	80,0	N= 277 [90 (1); 187 (2)]

Núm.	mROC de combinación	Combi.	ID	Curva ROC de AUC	IC 95 %	Umbral	Prec. (%)	Sp (%)	Se (%)	VPP (%)	VPN (%)	Población
	KCNJ15_CD GAB2_BCDJ IFNAR1.1 Site33355840 E											
4	MDM2_Site68821586_W MDM2_W MDM2_SWX MDM2_SW MDM2_T CAMK1D_AE PRKCB_J CAMK1D_BC CAMK1D_E MDM2_Site68821612_A MDM2_Site68821576_V MDM2_Site68821643 PRKCB_Site23920910_J MDM2_Site68821573 MDM2_Site68821616_X CAMK1D_AF PRKCB_DHIJK MDM2_Site68821570 MDM2_Site68821602 PRKCB DIJK MDM2_WX MDM2_SVW MDM2_SVWX CAMK1D_Site12378375_G PRKCB_I CAMK1D_A PDE8A_Site85096687_B CAMK1D_Site12378321_C PRKCB_Site23920952 MDM2_X LYN_A CAMK1D _ABG PRKCB_JK CAMK1D_FS GAB2_Site78331030 MDM2_K KCNJ15_C PRKCB_Site23920921_L PRKCB_IJK IFNAR 1,1_Site33355809 MDM2_Site68821572_T GAB2_BEJ GAB2_BDEHJ GAB2_H PRKCB_AD - PDE8A_B KCNJ15_CD GAB2_BCDJ IFNAR1,1 Site33355840 E	C49	17939	0,880	[0,835; 0,924]	- 0,2916	86,6	83,3	88,2	91,7	77,3	N= 277 [90 (1); 187 (2)]
	MDM2_Site68821586_W MDM2_W MDM2_SWX MDM2_SW MDM2_T CAMK1D_AE PRKCB_J CAMK1D_BC MDM2_Site68821612_A MDM2_Site68821643 PRKCB_Site23920910_J											

Núm.	mROC de combinación	Combi.	ID	Curva ROC de AUC	IC 95 %	Umbral	Prec. (%)	Sp (%)	Se (%)	VPP (%)	VPN (%)	Población
5	MDM2_Site68821573 PRKCB_DJ MDM2_Site68821616_X CAMK1D_AF PRKCB DHIJK MDM2_Site68821570 MDM2_Site68821602 PRKCB DIJK MDM2_WX MDM2_SVW MDM2_SVWX CAMK1DG PRKCB_I CAMK1D_A _Site12378375_ PDE8A_Site85096687_B CAMK1D_Site12378321_C PRKCB_Site23920952 MDM2_X LYN_A CAMK1D_ABG PRKCB_JK CAMK1D_FS GAB2_Site78331030 MDM2_K KCNJ15_C GAB2_IJ PRKCB_Site23920921_L PRKCB IJK IFNAR1,1Site33355809 MDM2_Site68821572_T GAB2_BEJ GAB2_BDEHJ GAB2_H PRKCB_AD PDE8A_B KCNJ15_CD GAB2_BCDJ IFNAR1,1 Site33355840 E	C49	18198	0,879	[0,835; 0,924]	- 0,3089	86,6	82,2	88,8	91,2	77,9	N= 277 [90 (1); 187 (2)]
6	MDM2_Site68821586_W MDM2_SWX GAB2_Site78331054 MDM2_SW MDM2_T CAMK1D_AE PRKCB_J CAMK1D_BC CAMK1D_E MDM2_Site68821612_A MDM2Site68821576_V MDM2_Site68821643 PRKCB_Site23920910_J MDM2_Site68821573 PRKCB_DJ MDM2_Site68821616_X CAMK1D_AF PRKCB_DHIJK MDM2_Site68821570	C31	4531	0,835	[0,782; 0,887]	- 0,2749	82,3	80,0	83,4	89,7	69,9	N= 277 [90 (1); 187 (2)]

Núm.	mROC de combinación	Combi.	ID	Curva ROC de AUC	IC 95 %	Umbral	Prec. (%)	Sp (%)	Se (%)	VPP (%)	VPN (%)	Población
	MDM2_Site68821602 PRKCB _ DIJK MDM2_WX MDM2_SVW PRKCB_I CAMK1D_A PDE8A_Site85096687_B CAMK1D_Site12378321_C PRKCB_Site23920952 MDM2_X LYN_A											
7	MDM2_Site68821586_W MDM2_SWX GAB2_Site78331054 MDM2_SW MDM2_T CAMK1D_AE PRKCB_J CAMK1D_BC CAMK1D_E MDM2_Site68821612_A MDM2_Site68821576_V MDM2_Site68821643 PRKCB_Site23920910_J MDM2_Site68821573 PRKCB_DJ MDM2_Site68821616_X CAMK1D_AF PRKCB DHIJK MDM2_Site68821570 PRKCB DIJK MDM2_WX MDM2_SVW CAMK1D Site12378375 G PRKCB I CAMK1D A PDE8A_Site85096687_B CAMK1D_Site12378321_C PRKCB_Site23920952 MDM2_X LYN_A	C31	4570	0,834	[0,781; 0,886]	-0,279	82,3	80,0	83,4	89,7	69,9	N= 277 [90 (1); 187 (2)]
8	MDM2_Site68821586_W MDM2_W MDM2_SWX GAB2_Site78331054 MDM2_SW MDM2_T CAMK1D_AE PRKCB_J CAMK1D_BC CAMK1D_E MDM2_Site68821612_A MDM2_Site68821576_V PRKCB_Site23920910_J MDM2_Site68821573 PRKCB DJ MDM2 Site68821616 X PRKCB_DHIJK MDM2_Site68821570	C31	1254	0,833	[0,78; 0,886]	- 0,1426	80,9	81,1	80,8	89,9	67,0	N= 277 [90 (1); 187 (2)]

Núm.	mROC de combinación	Combi.	ID	Curva ROC de AUC	IC 95 %	Umbral	Prec. (%)	Sp (%)	Se (%)	VPP (%)	VPN (%)	Población
	MDM2_Site68821602 PRKCB - DIJK MDM2_WX MDM2_SVW CAMK1D_Site12378375_G PRKCB_I CAMK1D_A PDE8A_Site85096687_B CAMK1D_Site12378321_C PRKCB_Site23920952 MDM2_X LYN_A											
9	MDM2_Site68821586_W MDM2_W GAB2_Site78331054 MDM2_T CAMK1D_AE PRKCB_J CAMK1D_BC CAMK1D_E MDM2_Site68821612_A MDM2_Site68821576_V MDM2_Site68821643	C11	202	0,781	[0,723; 0,838]	- 0,0475	74,7	78,9	72,7	87,7	58,2	N= 277 [90 (1); 187 (2)]
10	MDM2_W MDM2_SWX GAB2_Site78331054 MDM2_T CAMK1D_AE PRKCB_J D_BC CAMK1 CAMK1D_E MDM2_Site68821612_A MDM2_Site68821643 PRKCB_Site23920910_J	C11	325	0,787	[0,731; 0,842]	- 0,0183	73,7	80,0	70,6	88,0	56,7	N= 277 [90 (1); 187 (2)]
11	MDM2_Site68821586_W MDM2_W MDM2_SWX GAB2_Site78331054 MDM2_SW MDM2_T CAMK1D_AE PRKCB_J CAMK1D_BC CAMK1D_E PRKCB_Site23920910_J	C11	4	0,784	[0,729; 0,84]	- 0,1626	76,2	71,1	78,6	85,0	61,5	N= 277 [90 (1); 187 (2)]

El rendimiento de la población masculina fue aún mejor. De hecho, al analizar por separado la población masculina de la cohorte de validación (n = 134, 53 controles sanos y 81 pacientes con DEP) e incluir 8 variantes de edición en 6 BMK, se obtuvo un AUC de 0,831 (Sp 83,0 %; Se 75,3 %, figura 8A). En este caso, la ecuación usada es la siguiente:

$$Z = - (0,238400918804649) * GAB2_CH - (0,230156347297484) * CAMK1D_DS - (0,19154950723228) * KCNJ15_Site38287525_B - (0,21914392389275) * GAB2_CE - (0,155783908162178) * LYN_Site56012491 - (0,268738279594405) * MDM2_SVW - (0,107036976834934) * MDM2_Site68821643 + (0,400819265369961) * PRKCB_Site23920893_G,$$

Con 7 biomarcadores, incluyendo 39 variantes de edición de ARN, el AUC estaba por encima de 0,9 (AUC 0,936, Spe 92,5 %, Se 87,7 %, figura 8B). Cuando se combinaron los 8 biomarcadores incluidos en el panel, con 68 variantes de edición en estos BMK, se obtuvo un AUC sobresaliente de 0,976 (Spe 98,1 %, Se 93,8 %, figura 8C). Usando aprendizaje automático, se dividió aleatoriamente la población masculina en un conjunto de entrenamiento (el 70 % de toda la población) y un conjunto de prueba (el 30 % de toda la población) y se obtuvo un AUC con mROC de 0,859 (Sp 87,5 %, Se 87,5 %, figura 8D), mientras que el método de bosque aleatorio condujo a un AUC de 0,875 (Sp 82 %, Se 80 %, figura 8E).

La lista de combinaciones sometidas a prueba en la cohorte de validación masculina (n=134) se presenta en la tabla 8. Todas las combinaciones fueron estadísticamente significativas entre el control y los casos con un valor de $p < 10^{-4}$.

Tabla 8: Rendimientos diagnósticos obtenidos con pacientes de la población masculina (n=134) de la cohorte de validación analizados mediante secuenciación multiplexada de 8 dianas.

5

Núm.	mROC de combinación	Combi.	ID	Curva ROC de AUC	IC 95 %	Umbral	Prec. (%)	Sp (%)	Se (%)	VPP (%)	VPN (%)	Población
1	GAB2_CH CAMK1D_DS KCNJ15_Site38287525_B GAB2_CE LYN_Site56012491 LYN_AC KCNJ15_Site38287530_D MDM2_SVW MDM2_Site68821643 KCNJ15_ACE PRKCB_Site23920893_G KCNJ15_Site38287529_C CAMK1D_BF KCNJ15_CE PDE8A_E KCNJ15_ACD MDM2_KST GAB2_JK PRKCB_JM PRKCB_ADJK CAMK1D_D GAB2_Site78330939_H KCNJ15_Site38287526_A CAMK1D_R GAB2_ABDE KCNJ15_C KCNJ15_Site38287531 CAMK1D_AF GAB2_H GAB2_DH GAB2_Site78331054 LYN_Site56012495 MDM2_KSX LYN_Site56012468_C MDM2_BG LYN_Site56012521 LYN_Site56012556_F KCNJ15_AC PRKCB_DJK GAB2_Site78331011_C KCNJ15_ABC PRKCB_DJM PDE8A_Site85096729_F CAMK1D_F GAB2_DEHJ KCNJ15_Site38287591 LYN_Site56012567 PRKCB_Site23920911 LYN_Site56012553_A MDM2_SWX KCNJ15_Site38287586 CAMK1D_FS GAB2_DK KCNJ15_AB PRKCB_Site23920889_F PRKCB_AJ PRKCB_Site23920919_K LYN_Site56012469_D PRKCB_Site23920910_J IFNAR1,1_AC LYN_Site56012446 PRKCB_Site23920884_D IFNAR1,1_C	C68	1992	0,976	[0,952; 1]	-0,7316	95,5	98,1	93,8	98,7	91,2	N= 134 [53 (1); 81 (2)]

Núm.	mROC de combinación	Combi.	ID	Curva ROC de AUC	IC 95 %	Umbral	Prec. (%)	Sp (%)	Se (%)	VPP (%)	VPN (%)	Población
	GAB2_BCDE LYN_Site56012568 PRKCB_IJK IFNAR1,1_AB PDE8A_B											
2	GAB2_CH CAMK1D_DS KCNJ15_Site38287525_B LYN_Site56012491 LYN_AC LYN_AF KCNJ15_Site38287530_D MDM2_SVW MDM2_Site68821643 KCNJ15_ACE PRKCB_Site23920893_G KCNJ15_Site38287529_C CAMK1D_BF KCNJ15_CE PDE8A_E KCNJ15_ACD MDM2_KST GAB2_JK PRKCB_JM PRKCB_ADJK CAMK1D_D GAB2_Site78330939_H KCNJ15_Site38287526_A CAMK1D_R GAB2_ABDE KCNJ15_C KCNJ15_Site38287531 CAMK1D_AF GAB2_H GAB2_DH LYN_Site56012495 MDM2_KSX LYN_Site56012468_C MDM2_BG LYN_Site56012521 LYN_Site56012556_F KCNJ15_AC PRKCB DIJK GAB2_Site78331011_C KCNJ15_ABC PRKCB_DJM PDE8A_Site85096729_F CAMK1 D_F GAB2_DEHJ KCNJ15_Site38287591 LYN_Site56012567 PRKCB_Site23920911 LYN_Site56012553_A MDM2_SWX KCNJ15_Site38287586 CAMK1D_FS GAB2_DK KCNJ15_AB PRKCB_Site23920889_F PRKCB_AJ PRKCB_Site23920919_K LYN_Site56012469_D PRKCB_Site23920910_J IFNAR1,1_AC LYN_Site56012446 PRKCB_Site23920884_D IFNAR1,1_C GAB2_BCDE LYN_Site56012568 PRKCB_IJK IFNAR1,1 AB PDE8A B	C67	46501	0,975	[0,953; 0,997]	-0,679	95,5	100,0	92,6	100,0	89,8	N= 134 [53 (1); 81 (2)]
	GAB2_CH CAMK1D_DS KCNJ15_Site38287525_B											

Núm.	mROC de combinación	Combi.	ID	Curva ROC de AUC	IC 95 %	Umbral	Prec. (%)	Sp (%)	Se (%)	VPP (%)	VPN (%)	Población
3	GAB2_CE LYN_Site56012491 LYN_AC KCNJ15_Site38287530_D MDM2_SVW MDM2_Site68821643 KCNJ15_ACE PRKCB_Site23920893_G CAMK1D_BF KCNJ15_CE PDE8A_E KCNJ15_ACD MDM2_KST GAB2_JK PRKCB_JM PRKCB_ADJK CAMK1D_D GAB2_Site78330939_H KCNJ15_Site38287526 A CAMK1D R GAB2_ABDE KCNJ15_C KCNJ15_Site38287531 CAMK1D_AF GAB2_H GAB2_DH GAB2_Site78331054 LYN_Site56012495 MDM2_KSX LYN_Site56012468_C MDM2_BG LYN_Site56012521 LYN_Site56012556_F KCNJ15_AC PRKCB_ DIJK GAB2_Site78331011 _C KCNJ15 ABC PRKCB DJM PDE8A_Site85096729_F CAMK1D_F GAB2_DEHJ KCNJ15_Site38287591 LYN_Site56012567 PRKCB_Site23920911 LYN_Site56012553_A MDM2_SWX KCNJ15_Site38287586 CAMK1D_FS GAB2_DK KCNJ15_AB PRKCB_Site23920889_F PRKCB_AJ PRKCB_Site23920919_K LYN_Site56012469_D PRKCB_Site23920910_J IFNAR1,1_AC LYN_Site56012446 PRKCB_Site23920884_D IFNAR1,1_C GAB2_BCDE LYN_Site56012568 PRKCB_IJK IFNAR1,1 AB PDE8A B	C67	41346	0,977	[0,952; 1]	-0,7686	95,5	98,1	93,8	98,7	91,2	N= 134 [53 (1); 81 (2)]
	GAB2_CH CAMK1D_DS KCNJ15_Site38287525_B											
	GAB2_CE LYN_Site56012491 LYN_AC LYN_AF											
	KCNJ15_Site38287530_D MDM2_SVW											

Núm.	mROC de combinación	Combi.	ID	Curva ROC de AUC	IC 95 %	Umbral	Prec. (%)	Sp (%)	Se (%)	VPP (%)	VPN (%)	Población
4	MDM2_Site68821643 KCNJ15_ACE PRKCB_Site23920893_G KCNJ15_Site38287529_C CAMK1_D_BF KCNJ15_CE PDE8A_E KCNJ15_ACD MDM2_KST GAB2_JK PRKCB_JM PRKCB_ADJK CAMK1D_D GAB2_Site78330939_H KCNJ15_Site38287526_A CAMK1D_R GAB2_ABDE KCNJ15_C KCNJ15_Site38287531 CAMK1D_AF GAB2_H GAB2_DH PRKCB_B GAB2_Site78331054 LYN_Site56012495 MDM2_KSX LYN_Site56012468_C MDM2_BG LYN_Site56012556_F KCNJ15_AC PRKCB_DIJK GAB2_Site78331011_C KCNJ15_ABC PRKCB_DJM PDE8A_Site85096729_F CAMK1D_F GAB2_DEHJ KCNJ15_Site38287591 LYN_Site56012567 PRKCB_Site23920911 LYN_Site56012553_A MDM2_SWX KCNJ15_Site38287586 CAMK1D_FS GAB2_DK PRKCB_Site23920889_F PRKCB_AJ PRKCB_Site23920919_K LYN_Site56012469_D PRKCB_Site23920910_J IFNAR1,1_AC LYN_Site56012446 IFNAR1,1_C GAB2_BCDE LYN_Site56012568 PRKCB_IJK IFNAR1,1_AB PDE8A B	C67	5058	0,980	[0,958; 1]	-1,2498	95,5	96,2	95,1	97,5	92,7	N= 134 [53 (1); 81 (2)]
	GAB2_CH CAMK1D_DS KCNJ15_Site38287525_B GAB2_CE LYN_Site56012491 LYN_AC LYN_AF KCNJ15_Site38287530_D MDM2_SVW MDM2_Site68821643 KCNJ15_ACE PRKCB_Site23920893_G KCNJ15_Site38287529_C CAMK1_D_BF KCNJ15_CE PDE8A_E KCNJ15_ACD GAB2_JK											

Núm.	mROC de combinación	Combi.	ID	Curva ROC de AUC	IC 95 %	Umbral	Prec. (%)	Sp (%)	Se (%)	VPP (%)	VPN (%)	Población
5	PRKCB_ADJK CAMK1D_D GAB2_Site78330939_H KCNJ15_Site38287526_A CAMK1D_R GAB2_ABDE KCNJ15_C KCNJ15_Site38287531 CAMK1D_AF GAB2_H GAB2_DH PRKCB_B GAB2_Site78331054 LYN_Site56012495 MDM2_KSX LYN_Site56012468_C LYN_Site56012521 KCNJ15_AC KCNJ15_ABC PRKCB_DJM PDE8A_Site85096729_F	C39	2261	0,936	[0,898; 0,975]	-0,0614	88,1	100,0	80,3	100,0	76,8	N= 134 [53 (1); 81 (2)]
6	GAB2_CH CAMK1D_DS KCNJ15_Site38287525_B GAB2_CE LYN_Site56012491 LYN_AC LYN_AF KCNJ15_Site38287530_D MDM2_SVW MDM2_Site68821643 KCNJ15_ACE PRKCB_Site23920893_G KCNJ15_Site38287529_C KCNJ15_CE PDE8A_E KCNJ15_ACD MDM2_KST GAB2_JK PRKCB_ADJK CAMK1 D_D GAB2_Site78330939_H KCNJ15_Site38287526_A CAMK1D_R GAB2_ABDE KCNJ15_C KCNJ15_Site38287531 CAMK1D_AF GAB2_H GAB2_DH PRKCB_B GAB2_Site78331054 MDM2_KSX LYN_Site56012468_C MDM2_BG LYN_Site56012521 KCNJ15_AC KCNJ15_ABC PRKCB DJM PDE8A Site85096729_F	C39	3516	0,936	[0,896; 0,975]	-0,3892	89,6	92,5	87,7	94,7	83,1	N= 134 [53 (1); 81 (2)]
7	GAB2_CH CAMK1D_DS KCNJ15_Site38287525_B GAB2_CE LYN_Site56012491 LYN_AC LYN_AF KCNJ15_Site38287530_D MDM2_SVW MDM2_Site68821643 KCNJ15_ACE PRKCB_Site23920893_G KCNJ15_Site38287529_C KCNJ15_CE PDE8A_E	C39	3287	0,937	[0,898; 0,975]	-0,4206	89,6	88,7	90,1	92,4	85,5	N= 134 [53 (1); 81 (2)]

Núm.	mROC de combinación	Combi.	ID	Curva ROC de AUC	IC 95 %	Umbral	Prec. (%)	Sp (%)	Se (%)	VPP (%)	VPN (%)	Población
	KCNJ15_ACD MDM2_KST GAB2_JK PRKCB_JM PRKCB_ADJK CAMK1 D_D GAB2_Site78330939_H KCNJ15_Site38287526_A CAMK1D_R GAB2_ABDE KCNJ15_C KCNJ15_Site38287531 CAMK1D_AF GAB2_H GAB2_DH PRKCB B GAB2 Site78331054 LYN_Site56012495 MDM2_KSX LYN_Site56012468_C LYN_Site56012521 KCNJ15_AC KCNJ15_ABC PRKCB_DJM											
8	GAB2_CH CAMK1D_DS KCNJ15_Site38287525_B GAB2_CE LYN_Site56012491 LYN_AC LYN_AF KCNJ15_Site38287530_D MDM2_SVW MDM2_Site68821643 KCNJ15_ACE PRKCB_Site23920893_G KCNJ15_Site38287529_C KCNJ15_CE PDE8A_E KCNJ15_ACD GAB2_JK PRKCB_JM MDM2_KST PRKCB_ADJKCAMK1D_D GAB2_Site78330939_H KCNJ15_Site38287526_A CAMK1D_R GAB2_ABDE KCNJ15_C CAMK1D_AF GAB2_H GAB2_DH PRKCB_B GAB2_Site78331054 LYN_Site56012495 MDM2_KSX MDM2_BG LYN_Site56012521 KCNJ15_AC KCNJ15_ABC PRKCB_DJM PDE8A Site85096729 F	C39	3374	0,943	[0,908; 0,979]	-0,2029	88,1	96,2	82,7	97,1	78,5	N= 134 [53 (1); 81 (2)]
9	GAB2_CH CAMK1D_DS KCNJ15_Site38287525_B GAB2_CE LYN_Site56012491 LYN_AC KCNJ15_Site38287530_D MDM2_SVW PRKCB Site23920893 G	C9	8	0,833	[0,764; 0,901]	-0,1683	78,4	79,3	77,8	85,1	70,0	N= 134 [53 (1); 81 (2)]
10	GAB2_CH CAMK1 D_DS GAB2_CE LYN_Site56012491 LYN_AC LYN_AF MDM2 Site68821643 PRKCB Site23920893 G	C8	62	0,806	[0,731; 0,881]	-0,2557	79,9	71,7	85,2	82,1	76,0	N= 134 [53 (1); 81 (2)]

Núm.	mROC de combinación	Combi.	ID	Curva ROC de AUC	IC 95 %	Umbral	Prec. (%)	Sp (%)	Se (%)	VPP (%)	VPN (%)	Población
11	GAB2_CH CAMK1D_DS KCNJ15_Site38287525_B GAB2_CE LYN_Site56012491 MDM2_SVW MDM2_Site68821643 PRKCB_Site23920893_G	C8	20	0,831	[0,763; 0,899]	-0,1107	78,4	83,0	75,3	87,1	68,8	N= 134 [53 (1); 81 (2)]

Los datos obtenidos con ambos métodos dieron, por lo tanto, rendimientos diagnósticos excelentes para el diagnóstico específico por sexo de DEP, con AUC superiores a 0,850, especificidades y sensibilidades superiores al 80 % tanto para hombres como para mujeres, lo que indica un ensayo diagnóstico que puede usarse en la clínica.

Dado que se identificaron biomarcadores de edición de ARN para el diagnóstico del DEP, se buscó investigar si las isoformas no editadas de las dianas validadas podían ser útiles para discriminar las poblaciones de control y de pacientes y podían usarse en combinación con biomarcadores de edición de ARN. A continuación, se calculó la ROC de AUC para cada diana no editada, junto con la significación para cada isoforma no editada. El análisis de las isoformas no editadas de PRKCB, MDM2, IFNAR1 y PDE8A en la cohorte de validación (n = 411 sujetos) arrojó unas AUC superiores a 0,5; las isoformas no editadas tanto de PRKCB como de MDM2 fueron significativamente diferentes entre el control y DEP ($p < 0,05$, tabla 9).

Tabla 9: Rendimientos diagnósticos obtenidos con isoformas no editadas en la cohorte de validación analizada con secuenciación multiplexada

Diana	Valor de p	Curva ROC de AUC	% sin editar	Población
PRKCB sin editar	0,0001	0,622	43,6	N = 411 [143 (0); 268 (1)]
MDM2 sin editar	0,0132	0,573	33,7	N = 411 [143 (0); 268 (1)]
IFNAR1 sin editar	0,0785	0,553	74,4	N = 411 [143 (0); 268 (1)]
PDE8A sin editar	0,0915	0,551	78,1	N = 411 [143 (0); 268 (1)]

Además, cuando las isoformas no editadas se combinaron con una isoforma de las mismas dianas, se observaron diferencias significativas entre los pacientes de control sanos y con DEP (tabla 19), lo que sugiere que las isoformas no editadas también pueden usarse en el ensayo de diagnóstico.

Tabla 19: Rendimientos diagnósticos obtenidos con isoformas no editadas combinadas con dianas editadas en la cohorte de validación analizada con secuenciación multiplexada

Diana	Relación	Valor de p	Curva ROC de AUC	Población
CAMK1D	BGNS / sin editar	0,0002	0,615	N = 411 [143 (0); 268 (1)]
CAMK1D	ABCDEFRS / sin editar	0,0002	0,613	N= 411 [143 (0); 268 (1)]
CAMK1D	ABCEFGFR / sin editar	0,0007	0,599	N= 411 [143 (0); 268 (1)]
CAMK1D	BCM / sin editar	0,0008	0,599	N= 411 [143 (0); 268 (1)]
CAMK1D	BEIRS / sin editar	0,0009	0,599	N= 411 [143 (0); 268 (1)]
CAMK1D	ABDG / sin editar	0,0026	0,592	N= 411 [143 (0); 268 (1)]
GAB2	AEFHI / sin editar	0,0005	0,603	N= 411 [143 (0); 268 (1)]
IFNAR1	ABCEFGILNP / sin editar	0,001	0,601	N= 411 [143 (0); 268 (1)]
MDM2	KOSX / sin editar	0,0007	0,602	N= 411 [143 (0); 268 (1)]
MDM2	AJSWX / sin editar	0,0003	0,609	N= 411 [143 (0); 268 (1)]
MDM2	OP / sin editar	0,0007	0,602	N= 411 [143 (0); 268 (1)]
PRKCB	BDFIJK / sin editar	0,0006	0,601	N= 411 [143 (0); 268 (1)]

Diana	Relación	Valor de <i>p</i>	Curva ROC de AUC	Población
PRKCB	DJP / sin editar	0,0008	0,600	N= 411 [143 (0); 268 (1)]
PRKCB	BN / sin editar	0,0009	0,600	N= 411 [143 (0); 268 (1)]
PRKCB	CDFIJK / sin editar	0,0013	0,593	N= 411 [143 (0); 268 (1)]
PRKCB	BJKM / sin editar	0,0024	0,591	N= 411 [143 (0); 268 (1)]
PRKCB	AGHIJKLM / sin editar	0,0031	0,594	N= 411 [143 (0); 268 (1)]
PRKCB	DJ / sin editar	0,0003	0,612	N= 411 [143 (0); 268 (1)]

Conclusión

Según el conocimiento de los presentes inventores, esta investigación es el primer estudio que investiga modificaciones en la edición del ARN para usarlas como biomarcadores sanguíneos para diagnosticar trastornos del estado de ánimo. Se realizó un análisis editómico de la DEP en pacientes reclutados en los pacientes ambulatorios del Departamento de Psiquiatría de Emergencia y Cuidados Intermedios (CHRU de Montpellier). Se confirmó la presencia de diagnósticos psiquiátricos por psiquiatras capacitados, quienes emplearon la escala de clasificación de la depresión de Montgomery-Åsberg (MADRS) en francés y el Inventario de sintomatología depresiva de 30 puntos, calificado por médicos (IDS-C30) para puntuar la depresión. Tras un estudio de Sec-ARN, se buscó la asociación con la depresión en el editoma de los pacientes en una cohorte de descubrimiento de 64 sujetos. Se identificaron 646 variantes de edición de A a I, que representan 366 genes, que se editaron de manera diferencial entre pacientes deprimidos y controles sanos. Este conjunto de genes mostró un fuerte enriquecimiento para procesos biológicos, incluyendo múltiples categorías inmunitarias, confirmando de ese modo un fuerte componente inmunitario en la fisiopatología de la DEP⁴³. Luego, se redujo la lista de variantes a un subconjunto de 14 genes (17 variantes) usando criterios mucho más estrictos, y se validaron estos biomarcadores individualmente en una cohorte de priorización de 86 individuos. Estos genes también mostraron un fuerte enriquecimiento en los genes relacionados con la regulación del proceso del sistema inmunitario. Además, algunos estaban fuertemente relacionados con las vías implicadas en el sistema neuronal y las vías relacionadas con los receptores de neurotransmisores. Finalmente, se identificaron diferentes combinaciones de un panel de 8 marcadores y se validó esta combinación en una gran cohorte de 411 pacientes. Además, cuando se introdujo la especificidad sexual y se usó un enfoque de aprendizaje automático, el rendimiento del ensayo de diagnóstico aumentó sustancialmente. Los presentes hallazgos identificaron combinaciones de biomarcadores con un alto rendimiento diagnóstico, adecuadas para ser clínicamente útiles para diagnosticar objetivamente trastornos depresivos a partir de muestras de sangre.

Anteriormente se identificaron alteraciones en la edición del ARNm del receptor de serotonina 5HT2CR y PDE8A en muestras cerebrales de fallecidos suicidas mediante polimorfismo de conformación monocatenaria 5 por electroforesis capilar (CE-SSCP)^{5,44}. Recientemente, se demostró la alteración de la edición del ARN en pacientes psiquiátricos con virus de la hepatitis C (VHC) crónica sometidos a una terapia combinada antiviral con IFN-α y ribavirina, en quienes la depresión que surge con el tratamiento es una complicación común²². En este último estudio, las mediciones de los niveles de edición se basaron en secuenciación dirigida de próxima generación (NGS), que proporciona tanto un alto rendimiento como una profundidad de base suficiente para permitir una cuantificación fiable para todos los sitios estudiados.

Referencias

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 5ª ed. Washington, DC. 2013.
2. Dunlop K, Talishinsky A, Liston C. Intrinsic Brain Network Biomarkers of Antidepressant Response: a Review. Current psychiatry reports. 13 de agosto de 2019;21(9):87.
3. Jentsch MC, Van Buel EM, Bosker FJ, et al. Biomarker approaches in major depressive disorder evaluated in the context of current hypotheses. Biomarkers in medicine. 2015;9(3):277-297.
4. Smith KM, Renshaw PF, Bilello J. The diagnosis of depression: current and emerging methods. Comprehensive psychiatry. enero de 2013;54(1):1-6.
5. Chimienti F, Cavarec L, Vincent L, y col. Brain region-specific alterations of RNA editing in PDE8A mRNA in suicide decedents. Translational psychiatry. 15 de febrero de 2019;9(1):91.
6. Dobin A, Gingeras TR. Optimizing RNA-Seq Mapping with STAR. Methods in molecular biology. 2016;1415:245-262.

7. DePristo MA, Banks E, Poplin R, y col. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature genetics*. mayo de 2011;43(5):491-498.
8. John D, Weirick T, Dimmeler S, Uchida S. RNAEditor: easy detection of RNA editing events and the introduction of editing islands. *Briefings in bioinformatics*. 1 de noviembre de 2017;18(6):993-1001.
9. Yang H, Wang K. Genomic variant annotation and prioritization with ANNOVAR and wANNOVAR. *Nature protocols*. octubre de 2015;10(10):1556-1566.
10. Tempel S. Using and understanding RepeatMasker. *Methods in molecular biology*. 2012;859:29-51.
11. Kasprzyk A, Keefe D, Smedley D, y col. EnsMart: a generic system for fast and flexible access to biological data. *Genome research*. enero de 2004; 14(1): 160-169.
12. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nature methods*. 4 de marzo de 2012;9(4):357-359.
13. Li H, Handsaker B, Wysoker A, y col. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 15 de agosto de 2009;25(16):2078-2079.
14. Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*. 1 de noviembre de 2011;27(21):2987-2993.
15. Gentleman RC, Carey VJ, Bates DM, y col. Bioconductor: open software development for computational biology and bioinformatics. *Genome biology*. 2004;5(10):R80.
16. Linden A. Measuring diagnostic and predictive accuracy in disease management: an introduction to receiver operating characteristic (ROC) analysis. *Journal of evaluation in clinical practice*. abril de 2006; 12(2): 132-139.
17. Kramar A, Faraggi D, Fortune A, Reiser B. mROC: a computer program for combining tumour markers in predicting disease states. *Computer methods and programs in biomedicine*. septiembre de 2001;66(2-3):199-207.
18. Breiman L. Random Forests. *Machine Learning*. Kluwer Academic Publishers. 2001;45:5-32.
19. Consortium GO. Gene Ontology Consortium: going forward. *Nucleic acids research*. enero de 2015;43 (versión de base de datos):D1049-1056.
20. Raudvere U, Kolberg L, Kuzmin I, y col. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic acids research*. Jul 2 2019;47(W1):W191-W198.
21. Pomaznoy M, Ha B, Peters B. GOnet: a tool for interactive Gene Ontology analysis. *BMC bioinformatics*. 7 de diciembre de 2018;19(1):470.
22. Salvétat N, Van der Laan S, Vire B, y col. RNA editing blood biomarkers for predicting mood alterations in HCV patients. *Journal of neurovirology*. 22 de julio de 2019.
23. Szklarczyk D, Gable AL, Lyon D, y col. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic acids research*. 8 de enero de 2019;47(D1):D607-D613.
24. Shelton RC. The molecular neurobiology of depression. *The Psychiatric clinics of North America*. marzo de 2007;30(1):1-11.
25. Costas J, Gratacos M, Escaramis G, y col. Association study of 44 candidate genes with depressive and anxiety symptoms in post-partum women. *Journal of psychiatric research*. agosto de 2010;44(11):717-724.
26. Colledge M, Snyder EM, Crozier RA, y col. Ubiquitination regulates PSD-95 degradation and AMPA receptor surface expression. *Neuron*. 30 de octubre de 2003;40(3):595-607.
27. Liu B, Mink S, Wong KA, y col. PIAS1 selectively inhibits interferon-inducible genes and is important in innate immunity. *Nature immunology*. septiembre de 2004;5(9):891-898.
28. Murakami Y, Ishibashi T, Tomita E, y col. Depressive symptoms as a side effect of Interferon-alpha therapy induced by induction of indoleamine 2,3-dioxygenase 1. *Scientific reports*. 20 de julio de 2016;6:29920.

29. Brodie EJ, Infantino S, Low MSY, Tarlinton DM. Lyn, Lupus, and (B) Lymphocytes, a Lesson on the Critical Balance of Kinase Signaling in Immunity. *Frontiers in immunology*. 2018;9:401.
- 5 30. Hayashi T, Umemori H, Mishina M, Yamamoto T. The AMPA receptor interacts with and signals through the protein tyrosine kinase Lyn. *Nature*. 7 de enero de 1999;397(6714):72-76.
31. Moura-Alves P, Fae K, Houthuys E, y col. AhR sensing of bacterial pigments regulates antibacterial defence. *Nature*. 28 de agosto de 2014;512(7515):387-392.
- 10 32. Song MS, Song SJ, Kim SY, Oh HJ, Lim DS. The tumour suppressor RASSF1A promotes MDM2 self-ubiquitination by disrupting the MDM2-DAXX-HAUSP complex. *The EMBO journal*. 9 de julio de 2008;27(13):1863-1874.
- 15 33. Wang H, Nestor CE, Benson M, Zhang H. GAB2 regulates type 2 T helper cell differentiation in humans. *Cytokine*. agosto de 2017;96:234-237.
34. Verploegen S, Lammers JW, Koenderman L, Coffey PJ. Identification and characterization of CKLiK, a novel granulocyte Ca(++)/calmodulin-dependent kinase. *Sangre*. 1 de noviembre de 2000;96(9):3215-3223.
- 20 35. Lin E, Kuo PH, Liu YL, Yu YW, Yang AC, Tsai SJ. A Deep Learning Approach for Predicting Antidepressant Response in Major Depression Using Clinical and Genetic Biomarkers. *Frontiers in psychiatry*. 2018;9:290.
- 25 36. Zhou X, Chen Y, Mok KY, y col. Identification of genetic risk factors in the Chinese population implicates a role of immune system in Alzheimer's disease pathogenesis. *Proceedings of the National Academy of Sciences of the United States of America*. 20 de febrero de 2018; 115(8): 1697-1706.
37. Beurel E, Lowell JA. Th17 cells in depression. *Brain, behavior, and immunity*. marzo de 2018;69:28-34.
38. Fischer EH, Charbonneau H, Tonks NK. Protein tyrosine phosphatases: a diverse family of intracellular and transmembrane enzymes. *Science*. 26 de julio de 1991;253(5018):401-406.
39. Irie-Sasaki J, Sasaki T, Matsumoto W, y col. CD45 is a JAK phosphatase and negatively regulates cytokine receptor signalling. *Nature*. 18 de enero de 2001;409(6818):349-354.
- 35 40. Basterzi AD, Yazici K, Buturak V, y col. Effects of venlafaxine and fluoxetine on lymphocyte subsets in patients with major depressive disorder: a flow cytometric analysis. *Progress in neuro-psychopharmacology & biological psychiatry*. 1 de febrero de 2010;34(1):70-75.
41. Aston C, Jiang L, Sokolov BP. Transcriptional profiling reveals evidence for signaling and oligodendroglial abnormalities in the temporal cortex from patients with major depressive disorder. *Molecular psychiatry*. marzo de 2005;10(3):309-322.
- 40 42. Seney ML, Huo Z, Cahill K, y col. Opposite Molecular Signatures of Depression in Men and Women. *Biological psychiatry*. 1 de julio de 2018;84(1):18-27.
- 45 43. Wohleb ES, Franklin T, Iwata M, Duman RS. Integrating neuroimmune systems in the neurobiology of depression. *Nature reviews. Neuroscience*. agosto de 2016; 17(8):497-511.
44. Weissmann D, van der Laan S, Underwood MD, et al. Region-specific alterations of A-to-I RNA editing of serotonin 2c receptor in the cortex of suicides with major depression. *Translational psychiatry*. 30 de agosto de 2016;6(8):e878.
- 50

REIVINDICACIONES

1. Un método *in vitro* para el diagnóstico de trastornos de depresión, en un paciente humano a partir de una muestra biológica de dicho paciente, comprendiendo dicho método:

a) determinar, para una combinación de al menos dos biomarcadores de ARN de edición de A a I:

-para cada biomarcador seleccionado, la proporción relativa de edición de ARN en un sitio de edición dado para al menos uno o una combinación de sitios que se pueden editar en el transcrito de ARN de dichos al menos dos biomarcadores, y/o
-el porcentaje relativo de una isoforma o de una combinación de isoformas del transcrito de ARN de cada uno de dichos dos biomarcadores;

en donde dichos al menos dos biomarcadores de ARN de edición de A a I se seleccionan del grupo que consiste en los biomarcadores PRKCB, MDM2, PIAS1, IFNAR1, IFNAR2, LYN, AHR, RASSF1, GAB2, CAMK1D, KCNJ15, IL17RA, PTPRC y PDE8A;

b) determinar un valor de resultado obtenido para cada uno de dichos al menos dos biomarcadores y un valor de resultado final basándose en un algoritmo o ecuación que incluye el valor de resultado obtenido para cada biomarcador seleccionado;

c) determinar el valor de resultado final o la puntuación de depresión obtenida para el paciente y un valor de resultado de control obtenido para un sujeto de control sano, en donde el valor de resultado de control se determinó de una manera comparable a la del valor de resultado final; y

d) si dicho resultado final o valor de puntuación de depresión es superior a un umbral/punto de corte, clasificar a dicho paciente como positivo o con trastornos de depresión o, si dicho valor de resultado final no es superior a dicho umbral, clasificar a dicho paciente como negativo o sin trastornos de depresión.

2. El método *in vitro* para el diagnóstico de depresión según la reivindicación 1, en donde, en la etapa b), dichos al menos dos biomarcadores de ARN de edición de A a I se seleccionan del grupo que consiste en PRKCB, MDM2, IFNAR1, KCNJ15, LYN, CAMK1D, GAB2 y PDE8A.

3. El método *in vitro* para el diagnóstico de trastornos de depresión, preferiblemente DEP, según una de las reivindicaciones 1 y 2, en donde, en la etapa b), el valor de resultado o puntuación de depresión se calcula mediante un algoritmo que implementa un método multivariante que incluye:

- programa mROC, en particular para identificar la combinación lineal, que maximiza la curva ROC de AUC (área bajo la curva) y en donde se proporciona la ecuación para la combinación respectiva y se puede usar como un nuevo marcador virtual Z, de la siguiente manera:

$$Z = a.(\text{Biomarcador 1}) + b.(\text{Biomarcador 2}) + \dots i.(\text{Biomarcador i}) + \dots n.(\text{Biomarcador n})$$

donde i son coeficientes calculados y (Biomarcador i) son el nivel del biomarcador considerado; y/o
-un enfoque de bosques aleatorios (RF) aplicado para evaluar el/los sitio(s) de edición de ARN y/o combinaciones de isoformas, en particular para clasificar la importancia del/de los sitio(s) de edición de ARN y/o isoforma(s), y para combinar el/los mejor(es) sitio(s) de edición de ARN y/o isoforma(s), y/u opcionalmente

-un análisis multivariante aplicado para evaluar el/los sitio(s) de edición de ARN y/o combinaciones de isoformas para el diagnóstico, seleccionándose dicho análisis multivariante, por ejemplo, del grupo que consiste en:

-modelo de regresión logística aplicado para análisis univariante y multivariante para estimar el riesgo relativo de un paciente a diferentes valores de nivel de sitio de edición de ARN o de isoformas; y/o

-enfoque CART (árboles de clasificación y regresión) aplicado para evaluar el/los sitio(s) de edición de ARN y/o combinaciones de isoformas; y/o

-enfoque de máquina de vectores de soporte (SVM);

-enfoque de red neuronal artificial (ANN);

-enfoque de red bayesiana;

-enfoque de WKNN (k vecinos más cercanos ponderados);

-mínimos cuadrados parciales - análisis discriminante (PLS-DA);

-análisis discriminante lineal y cuadrático (LDA/ QDA), y

-cualquier otro método matemático que combina biomarcadores.

4. El método de una de las reivindicaciones 1 a 3, en donde dicho trastorno de depresión es DEP.

5. El método de una de las reivindicaciones 1 a 4, en donde dicha muestra biológica es sangre completa, suero, plasma, orina, líquido cefalorraquídeo o saliva, prefiriéndose la muestra de sangre completa.

6. El método de una de las reivindicaciones 1 a 5, en donde para cada biomarcador seleccionado, dicho valor de resultado es estadística/específicamente diferente de dicho valor de resultado de control a $p < 0,05$.
- 5 7. El método de una de las reivindicaciones 1 a 6, en donde han de satisfacerse los siguientes criterios para el biomarcador o la combinación de biomarcadores seleccionados en la etapa a):
- 10 -la cobertura > 30 ;
 -AUC $> 0,8$;
 - $0,95 > \text{múltiplo de variación} > 1,05$, y
 - $p < 0,05$.
- 15 8. El método de una de las reivindicaciones 1 a 7, en donde dicho(s) biomarcador(es) de ARN de edición de A a I seleccionado(s) comprendido(s) en la etapa a) se selecciona(n) del grupo que consiste en:
- una combinación que comprende al menos 3, 4, 5, 6, 7 u 8 de los siguientes biomarcadores: PRKCB, MDM2, IFNAR1, KCNJ15, LYN, CAMK1D, GAB2 y PDE8A, preferiblemente la combinación de los 8 biomarcadores citados.
- 20 9. El método de una de las reivindicaciones 1 a 8, en donde dicho(s) biomarcador(es) de ARN de edición de A a I seleccionado(s) comprendido(s) en la etapa a) la combinación de los siguientes 8 biomarcadores GAB2, IFNAR1, KCNJ15, LYN, MDM2, PRKCB, CAMK1D y PDE8A cuando el modelo se ajusta por edad, sexo, tratamiento psiquiátrico y adicción, o cuando el modelo se ajusta por edad y sexo y usando un algoritmo de bosques aleatorios (RF) aplicado para evaluar el/los sitio(s) de edición de ARN y/o combinaciones de isoformas, en particular para clasificar la importancia del/de los sitio(s) de edición de ARN y/o isoforma(s), y para combinar el/los mejor(es) sitio(s) de edición de ARN y/o isoforma(s), y preferiblemente en donde el cálculo del porcentaje relativo de al menos uno del sitio de edición de ARN o isoforma se basa en el sitio de edición de ARN o isoforma enumerados en las tablas 5.1 y 5.2.
- 25 10. El método de una de las reivindicaciones 1 a 9, en donde dicha combinación de al menos 2, 3, 4, 5, 6, 7 u 8 biomarcadores seleccionados comprende el cálculo del porcentaje relativo de al menos uno del sitio de edición de ARN o isoforma enumerados en la tabla 2A (sitios de edición), 2B (isoformas) o 3 (sitios de edición) para cada uno de los biomarcadores seleccionados.
- 30 11. El método de una de las reivindicaciones 2 a 10, en donde dicha combinación de al menos 2, 3, 4, 5, 6, 7 u 8 biomarcadores seleccionados se selecciona de las combinaciones de biomarcadores proporcionadas en las tablas 4 y 5, preferiblemente las combinaciones que comprenden los siguientes 8 biomarcadores PRKCB, MDM2, IFNAR1, KCNJ15, LYN, CAMK1D, GAB2 y PDE8A.
- 35 12. El método de una de las reivindicaciones 1 a 11, en donde el algoritmo o la ecuación que permite el cálculo del valor de resultado o puntuación de depresión Z se seleccionan de las ecuaciones de Z enumeradas en los ejemplos y las figuras, preferiblemente la ecuación implementa una combinación de los siguientes 8 biomarcadores PRKCB, MDM2, IFNAR1, KCNJ15, LYN, CAMK1D, GAB2 y PDE8A.
- 40 13. Un método *in vitro* para diagnosticar trastornos de depresión, preferiblemente MDD, en un paciente humano según una de las reivindicaciones 1 a 12, en donde si el paciente a analizar es una mujer, en la etapa a) los biomarcadores de dicha combinación de al menos dos biomarcadores de ARN de edición de A a I se seleccionan del grupo de biomarcadores que consiste en los biomarcadores PRKCB, MDM2, PIAS1, IFNAR1, IFNAR2, LYN, AHR, RASSF1, GAB2, CAMK1D, KCNJ15, IL17RA, PTPRC y PDE8A, preferiblemente los biomarcadores PRKCB, MDM2, IFNAR1, KCNJ15, LYN, CAMK1D, GAB2 y PDE8A, preferiblemente la combinación de los 8 biomarcadores preferidos.
- 45 50 14. Un método *in vitro* para diagnosticar trastornos de depresión, preferiblemente MDD, en un paciente humano según una de las reivindicaciones 1 a 12, en donde si el paciente a analizar es un hombre, en la etapa a) los biomarcadores de dicha combinación de al menos dos biomarcadores de ARN de edición de A a I se seleccionan del grupo de biomarcadores que consiste en los biomarcadores PRKCB, MDM2, PIAS1, IFNAR1, IFNAR2, LYN, AHR, RASSF1, GAB2, CAMK1D, KCNJ15, IL17RA, PTPRC y PDE8A, preferiblemente los biomarcadores PRKCB, MDM2, IFNAR1, KCNJ15, LYN, CAMK1D, GAB2 y PDE8A, preferiblemente la combinación de los 8 biomarcadores preferidos.
- 55 60 15. Un método *in vitro* para el diagnóstico de trastornos de depresión, preferiblemente DEP, en un paciente humano, en donde dicho método comprende el método para el diagnóstico de depresión para un paciente humano según una de las reivindicaciones 1 a 14, en donde la combinación de biomarcadores y/o la ecuación de Z asociada con dicha combinación usada para el diagnóstico de trastornos de depresión es diferente si el paciente que se va a analizar es una mujer o un hombre.
- 65

16. El método de una de las reivindicaciones 1 a 15, en donde, en la etapa a), se miden la proporción relativa de edición de ARN en una edición dada y/o el porcentaje de una isoforma mediante NGS en dicha muestra biológica.
- 5 17. El método de una de las reivindicaciones 1 a 17, en donde en la etapa a), la secuencia de amplicón/nucleica usada para la detección de los sitios de edición de ARN y/o las isoformas del transcrito de ARN de dicho biomarcador se obtiene o se puede obtener con:
- 10 -el conjunto de cebadores enumerados en la tabla 1 para cada uno de los biomarcadores seleccionados (Id. de sec. n.º 1 a 36), o
 -un conjunto o una combinación de conjuntos de cebadores que permiten obtener un(os) amplicón/amplicones que incluye(n), o idéntico(s) a, el/los amplicón/amplicones que se puede(n) obtener mediante el conjunto de cebadores enumerados en la tabla 1 (Id. de sec. n.º 1 a Id. de sec. n.º 36).
- 15 18. Un método para monitorizar el tratamiento para trastornos de depresión, preferiblemente DEP, en un sujeto humano, comprendiendo dicho método:
- 20 A) diagnosticar trastornos de depresión en dicho ser humano mediante el método según una de las reivindicaciones 1 a 17 antes del inicio del tratamiento que se desea monitorizar;
 B) repetir las etapas (a) a c) del método de una de las reivindicaciones 1 a 17, después de un período de tiempo durante el cual dicho paciente recibió tratamiento para dichos trastornos de depresión para obtener una puntuación de depresión postratamiento;
 C) comparar la puntuación de depresión postratamiento de la etapa (c) con:
- 25 -la puntuación de depresión obtenida antes del período de tiempo durante el cual dicho paciente recibe tratamiento para dichos trastornos de depresión, y
 -con la puntuación de depresión (valor final de control) para sujetos normales/sanos, y clasificar dicho tratamiento como eficaz si la puntuación de depresión postratamiento obtenida en la etapa B) está más cerca que la puntuación de depresión obtenida antes del período de tiempo durante el cual dicho paciente recibe tratamiento para trastornos de depresión de la serotonina puntuación obtenida para sujetos normales/sanos.
- 30 19. Kit para determinar si un paciente presenta un trastorno de depresión, preferiblemente MDD, comprendiendo dicho kit:
- 35 1) opcionalmente, instrucciones para aplicar el método según una de las reivindicaciones 1 a 17, con el fin de obtener el valor de resultado o la puntuación de depresión cuyo análisis determina si dicho paciente presenta un trastorno de depresión, preferiblemente DEP; y
- 40 2)
- 45 i) una combinación de al menos dos conjuntos de cebadores del grupo de pares de cebadores que se seleccionan de la Id. de sec. n.º 1 a Id. de sec. n.º 36, cuya selección depende de la combinación de biomarcadores usada para el diagnóstico de trastornos de depresión, preferiblemente DEP;
 o
 ii) una combinación de al menos dos conjuntos de cebadores que permiten obtener amplicones idénticos a los amplicones obtenidos mediante los al menos dos conjuntos de cebadores seleccionados en i).

Fig. 1A

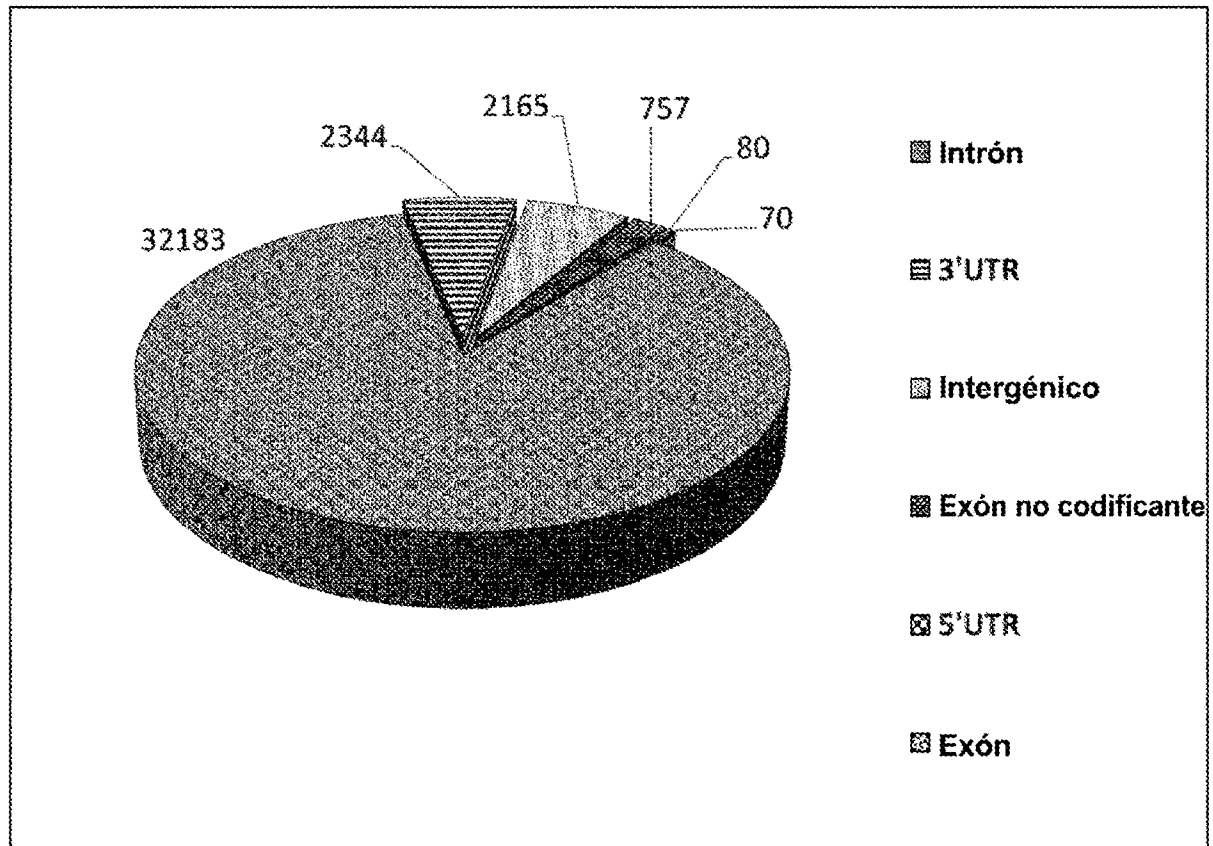


Fig. 1B

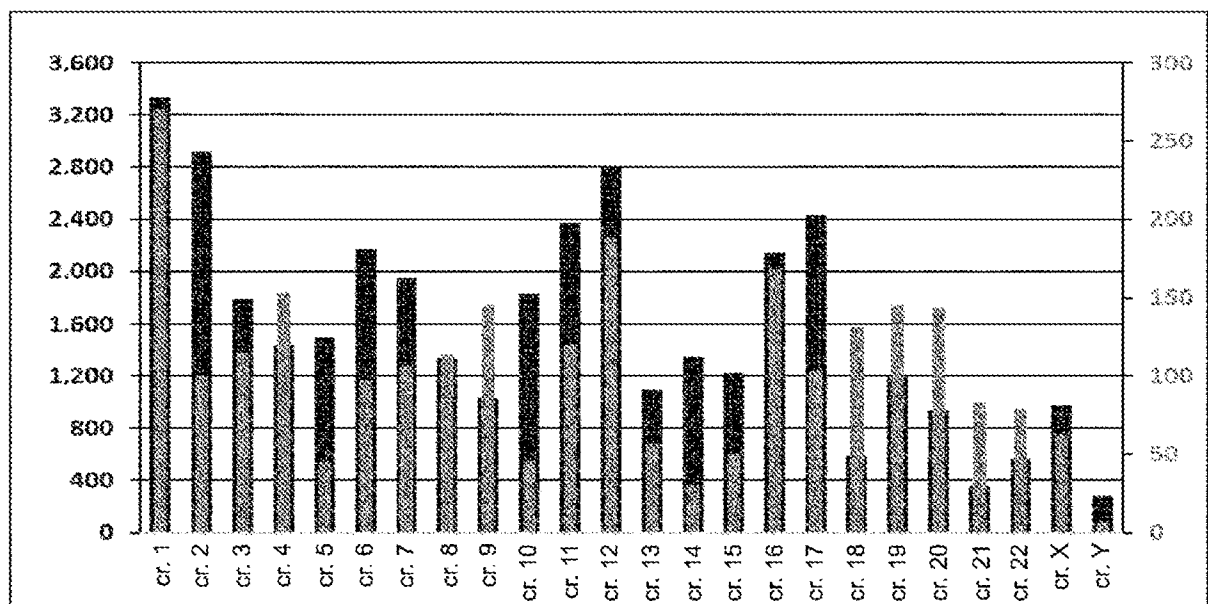


Fig. 1C

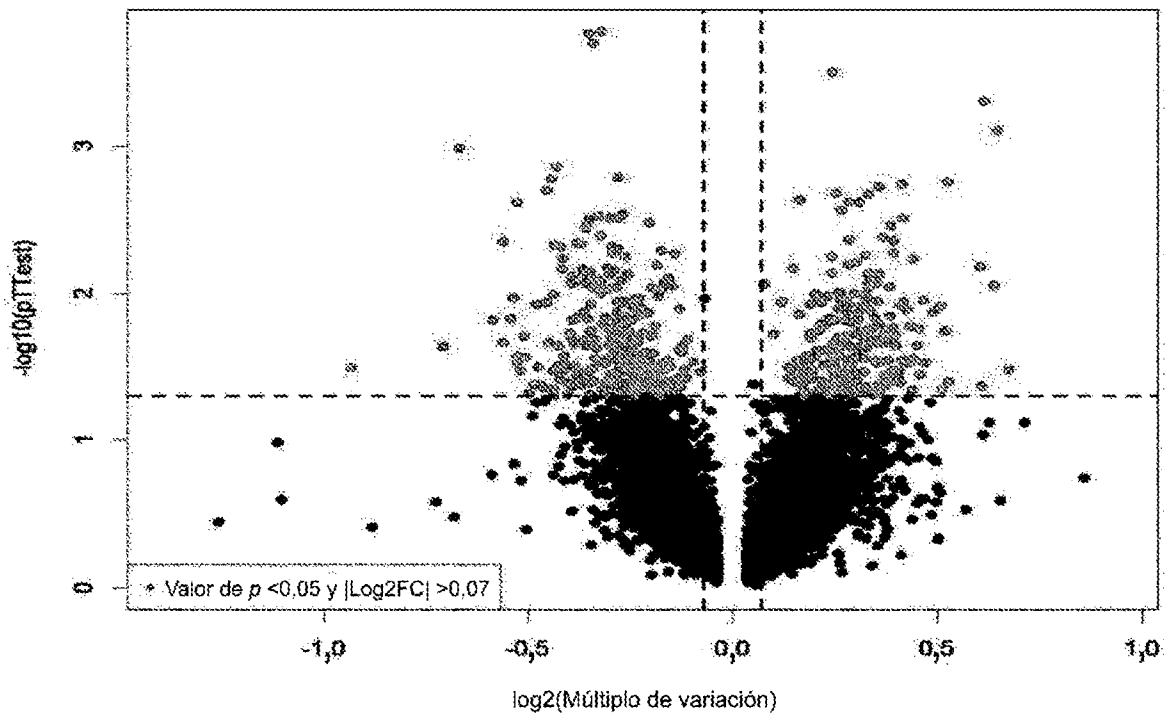


Fig. 1D

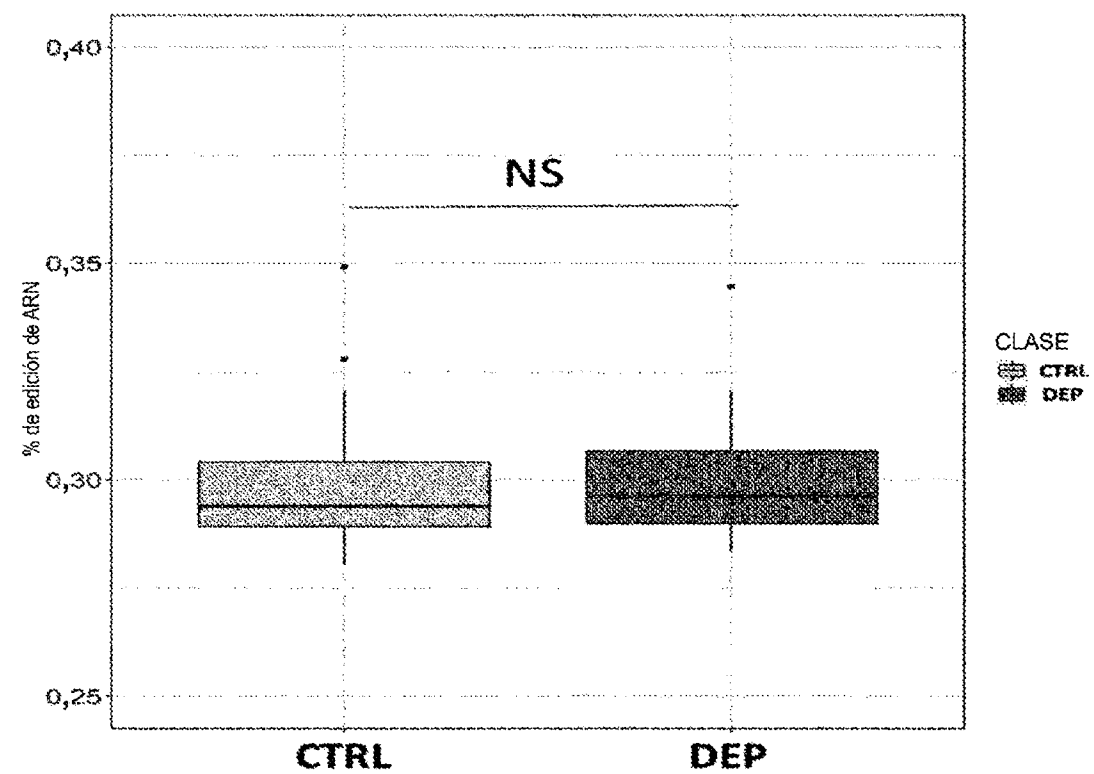


Fig. 2

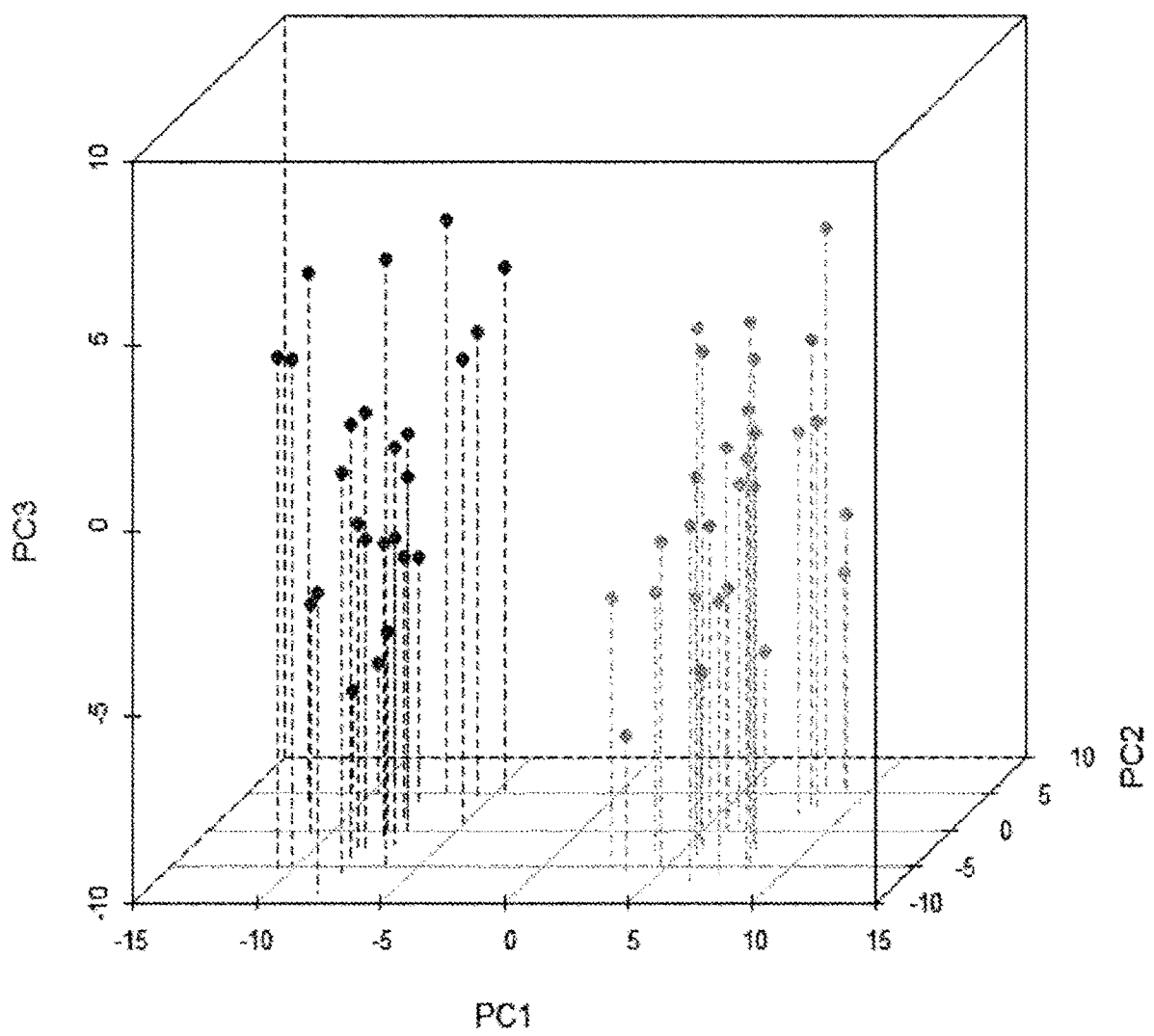


Fig. 3

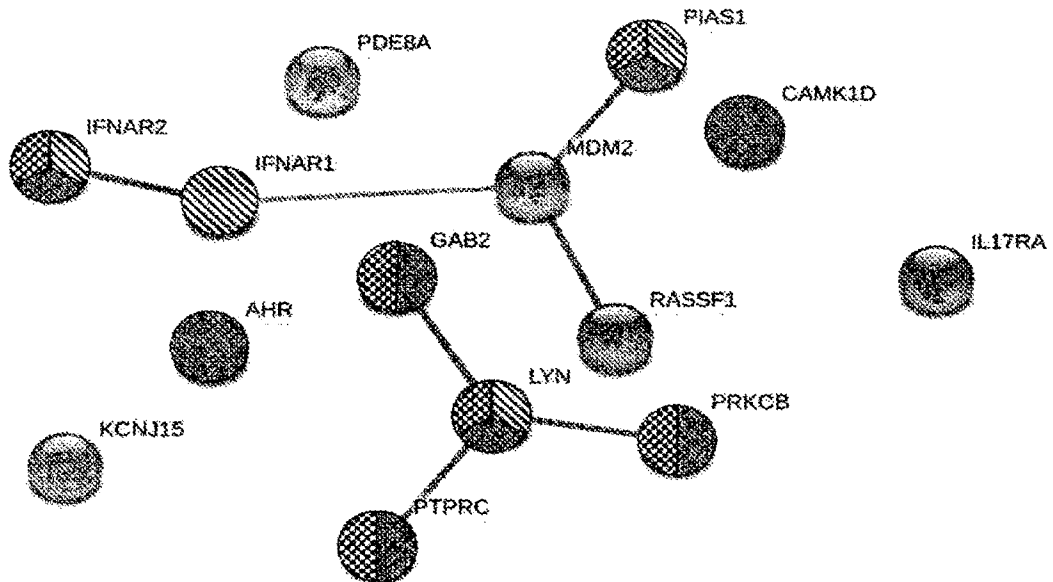


Fig. 4A

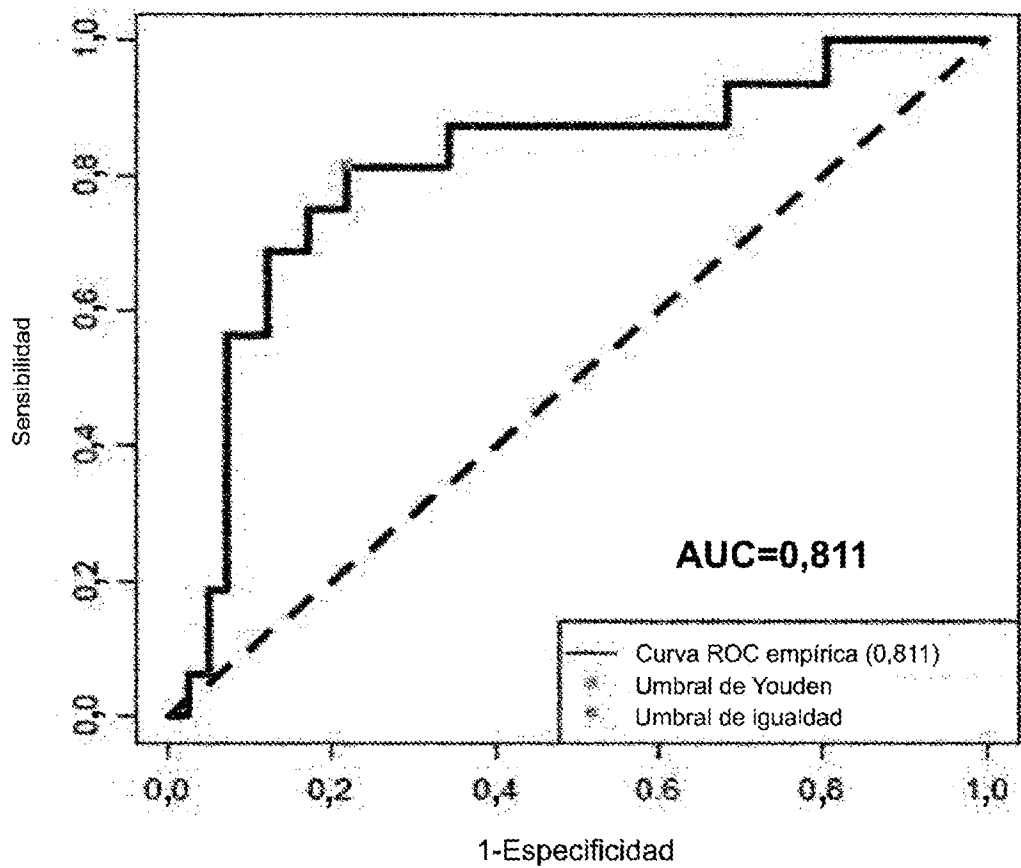


Fig. 4B

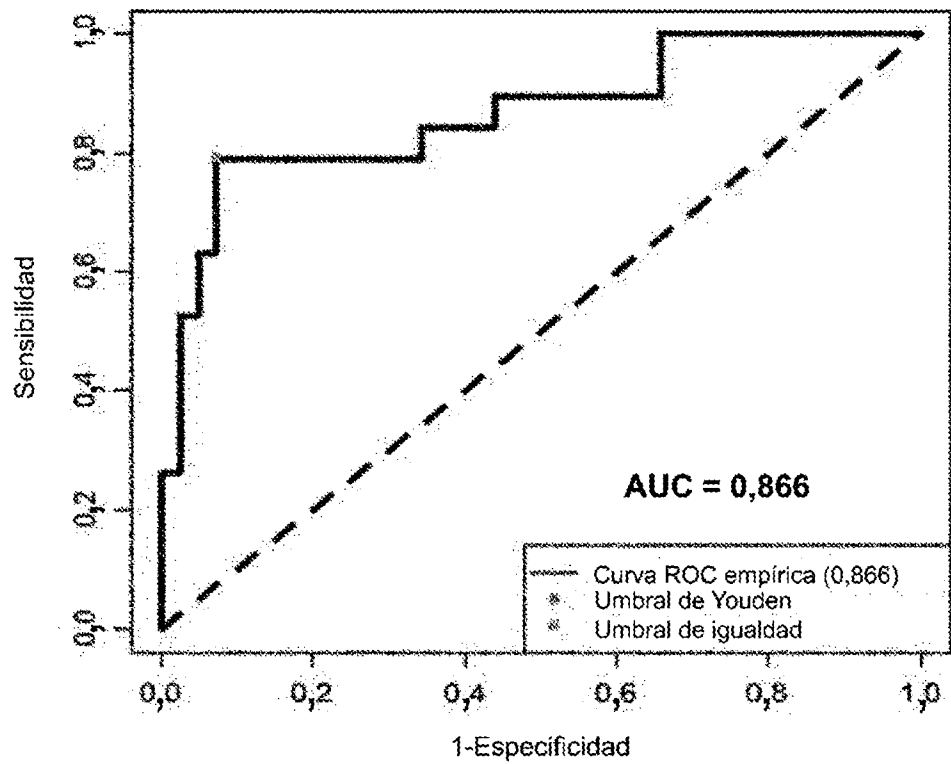


Fig. 4C

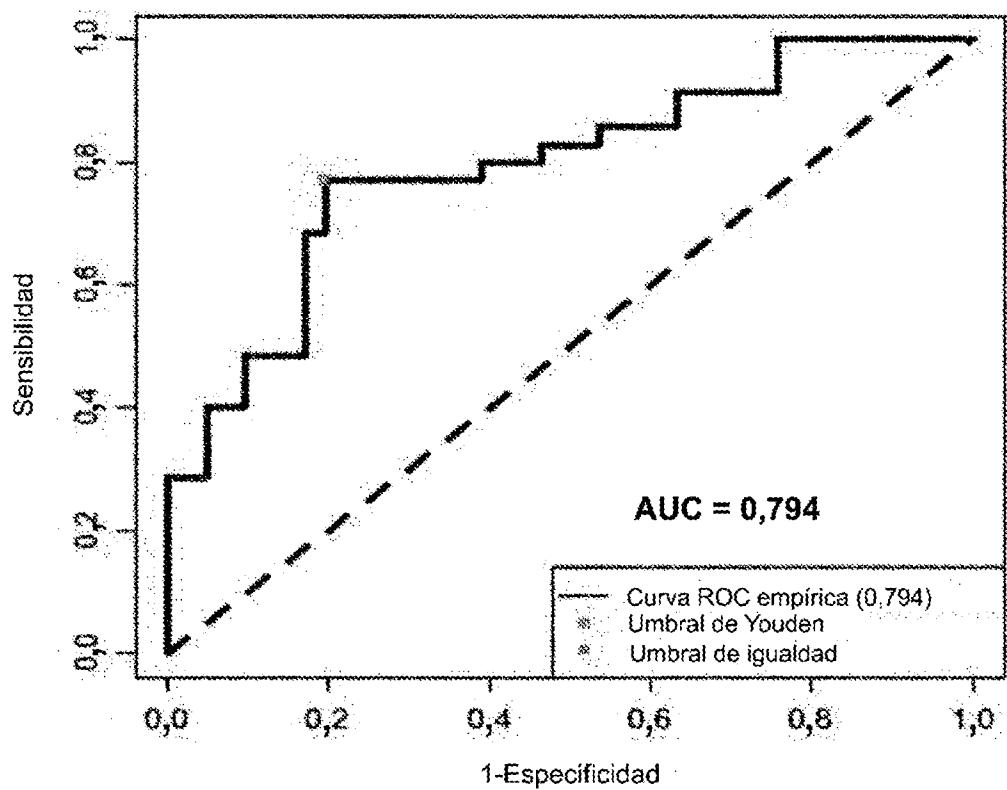


Fig. 4D

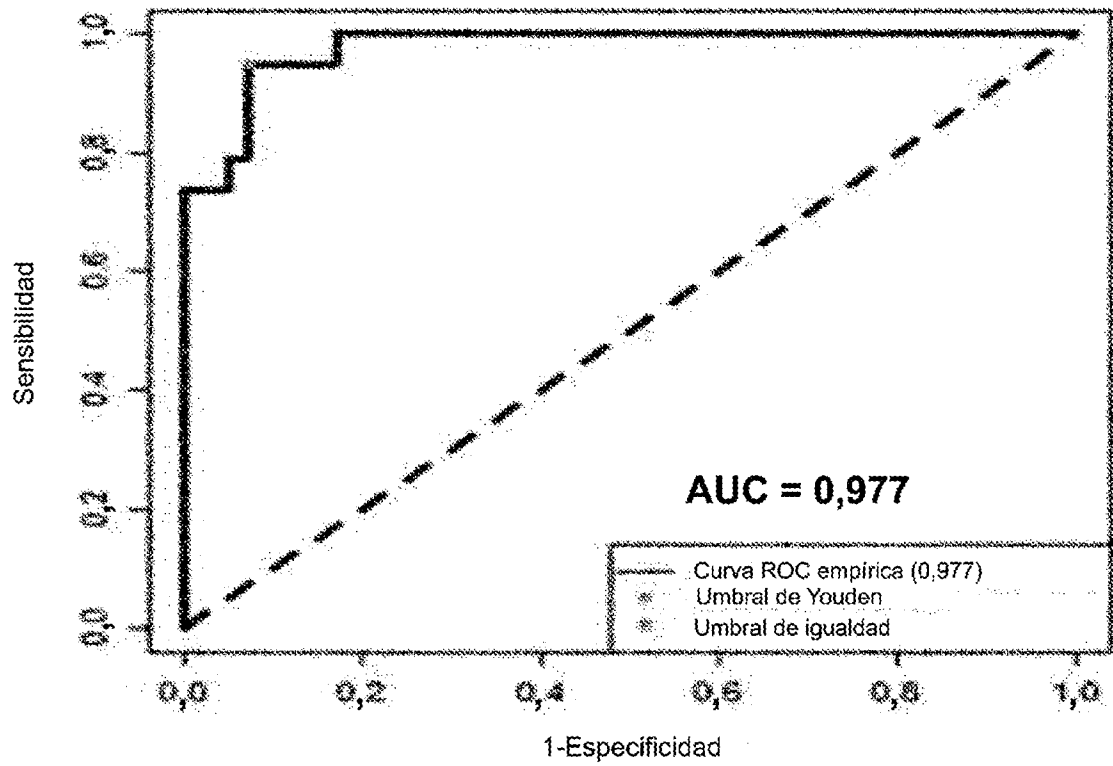


Fig. 5A

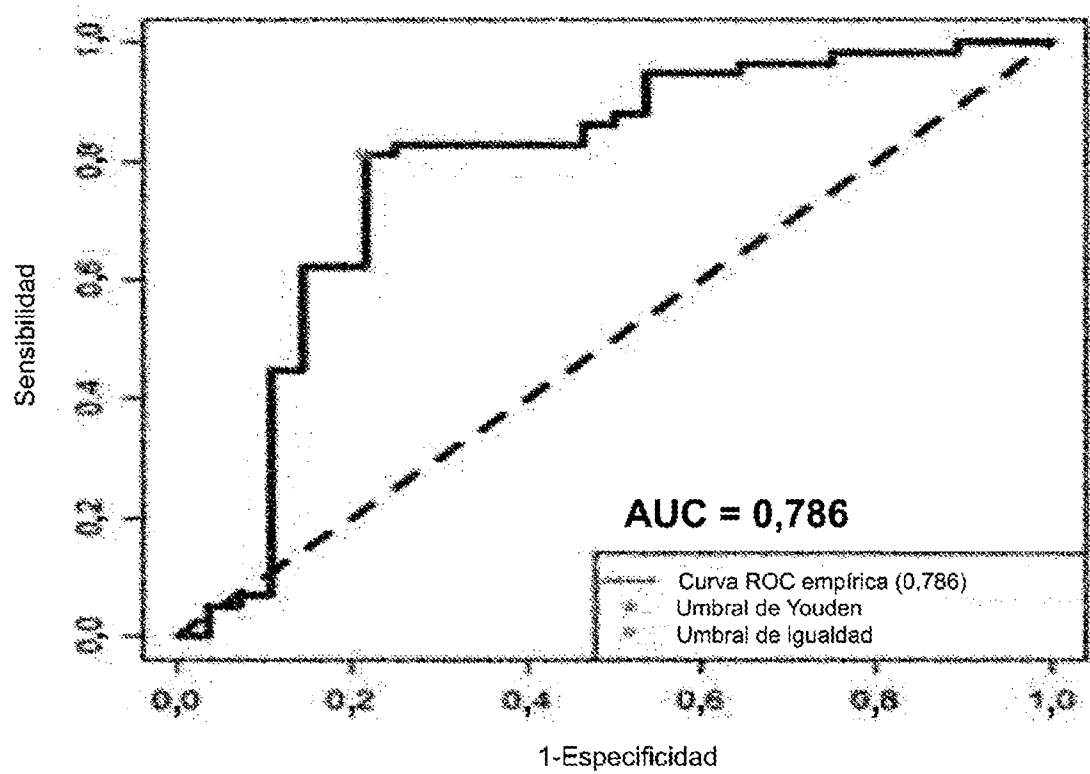


Fig. 5B

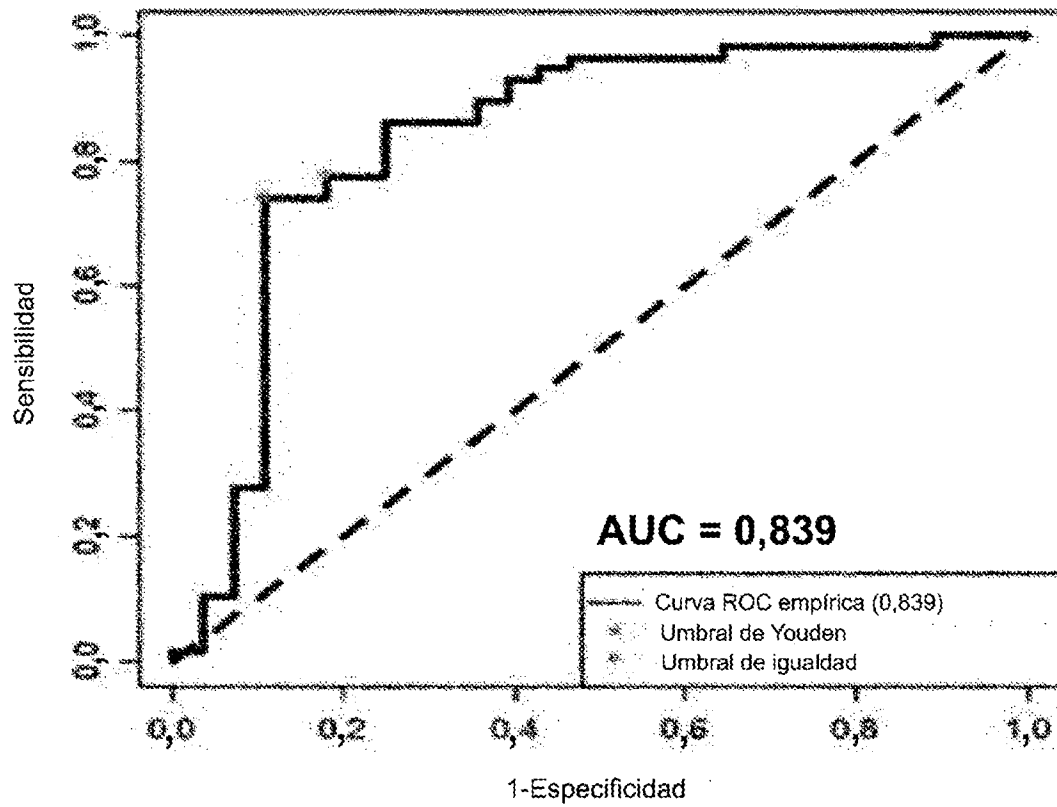


Fig. 5C

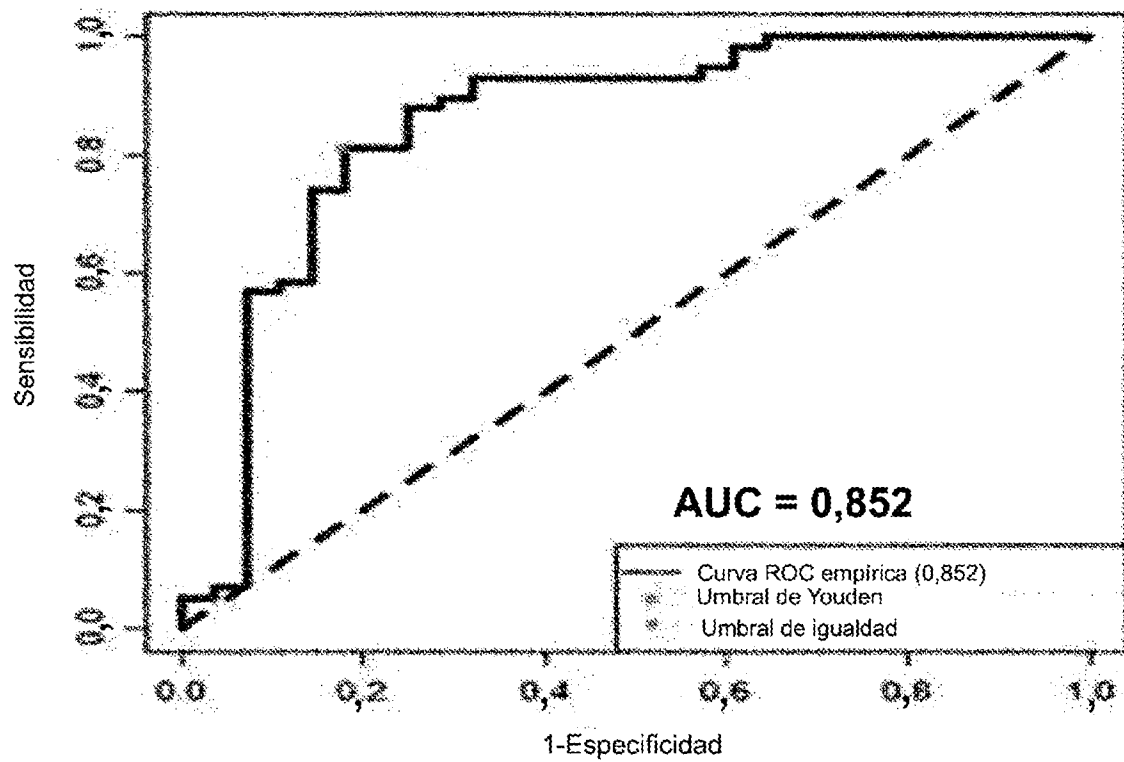


Fig. 5D

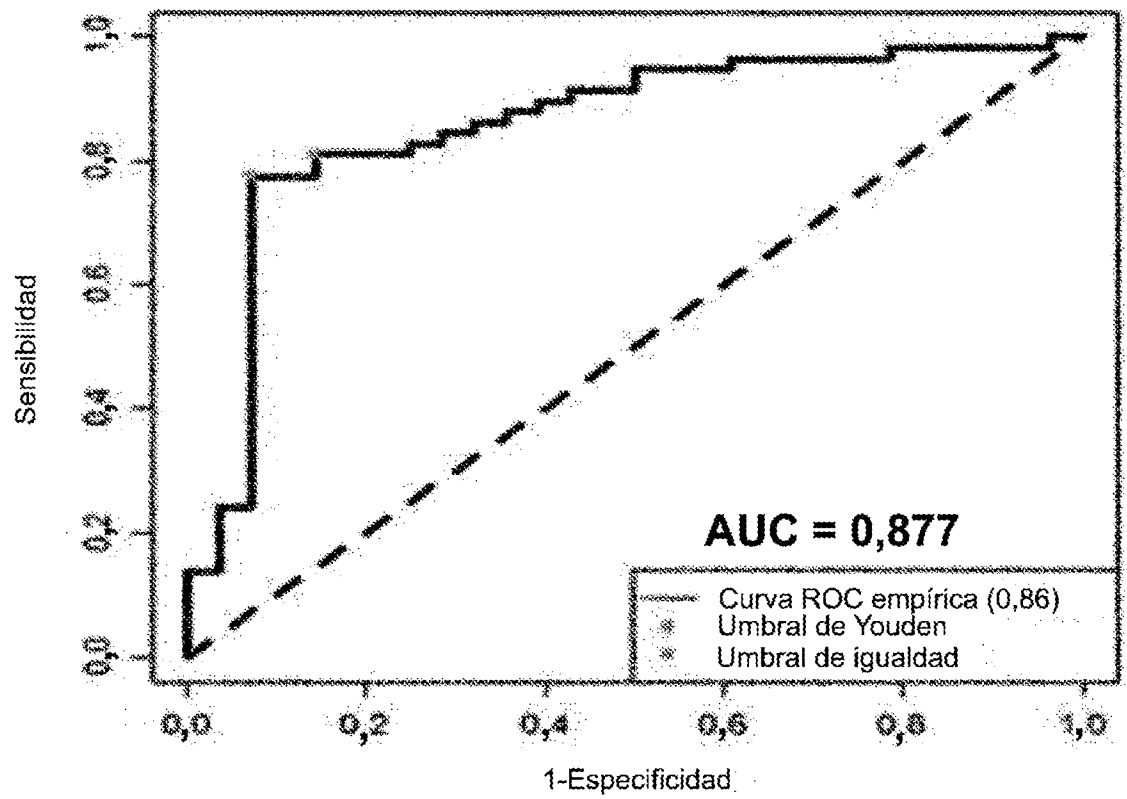


Fig. 5E

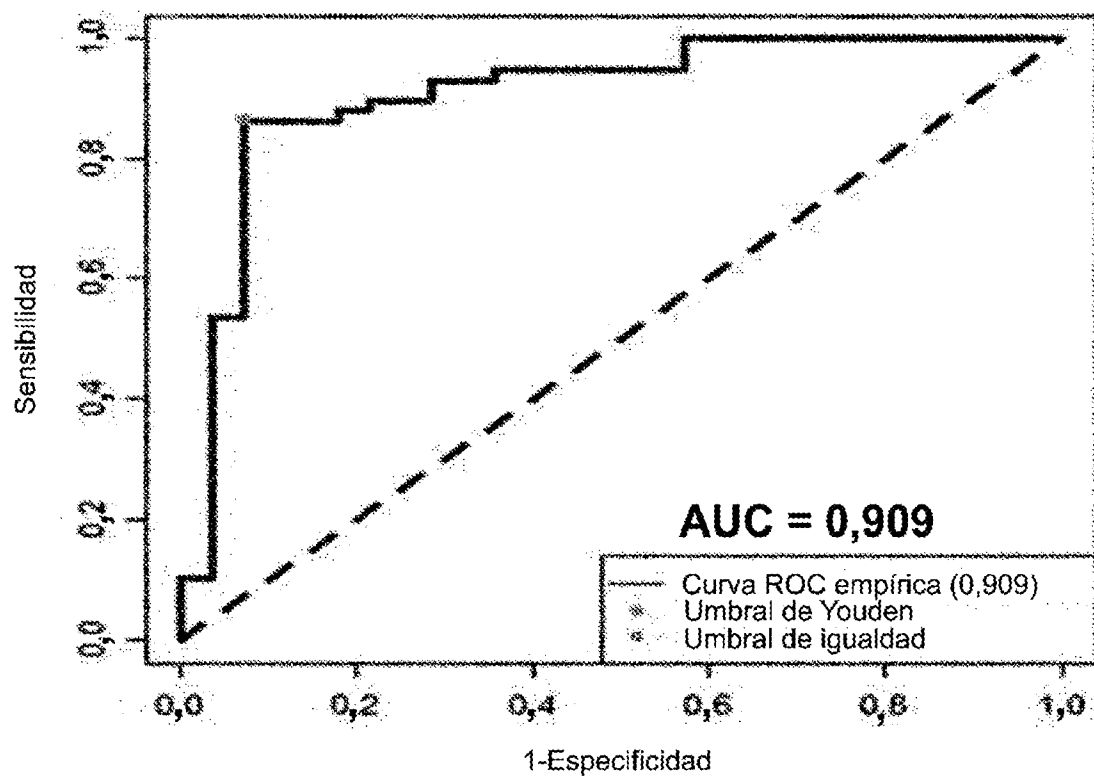


Fig. 5F

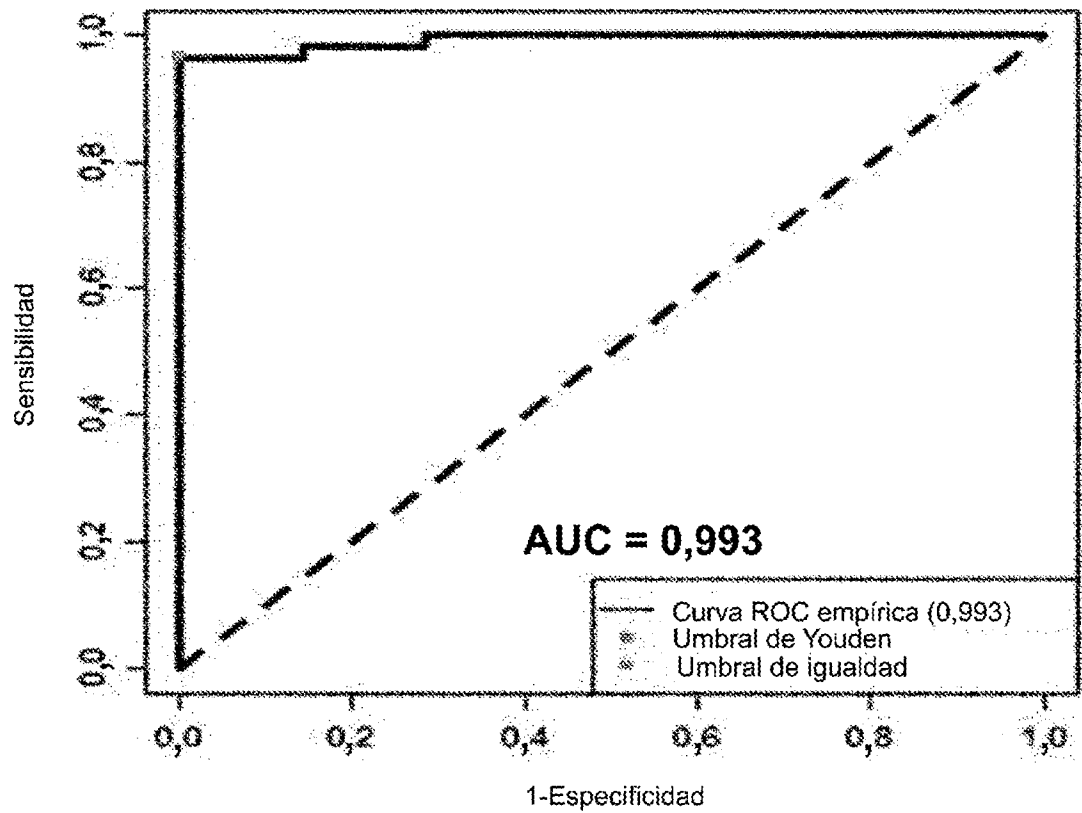


Fig. 6A

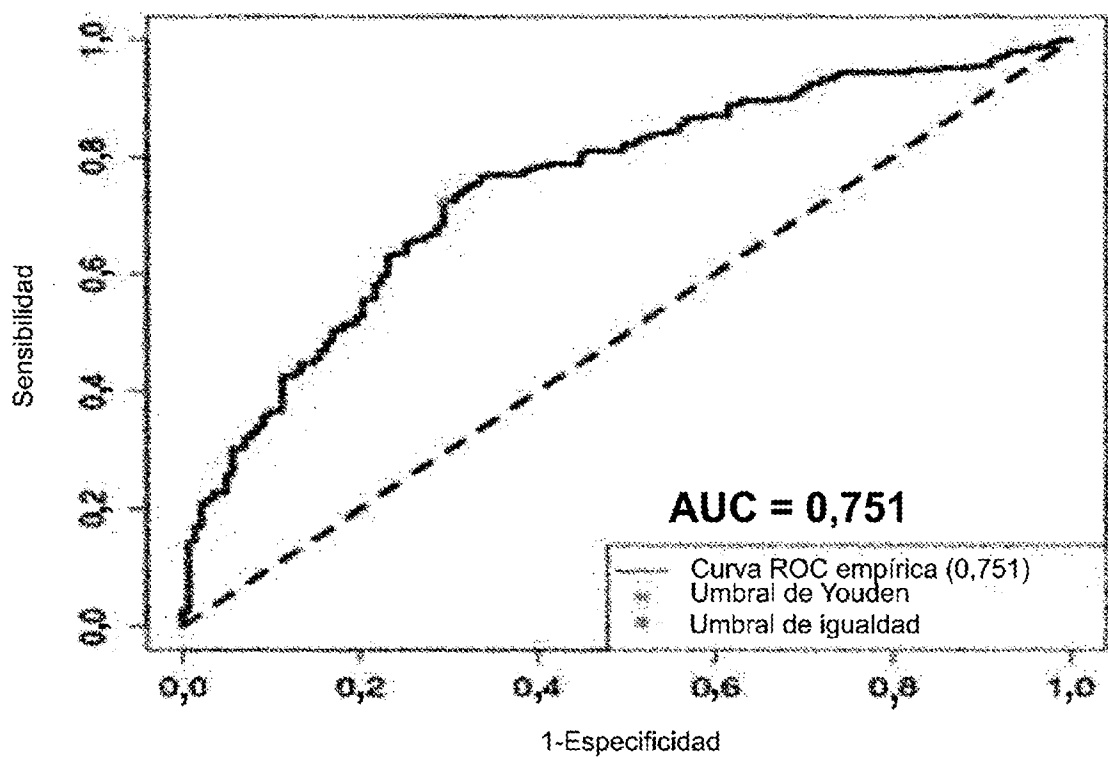


Fig. 6B

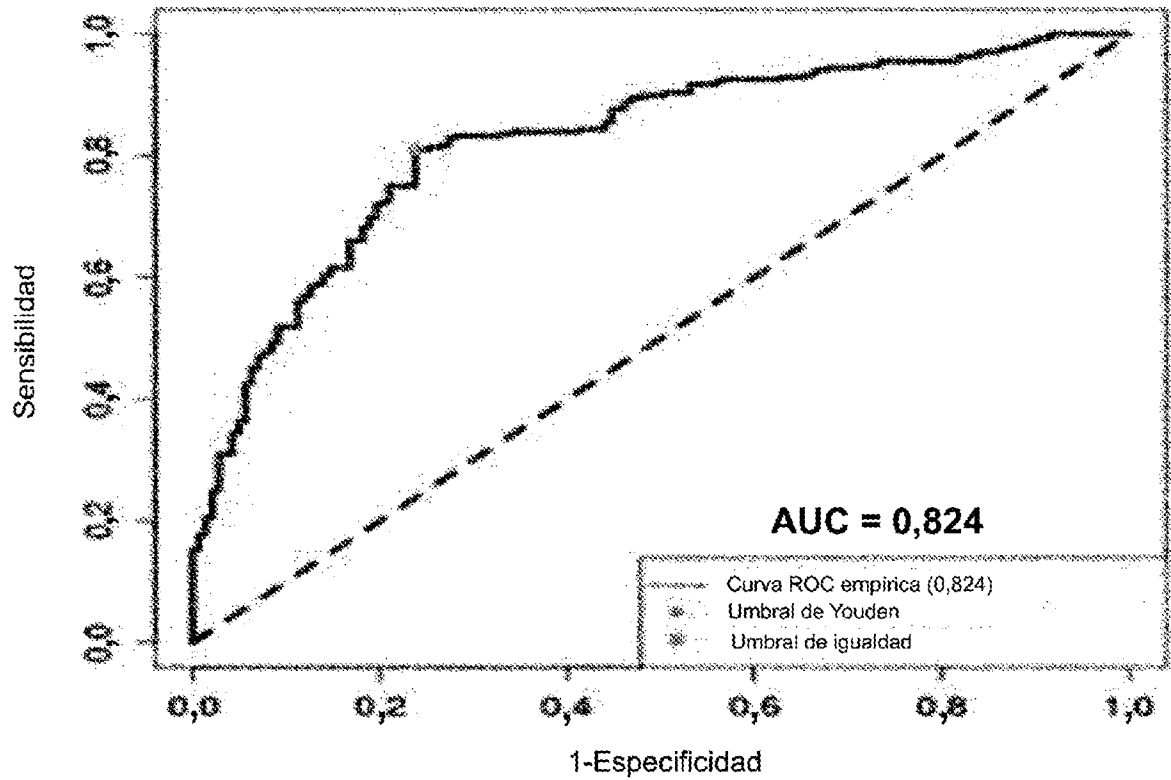


Fig. 6C

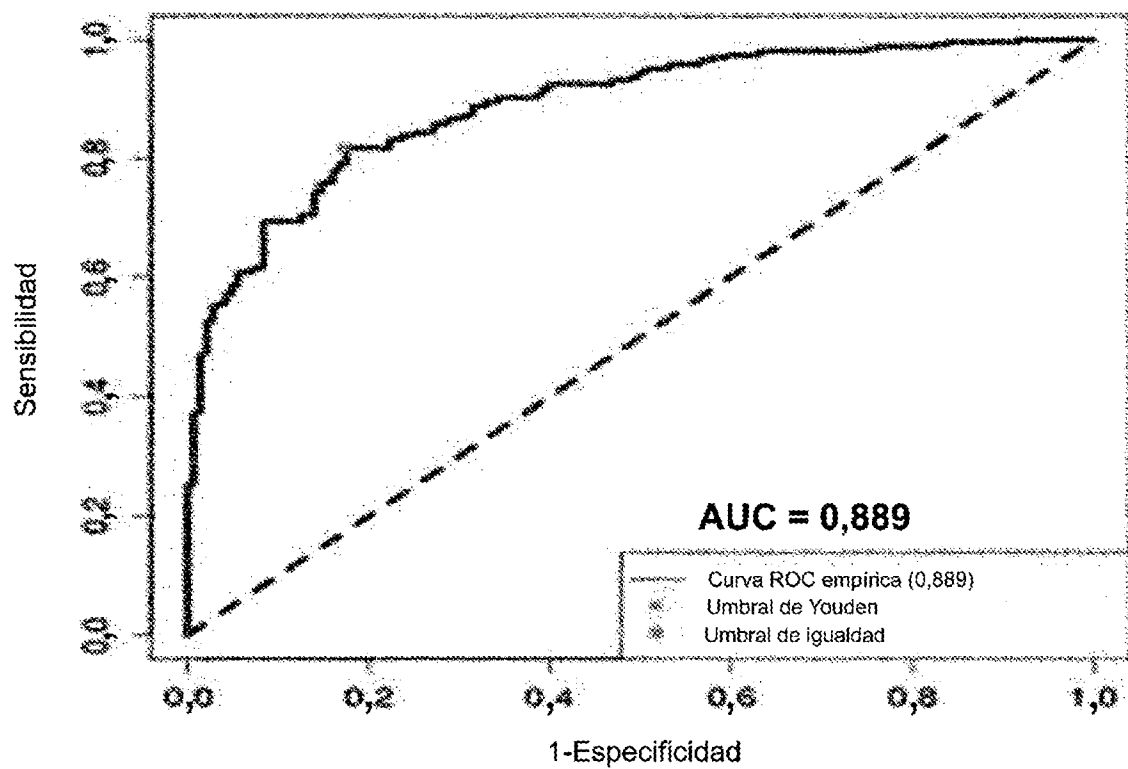


Fig. 7A

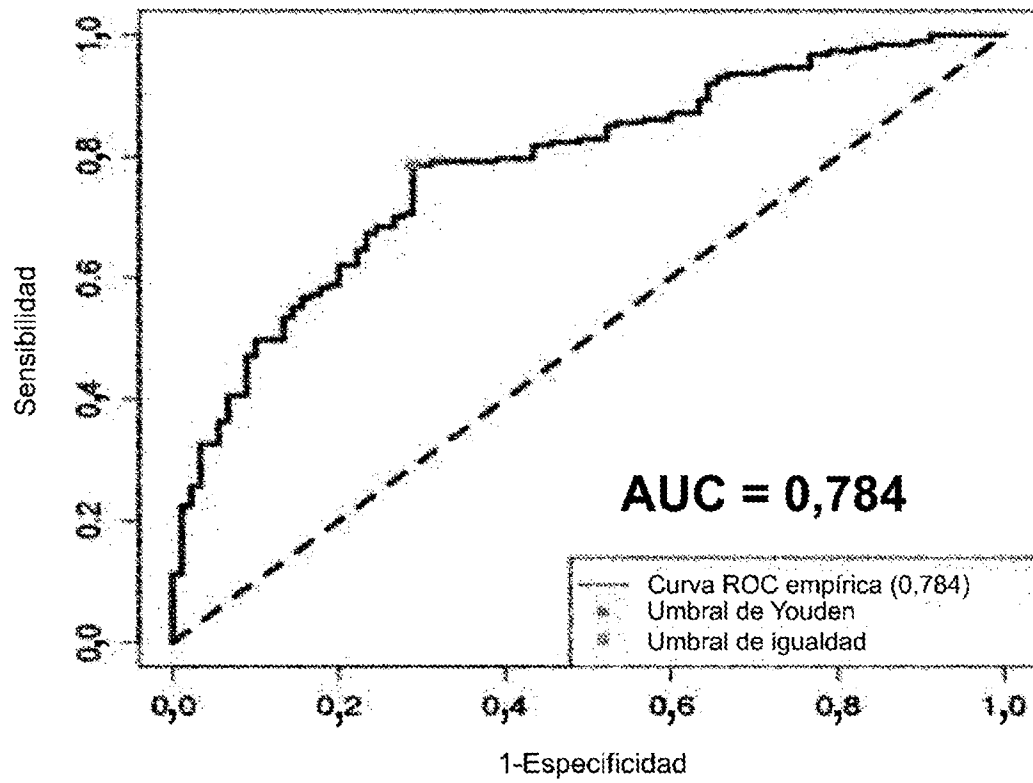


Fig. 7B

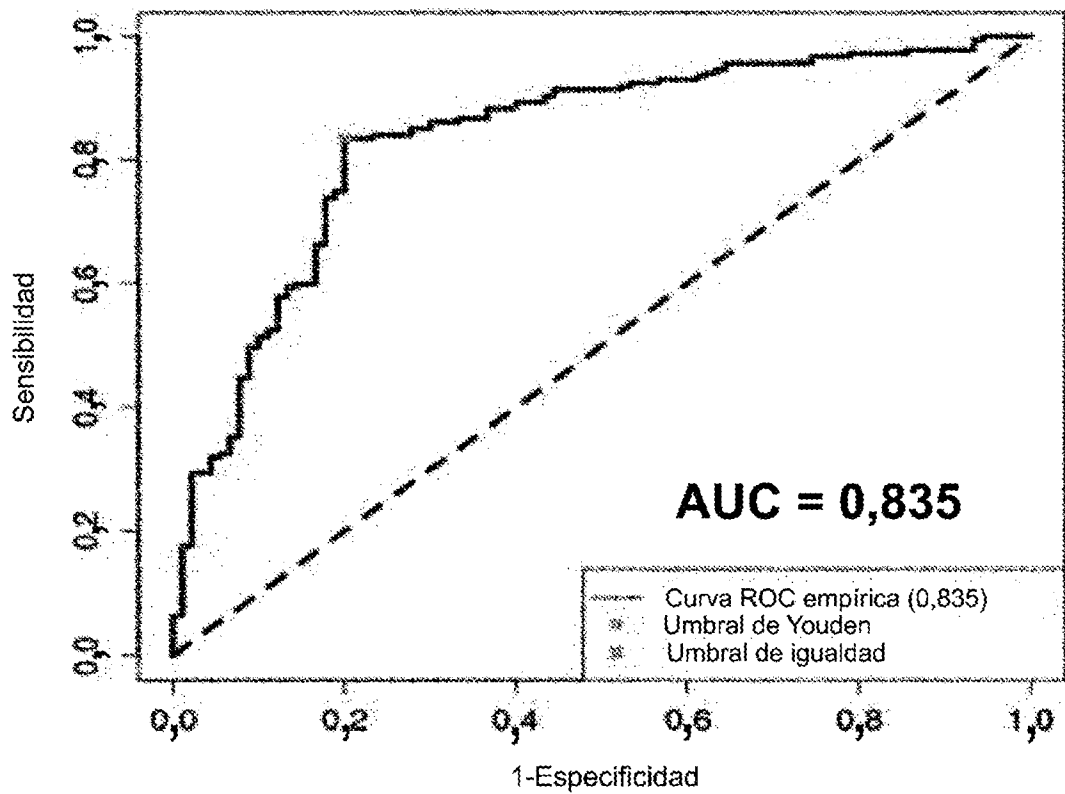


Fig. 7C

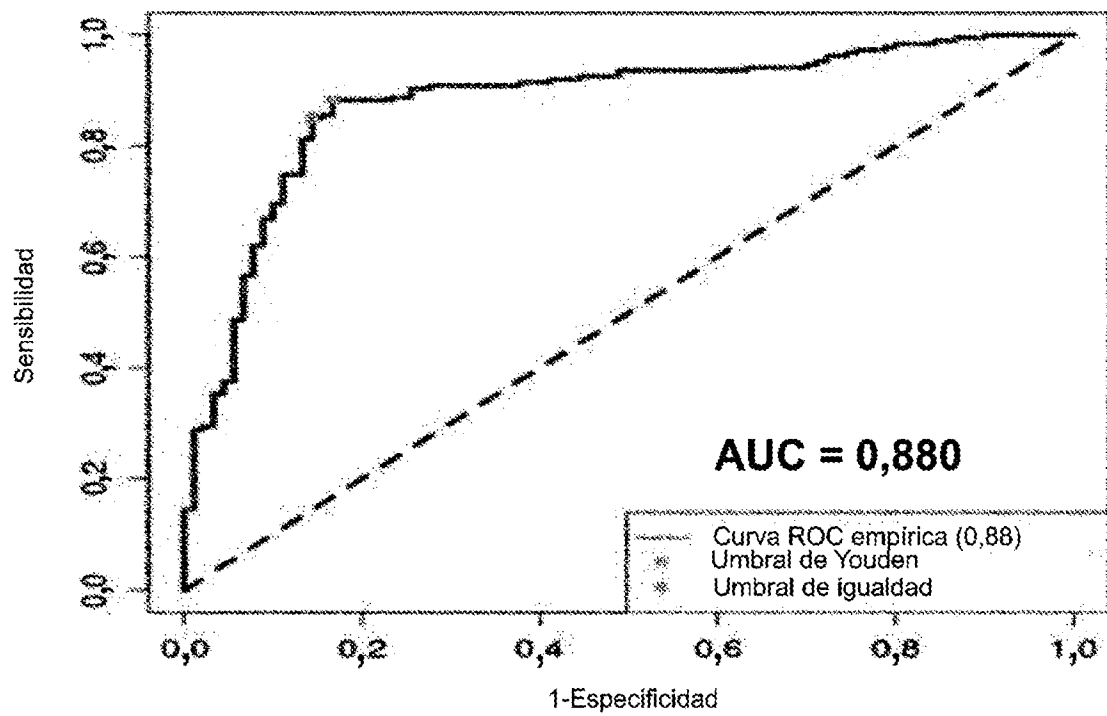


Fig. 7D

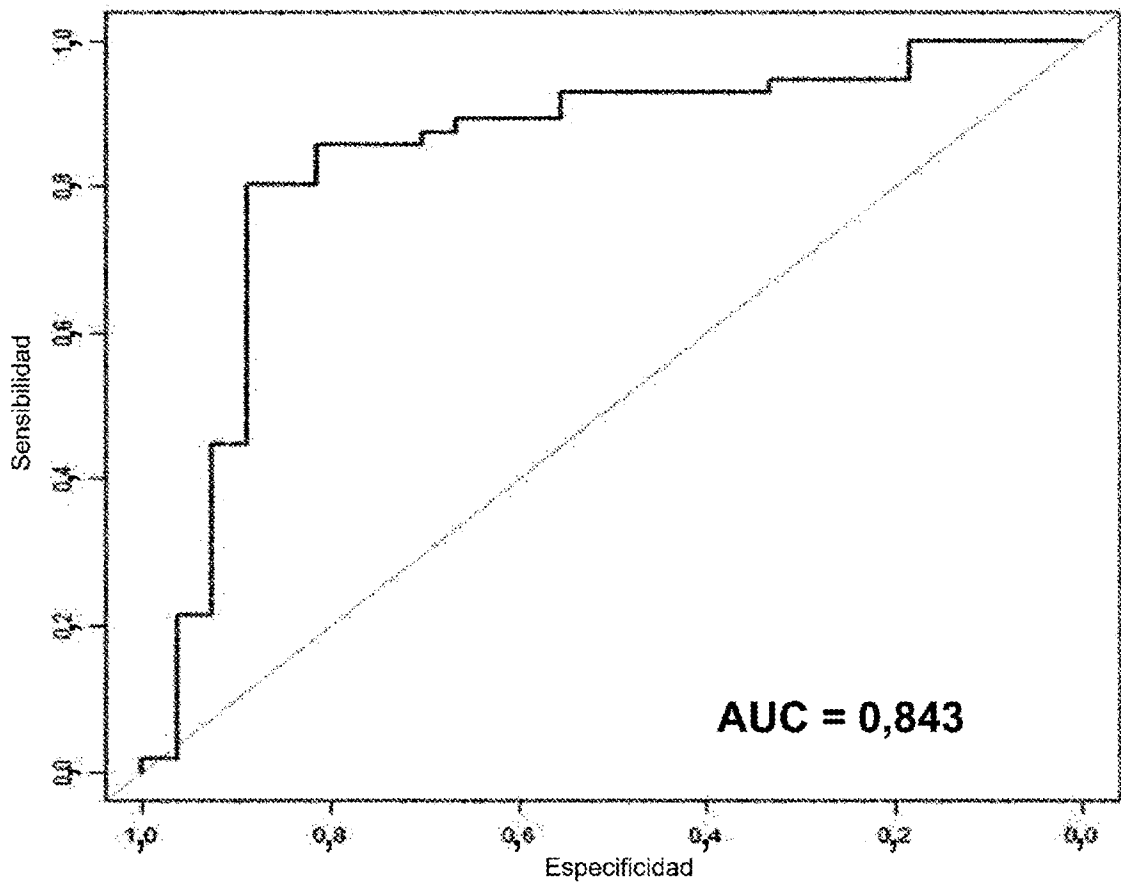


Fig. 7E

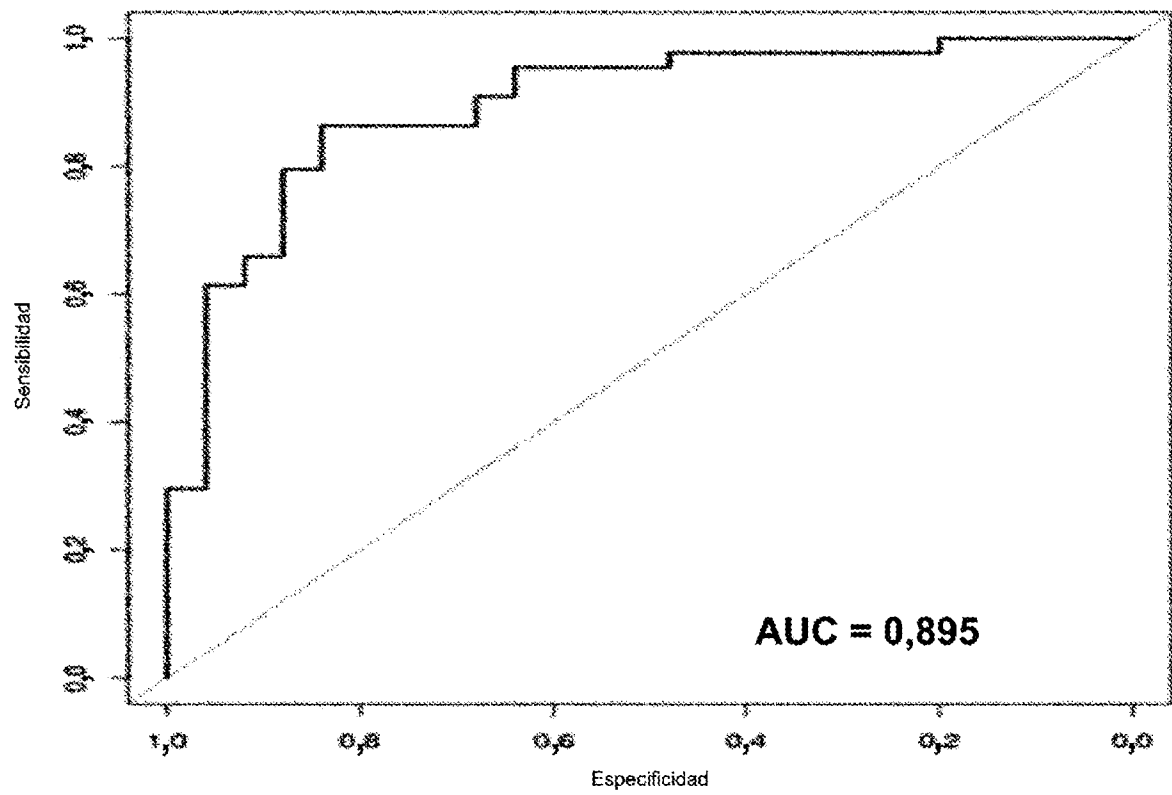


Fig. 8A

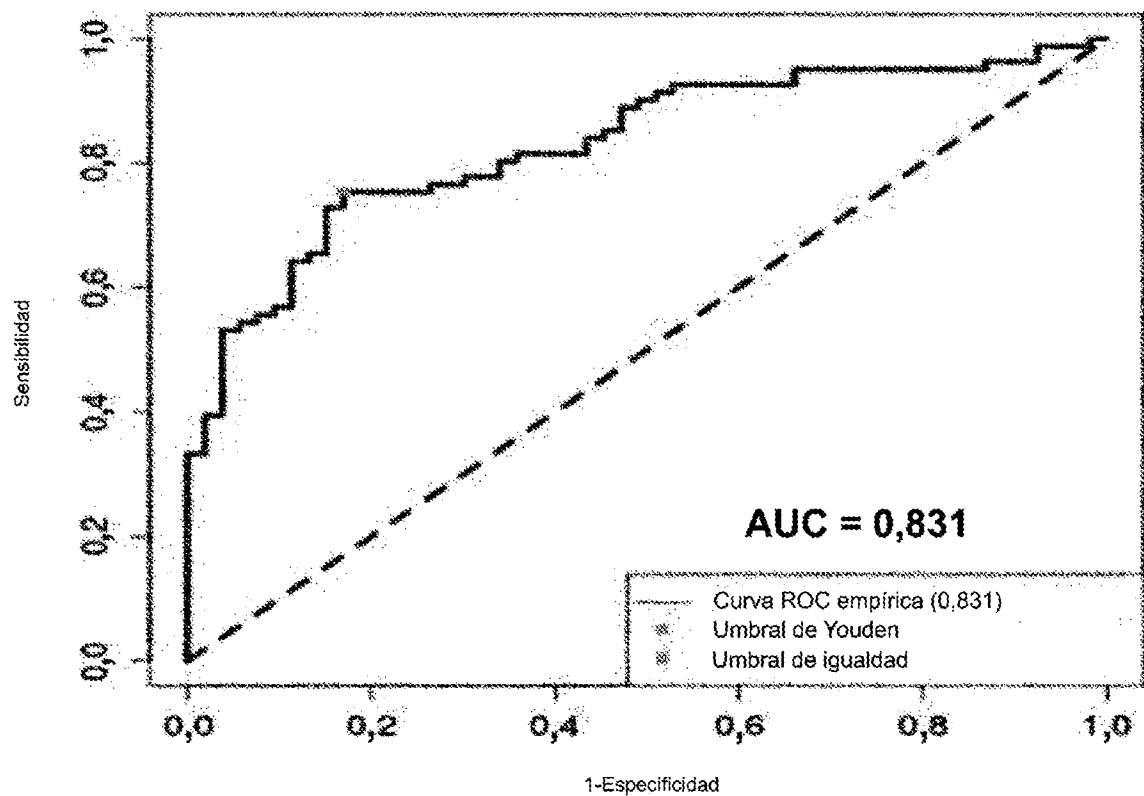


Fig. 8B

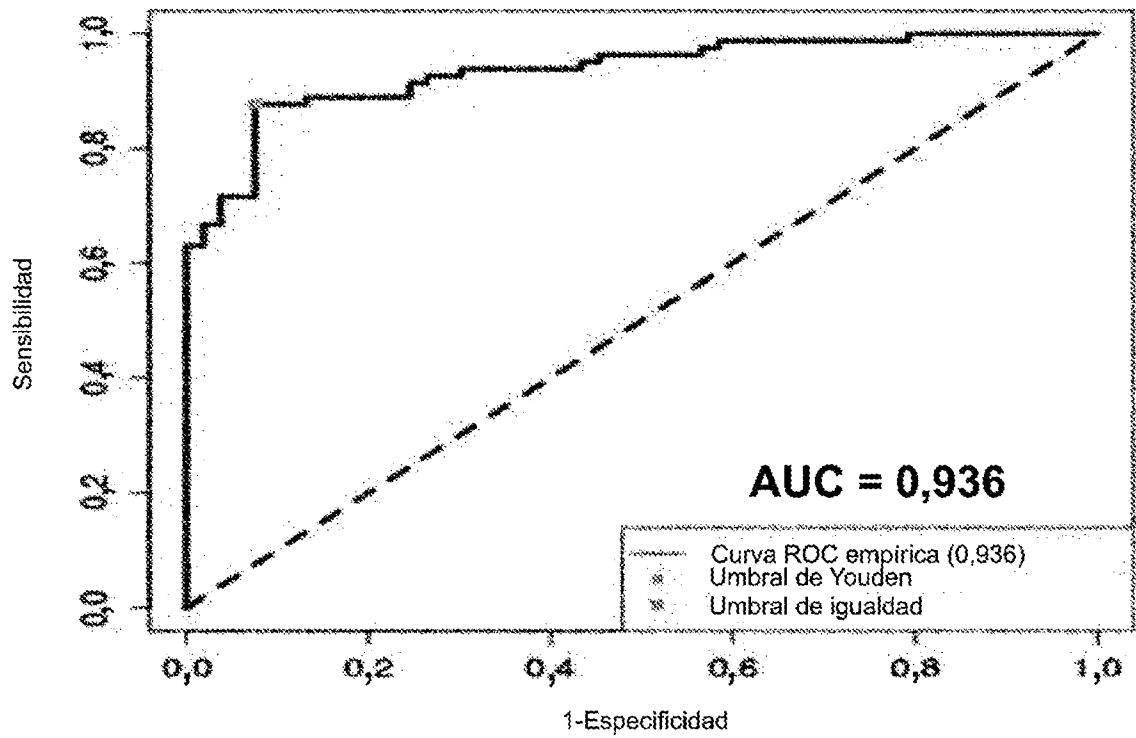


Fig. 8C

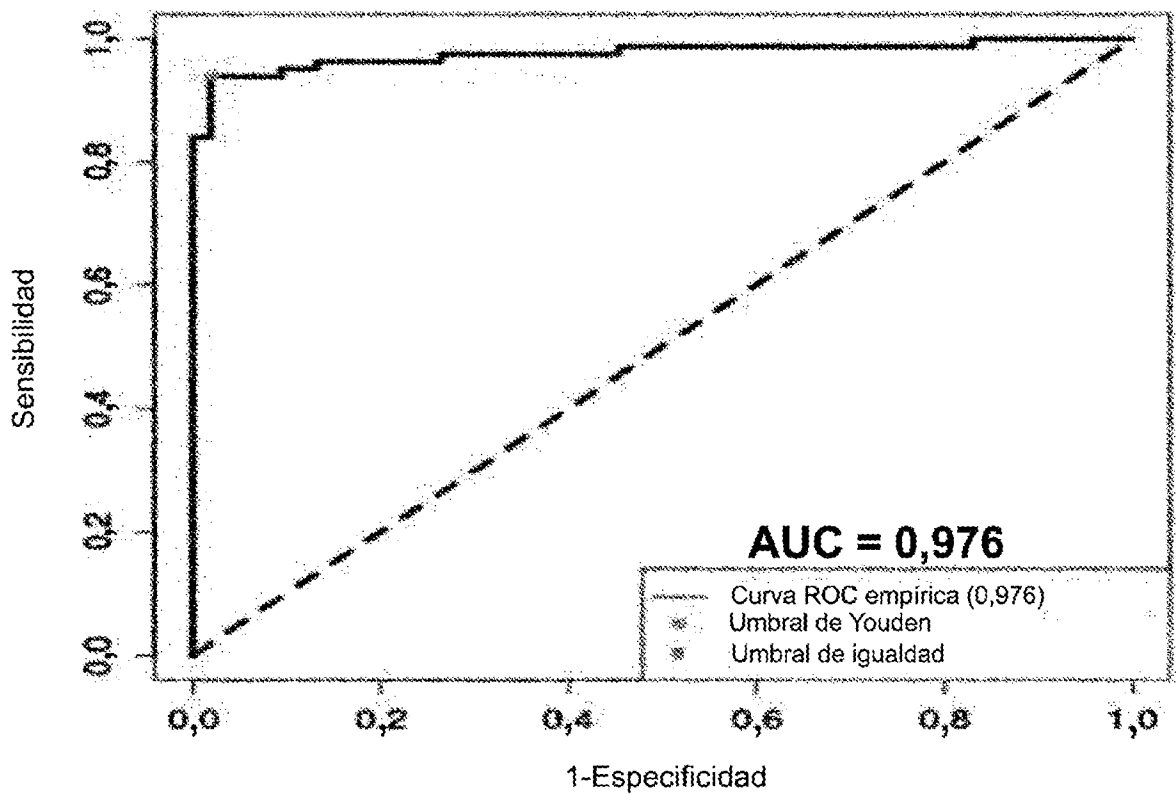


Fig. 8D

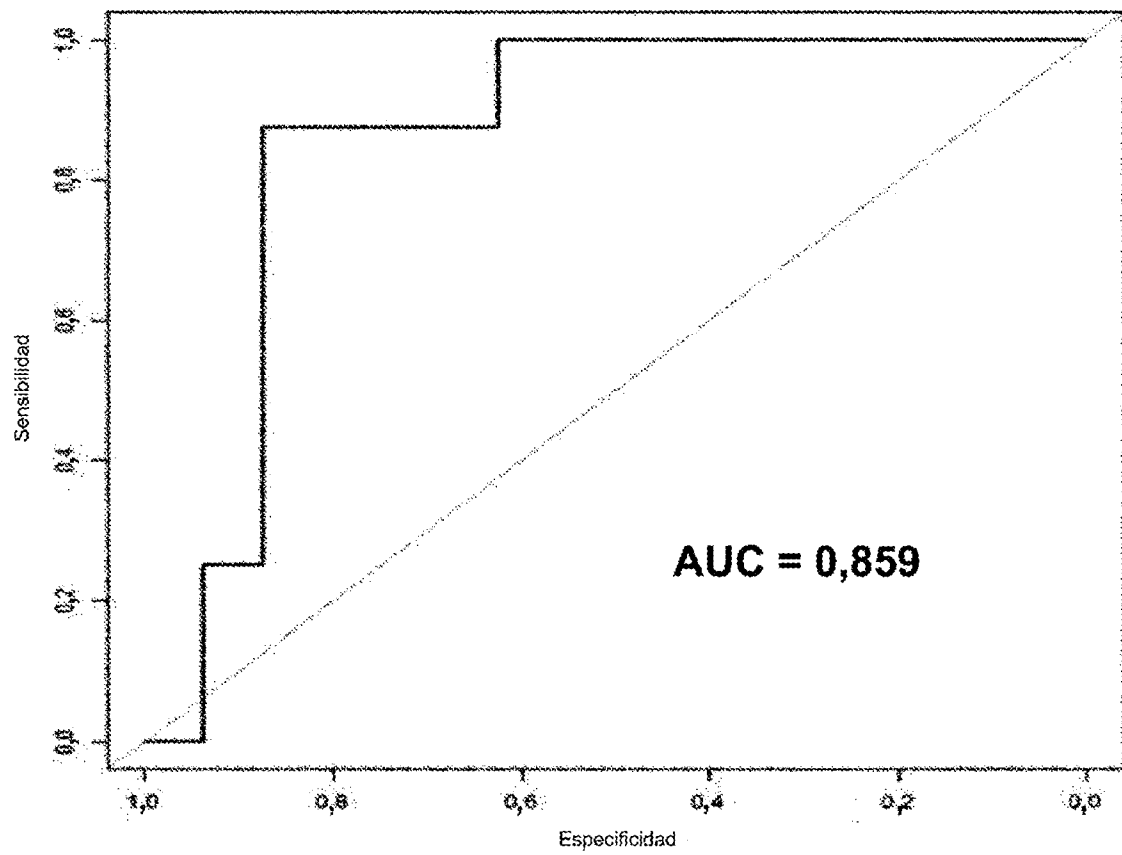


Fig. 8E

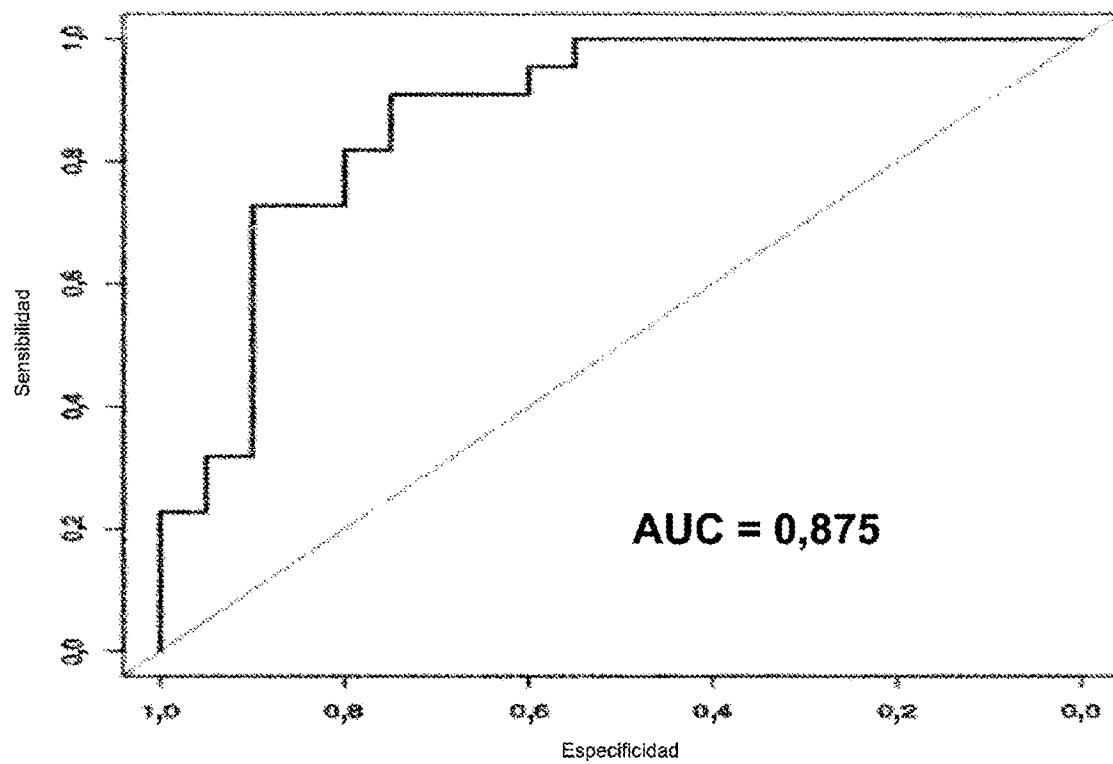


Fig. 9

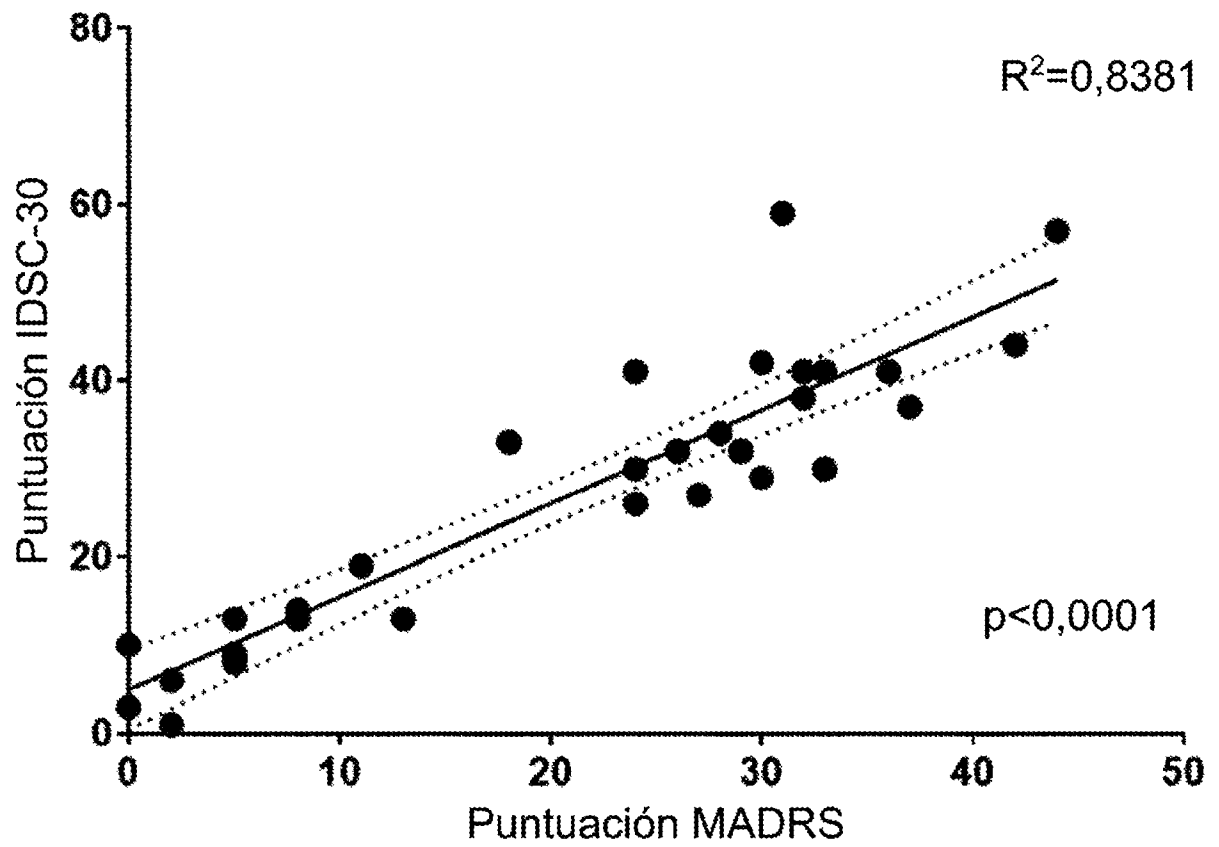


Fig. 10A

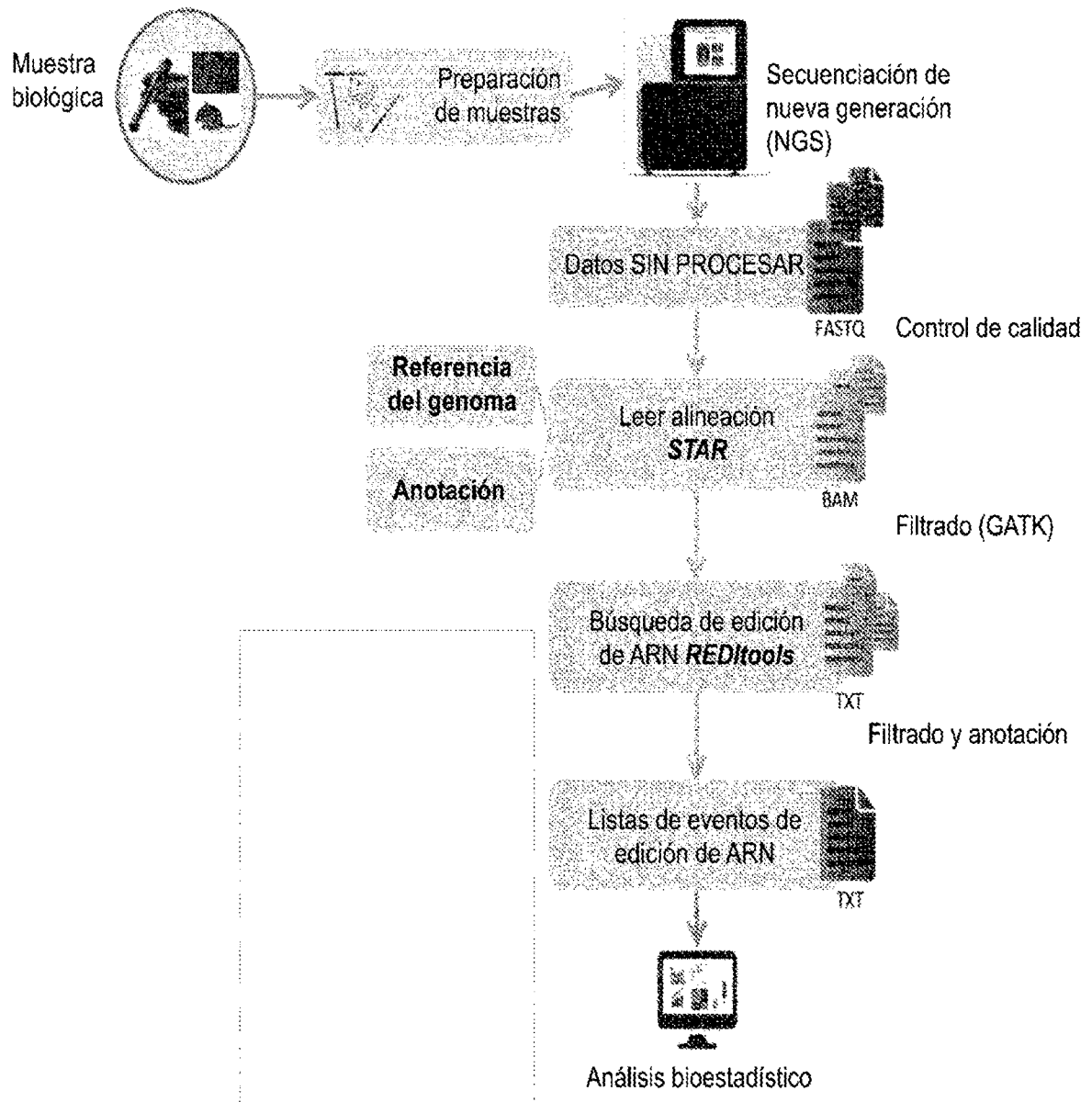


Fig. 10B

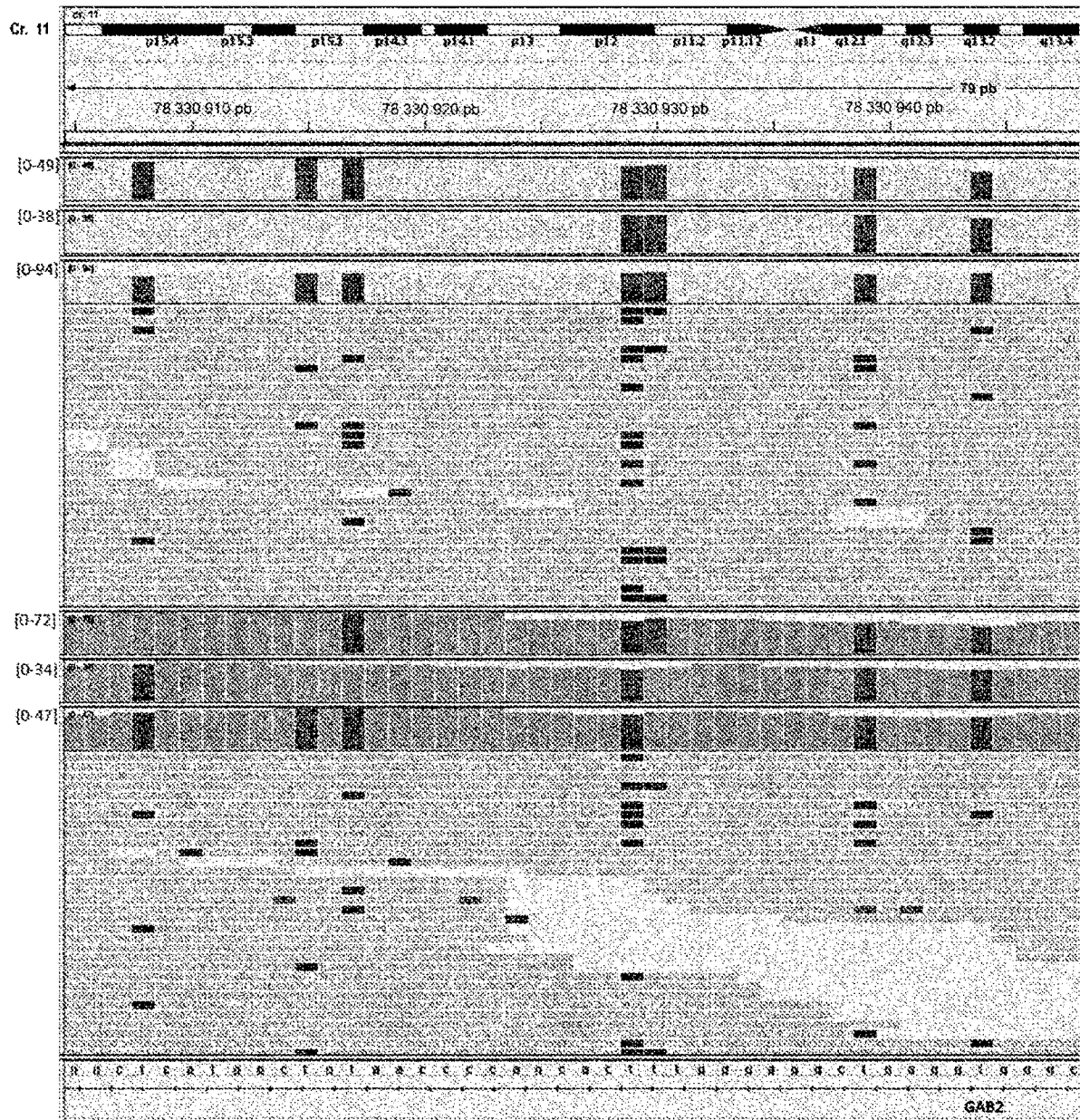


Fig. 11

	Muestra total	CTRL	DEP
Datos demográficos			
	n	n	n
Número	64	32	32
Edad (mín.-máx.)	24-61	24-61	26-61
Edad (media $\pm$ DE)	42,5 ($\pm$ 10,1)	42,5 ($\pm$ 10,1)	42,6 ($\pm$ 10,3)
Valor de p (frente a control)			0,46
Género			
Masculino (n (%))	32 (50 %)	16 (50 %)	16 (50 %)
Femenino (n (%))	32 (50 %)	16 (50 %)	16 (50 %)
IMC (kg/m²)		25,6	23,9
Valor de p (frente a control)			0,15
Características clínicas			
Puntuación de MADRS promedio (media $\pm$ EEM)	-	1,06 ($\pm$ 0,3)	21,3 ($\pm$ 2,5)
Valor de p (frente a control)	-		5,93E-09
Puntuación de IDSC30 promedio (media $\pm$ EEM)		2,8 ($\pm$ 0,5)	27,4 ($\pm$ 2,9)
Valor de p (frente a control)			2,59E-09

Los datos son la media $\pm$ EEM. Los valores p de las características principales se visualizan con la prueba t de Student.

CTRL: Controles sanos; DEP: episodio depresivo mayor; IMC: Índice de masa corporal

Fig. 12

Categoría	ID de GO	Término de GO	Relación	p(TDF)
Proceso biológico	GO:0002682	Proceso de regulación del sistema inmunitario	66/1391	2,37E-12
Proceso biológico	GO:0050776	Regulación de la respuesta inmunitaria	47/873	5,37E-10
Proceso biológico	GO:0002376	Proceso del sistema inmunitario	83/2370	1,09E-09
Proceso biológico	GO:0050778	Regulación positiva de la respuesta inmunitaria	36/589	7,68E-09
Proceso biológico	GO:0002768	Ruta de señalización del receptor de superficie celular que regula la respuesta inmunitaria	24/266	1,62E-08
Proceso biológico	GO:0002757	Transducción de señales que activan la respuesta inmunitaria	26/332	3,53E-08
Proceso biológico	GO:0002252	Proceso efector inmunitario	44/927	3,83E-08
Proceso biológico	GO:0002253	Activación de la respuesta inmunitaria	28/393	3,83E-08
Proceso biológico	GO:0002429	Ruta de señalización del receptor de superficie celular que activa la respuesta inmunitaria	22/234	3,83E-08
Proceso biológico	GO:0002684	Regulación positiva del proceso del sistema inmunitario	43/882	3,83E-08
Proceso biológico	GO:0002764	Ruta de señalización que regula la respuesta inmunitaria	27/365	3,83E-08
Proceso biológico	GO:0045321	Activación de leucocitos	42/894	1,13E-07
Proceso biológico	GO:0001775	Activación celular	45/1024	1,67E-07
Proceso biológico	GO:0080090	Regulación del proceso metabólico primario	144/5982	1,67E-07
Proceso biológico	GO:0044093	Regulación positiva de la función molecular	61/1713	2,90E-07
Proceso biológico	GO:0005464	Proceso de modificación de proteínas celulares	88/2999	3,36E-07
Proceso biológico	GO:0060255	Regulación del proceso metabólico de macromolecular	144/6072	3,63E-07
Proceso biológico	GO:0048584	Regulación positiva de la respuesta al estímulo	63/2054	3,89E-07
Proceso biológico	GO:0051171	Regulación del proceso metabólico de compuestos nitrogenados	139/5827	5,79E-07
Proceso biológico	GO:0006955	Respuesta inmunitaria	56/1560	8,88E-07

Fig. 13

ID de ruta	Nombre de ruta	Jerarquía de eventos	Relación de entidades	p(TDF)
R-HSA-202427	Fosforilación de las cadenas zeta de CD3 y TCR	Sistema inmunitario adaptativo	16/45	8,99E-10
R-HSA-202433	Generación de moléculas de segundo mensajero	Sistema inmunitario adaptativo	16/58	1,79E-08
R-HSA-202430	Translocación de ZAP-70 a sinapsis inmunitaria	Sistema inmunitario adaptativo	14/42	1,79E-08
R-HSA-388948	Señalización de PD-1	Sistema inmunitario adaptativo	14/45	3,27E-08
R-HSA-202403	Señalización de TCR	Sistema inmunitario adaptativo	21/146	8,91E-07
R-HSA-388841	Coestimulación mediante la familia CD28	Sistema inmunitario adaptativo	19/97	1,30E-06
R-HSA-202424	Señalización de TCR en sentido 3'	Sistema inmunitario adaptativo	17/124	3,53E-05
R-HSA-877300	Señalización del interferón gamma	Señalización de citocinas en el sistema inmunitario	21/250	0,00273264
R-HSA-2132295	Presentación del antígeno de clase MHC II	Sistema inmunitario adaptativo	15/148	0,00852889
R-HSA-9616222	Regulación transcripcional de la granulopoyesis	Biología del desarrollo	9/71	0,02720614

Fig. 14

ID de gen	Nombre de gen	Región	Cr.	Posición (GRCh38)	Cobertura	Múltiplo de variación	Valor de p	AUC
ENSG00000177663	IL17RA	3UTR	22	17110591	107	0,74	5,53E-03	0,929
ENSG00000183049	CAMK1D	Intrón	10	12387673	34	0,66	6,32E-03	0,912
ENSG00000142166	IFNAR1	3UTR	21	33355807	32	1,42	1,31E-02	0,910
ENSG00000177663	IL17RA	Intrón	22	17089708	48	1,25	6,94E-03	0,906
ENSG00000033327	GAB2	Intrón	11	76330944	37	1,43	3,71E-03	0,896
ENSG00000166501	LYN	Intrón	8	55890893	130	1,20	4,20E-02	0,880
ENSG00000106546	AHR	3UTR	7	17345159	61	1,36	1,47E-02	0,859
ENSG00000033800	PIAS1	Intrón	15	68138039	47	1,32	1,52E-02	0,857
ENSG00000166501	PRKCB	Intrón	16	23920896	78	0,74	3,43E-03	0,853
ENSG00000135679	MDM2	Intrón	12	68821612	38	1,31	4,45E-02	0,850
ENSG00000081237	PTPRC	Intrón	1	198680332	74	0,77	4,31E-03	0,846
ENSG00000166501	LYN	3UTR	8	56012553	44	1,48	2,11E-03	0,846
ENSG00000159110	IFNAR2	Intrón	21	33231766	37	0,73	6,82E-03	0,840
ENSG00000068028	RASSF1	Intrón	3	50332838	43	0,84	3,76E-02	0,825
ENSG00000157551	KCNJ15	Intrón	21	38288302	37	1,47	1,69E-02	0,819

Fig. 15

Categoría	ID del término	Descripción del término	p(TDF)	Proteínas coincidentes
Proceso biológico	GO:0035356	Transducción de señales intracelulares	7,64E-06	AHR, CAMK1D, GAB2, IFNAR1, IFNAR2, LYN, MDM2, PIAS1, PRKCB, RASSF1
Proceso biológico	GO:0007259	Ruta de señalización del receptor a través de JAK-STAT	1,41E-05	IFNAR1, IFNAR2, LYN, PIAS1
Proceso biológico	GO:0007166	Transducción de señales	3,19E-05	AHR, CAMK1D, GAB2, IFNAR1, IFNAR2, IL17RA, LYN, MDM2, PIAS1, PRKCB, PTPRC, RASSF1
Proceso biológico	GO:0050857	Regulación positiva de la ruta de señalización mediada por el receptor de antígenos	9,82E-05	LYN, PRKCB, PTPRC
Proceso biológico	GO:0002582	Proceso de regulación del sistema inmunitario	0,0002	AHR, CAMK1D, GAB2, IFNAR2, LYN, PIAS1, PRKCB, PTPRC
Proceso biológico	GO:0050853	Ruta de señalización del receptor de células B	0,00029	LYN, PRKCB, PTPRC
Proceso biológico	GO:0071310	Respuesta celular a sustancia orgánica	0,00039	AHR, GAB2, IFNAR1, IFNAR2, IL17RA, LYN, MDM2, PIAS1, PRKCB, PTPRC
Proceso biológico	GO:0024888	Regulación de la proliferación de células B	0,0011	PIAS1
Proceso biológico	GO:0050776	Regulación de la respuesta inmunitaria	0,0014	AHR, LYN, PTPRC
Proceso biológico	GO:0050861	Regulación positiva de la ruta de señalización del receptor de células B	0,0014	GAB2, IFNAR2, LYN, PIAS1, PRKCB, PTPRC
Proceso biológico	GO:0002768	Ruta de señalización del receptor de superficie celular que regula la respuesta inmunitaria	0,0021	LYN, PRKCB
Proceso biológico	GO:0007166	Ruta de señalización del receptor de superficie celular	0,0025	GAB2, LYN, PRKCB, PTPRC
Proceso biológico	GO:0048522	Regulación positiva del proceso celular	0,0025	GAB2, IFNAR1, IFNAR2, IL17RA, LYN, PIAS1, PRKCB, PTPRC
Proceso biológico	GO:0048583	Regulación de la respuesta al estímulo	0,0025	AHR, CAMK1D, GAB2, IL17RA, LYN, MDM2, PIAS1, PRKCB, PTPRC, RASSF1
Proceso biológico	GO:0069950	Respuesta al estrés	0,0041	CAMK1D, GAB2, IFNAR2, IL17RA, LYN, MDM2, PIAS1, PRKCB, PTPRC
Proceso biológico	GO:0009893	Regulación positiva del proceso metabólico	0,0041	S1, PRKCB, PTPRC
Proceso biológico	GO:0051049	Regulación del transporte	0,0041	TPTC, RASSF1
Proceso biológico	GO:0032879	Regulación de la localización	0,0046	AHR, CAMK1D, LYN, MDM2, PIAS1, PRKCB, PTPRC, RASSF1
Proceso biológico	GO:0006552	Respuesta de defensa	0,0047	CAMK1D, GAB2, IL17RA, KCMJ15, LYN, MDM2, PRKCB
Proceso biológico	GO:0007049	Ciclo celular	0,0048	CAMK1D, GAB2, IL17RA, KCMJ15, LYN, MDM2, PRKCB, PTPRC

Entidades coincidentes	p(TDF)	Nombre de ruta	ID de ruta
PRKCB	0,01235604	RUNX1 regula la transcripción de genes implicados en la diferenciación de células mieloides	R-HSA-8939246
MDM2;PIAS1	0,0203032	SUMOilación de factores de transcripción	R-HSA-3232118
IFNAR2;IFNAR1	0,02527928	Regulación de la señalización de IFNA	R-HSA-912694
PRKCB;MDM2	0,02527928	Trafico de receptores AMPA	R-HSA-399719
PRKCB;MDM2	0,02527928	Unión de glutamato, activación de receptores AMPA y plasticidad sináptica	R-HSA-399721
MDM2	0,02527928	Regulación de la degradación de TP53	R-HSA-6804757
MDM2;PIAS1	0,02527928	SUMOilación de proteínas de ubiquitinación	R-HSA-3232142
MDM2	0,02527928	Regulación de la expresión y degradación de TP53	R-HSA-6806003
LYN;GAB2	0,02527928	Señalización de interleucina 3, interleucina 5 y GM-CSF	R-HSA-512988
LYN;GAB2	0,02527928	Señalización por SCF-KIT	R-HSA-1433557
LYN;IFNAR2;GAB2;IL17RA;PIAS1;IFNAR1	0,02640054	Señalización de citoquinas en el sistema inmunitario	R-HSA-1280215
PRKCB;MDM2;KCNJ15	0,02887316	Receptores de neurotransmisores y transmisión de señales postsinápticas	R-HSA-112314

Fig. 17

Muestra total			CIRL	DEP
Datos demográficos				
Número	n	n		
	76	41		35
Edad (mín.-máx.)	19-61	21-61		19-60
Edad (media $\pm$ DE)	43,1 ($\pm$ 11,7)	43,5 ($\pm$ 12,7)		42,5 ($\pm$ 10,6)
Valor de p (frente a control)				0,36
Género				
Masculino (n (%))	39 (51,32 %)	21 (51,22 %)		16 (50 %)
Femenino (n (%))	37 (48,68 %)	20 (48,78 %)		16 (50 %)
IMC (kg/m ²)		25,2		23,2
Valor de p (frente a control)				0,50
Características clínicas				
Puntuación de MADRS promedio (media $\pm$ EEM)		0,76 ($\pm$ 0,2)		24,3 ($\pm$ 1,8)
Valor de p (frente a control)				7,16E-27
Puntuación de IDSC30 promedio (media $\pm$ EEM)		2,6 ($\pm$ 0,4)		30,6 ($\pm$ 2,1)
Valor de p (frente a control)				4,25E-26

Fig. 18

	Muestra total	CTRL	DEP
Datos demográficos			
	n	n	n
Número	86	28	58
Edad (mín.-máx.)	18-64	21-64	18-62
Edad (media $\pm$ DE)	43,2 ($\pm$ 12,3)	45,8 ($\pm$ 14,4)	41,4 ($\pm$ 11,2)
Valor de p (frente a control)			0,16
IMC (kg/m²)		24,2	23,5
Valor de p (frente a control)			0,44
Género			
Masculino (n (%))	40 (46,5 %)	13 (46,5 %)	27 (46,6 %)
Femenino (n (%))	46 (53,5 %)	15 (53,5 %)	31 (53,4 %)
Características clínicas			
Puntuación de MADRS promedio (media $\pm$ EEM)		0,5 ($\pm$ 1,1)	33,1 ($\pm$ 0,9)
Valor de p (frente a control)			<0,00001
Puntuación de IDSC30 promedio (media $\pm$ EEM)		2 ($\pm$ 1,9)	38,6 ($\pm$ 1,1)
Valor de p (frente a control)			<0,00001

Fig. 19

	Muestra total	CTRL	DEP
Datos demográficos			
	n	n	n
Número	411	143	268
Edad (min.-máx.)	18-67	18-65	18-67
Edad (media $\pm$ DE)	39,4 ($\pm$ 13,4)	39,4 ($\pm$ 14,1)	39,4 ($\pm$ 13,1)
Valor de p (frente a control)			0,99
Género			
Masculino (n (%))	134 (32,6 %)	53 (37,1 %)	81 (30,2 %)
Femenino (n (%))	277 (67,4 %)	90 (62,9 %)	187 (69,8 %)
IMC (kg/m ²)		24,6	23,8
Valor de p (frente a control)			0,12
Características clínicas			
Puntuación de MADRS promedio (media $\pm$ EEM)	-	0,8 ($\pm$ 0,1)	29,7 ($\pm$ 0,5)
Valor de p (frente a control)	-		<0,00001
Puntuación de IDSC30 promedio (media $\pm$ EEM)	-	1,97 ($\pm$ 0,2)	36,5 ($\pm$ 0,5)
Valor de p (frente a control)	-		<0,00001

Fig. 20

Categoría	ID de GO	Término de GO	TDF
Proceso biológico	GO:0002376	Proceso del sistema inmunitario	1,77E-10
Proceso biológico	GO:0002682	Proceso de regulación del sistema inmunitario	7,45E-10
Proceso biológico	GO:0002252	Proceso efector inmunitario	1,52E-08
Proceso biológico	GO:0002253	Activación de la respuesta inmunitaria	7,50E-08
Proceso biológico	GO:0050776	Regulación de la respuesta inmunitaria	1,19E-07
Proceso biológico	GO:0006955	Respuesta inmunitaria	1,51E-07
Proceso biológico	GO:0002757	Transducción de señales que activan la respuesta inmunitaria	1,72E-07
Proceso biológico	GO:0002764	Ruta de señalización que regula la respuesta inmunitaria	1,74E-07
Proceso biológico	GO:0050778	Regulación positiva de la respuesta inmunitaria	2,22E-07
Proceso biológico	GO:0002684	Regulación positiva del proceso del sistema inmunitario	4,15E-07
Proceso biológico	GO:0048584	Regulación positiva de la respuesta al estímulo	6,90E-07
Proceso biológico	GO:0002444	Inmunidad mediada por leucocitos mieloides	1,0526E-06
Proceso biológico	GO:0002274	Activación de leucocitos mieloides	4,8148E-06
Proceso biológico	GO:0002446	Inmunidad mediada por neutrófilos	8,1445E-06
Proceso biológico	GO:0002366	Activación de leucocitos implicada en la respuesta inmunitaria	8,4827E-06
Proceso biológico	GO:0035556	Transducción de señales intracelulares	8,6505E-06
Proceso biológico	GO:0043299	Desgranulación leucocitaria	8,8488E-06
Proceso biológico	GO:0045321	Activación de leucocitos	9,2368E-06
Proceso biológico	GO:0002263	Activación celular implicada en la respuesta inmunitaria	9,7805E-06
Proceso biológico	GO:0002768	Ruta de señalización del receptor de superficie celular que regula la respuesta inmunitaria	1,059E-05

Fig. 21

Categoría	ID de GO	Término de GO	TDF	N de genes
Proceso biológico	GO:0002682	proceso de regulación del sistema inmunitario	1,00E-10	67
Proceso biológico	GO:0002376	proceso del sistema inmunitario	1,51E-07	84
Proceso biológico	GO:0050776	regulación de la respuesta inmunitaria	5,90E-09	50
Proceso biológico	GO:0050778	regulación positiva de la respuesta inmunitaria	2,68E-07	38
Proceso biológico	GO:0045321	activación de leucocitos	3,61E-07	42
Proceso biológico	GO:0002684	regulación positiva del proceso del sistema inmunitario	5,45E-07	45
Proceso biológico	GO:0048584	regulación positiva de la respuesta al estímulo	5,99E-07	73
Proceso biológico	GO:0001775	activación celular	6,94E-07	45
Proceso biológico	GO:0036211	proceso de modificación de proteínas	7,43E-07	86
Proceso biológico	GO:0006464	proceso de modificación de proteínas celulares	7,43E-07	86
Proceso biológico	GO:0060255	regulación del proceso metabólico de macromolecular	9,07E-07	139
Proceso biológico	GO:0002757	transducción de señales que activan la respuesta inmunitaria	1,06E-06	28
Proceso biológico	GO:0002253	activación de la respuesta inmunitaria	1,07E-06	30
Proceso biológico	GO:0002252	proceso efector inmunitario	1,07E-06	43
Proceso biológico	GO:0002764	ruta de señalización que regula la respuesta inmunitaria	1,12E-06	29
Proceso biológico	GO:0043412	modificación de macromoléculas	1,73E-06	88
Proceso biológico	GO:0019222	regulación del proceso metabólico	4,11E-06	145
	GO:0002366	activación de leucocitos implicada en la respuesta inmunitaria	4,11E-06	31
Proceso biológico	GO:0002263	activación celular implicada en la respuesta inmunitaria	4,55E-06	31
Proceso biológico	GO:0002274	activación de leucocitos mieloides	5,48E-06	30

Fig. 22A

Modelo de bosques aleatorios 1
Ajustado por edad, sexo, tratamientos y adicciones

	Rendimiento
AUC	0,929
IC 95 %	[0,876 ; 0,981]
Se	82,9 %
Sp	87,1 %
PPV	94,4 %
NPV	65,9 %
Prec	84,1 %
Corte	0,5

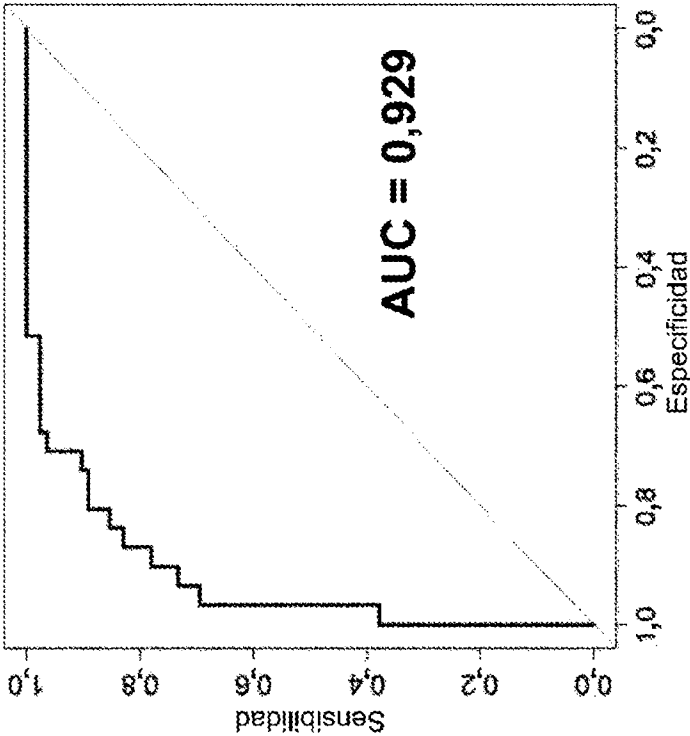


Fig. 22B

Modelo de bosques aleatorios 2
Ajustado por edad, sexo

	Rendimiento
AUC	0,911
IC 95 %	[0,854 ; 0,968]
Se	80,0 %
Sp	82,4 %
PPV	89,7 %
NPV	68,3 %
Prec	80,8 %
Corte	0,5

