

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 966 015**

51 Int. Cl.:

C12Q 1/6886 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **25.07.2019 PCT/NL2019/050487**

87 Fecha y número de publicación internacional: **30.01.2020 WO20022895**

96 Fecha de presentación y número de la solicitud europea: **25.07.2019 E 19787488 (6)**

97 Fecha y número de publicación de la concesión europea: **08.11.2023 EP 3827101**

54 Título: **Firmas génicas para predecir metástasis del melanoma y pronóstico del paciente**

30 Prioridad:

25.07.2018 US 201862703221 P

25.10.2018 US 201862750294 P

12.12.2018 US 201862778777 P

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

17.04.2024

73 Titular/es:

SKYLINEDX B.V. (50.0%)

Lichtenauerlaan 40

3062 ME Rotterdam, NL y

MAYO FOUNDATION FOR MEDICAL EDUCATION AND RESEARCH (50.0%)

72 Inventor/es:

MEVES, ALEXANDER y

BELLOMO, DOMENICO

74 Agente/Representante:

SÁEZ MAESO, Ana

ES 2 966 015 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

Firmas génicas para predecir metástasis del melanoma y pronóstico del paciente

Campo de la invención

5 La divulgación proporciona firmas génicas para clasificar a un individuo que padece melanoma cutáneo. Las "firmas génicas SLN" proporcionadas en la presente clasifican a un individuo basado en el pronóstico y/o clasifican a un individuo como que tiene un ganglio linfático centinela (SLN) positivo- o negativo- para metástasis. Las "firmas génicas N-SLN" proporcionadas en la presente clasifican a un individuo como que tiene un ganglio linfático no centinela (N-SLN) positivo- o negativo- para metástasis.

Antecedentes de la invención

10 El melanoma cutáneo es un tumor maligno que surge de las células melanocíticas productoras de pigmento ubicadas principalmente en el estrato basal de la epidermis de la piel. La mayoría de las lesiones aparecen con signos descritos por las reglas ABCDE: Asimetría, Bordes irregulares, Cambios de color, Diámetro superior a 6 mm y Evolución, es decir, la tendencia a cambiar rápidamente Abbasi NR, Shaw HM, Rigel DS, et al. Early Diagnosis of Cutaneous Melanoma Revisiting the ABCD Criteria. JAMA. 2004;292(22):2771-2776). Los sitios suelen ser asintomáticos, sin embargo, pueden causar picazón y/o sangrado, especialmente en etapas tardías. La detección de lesiones sospechosas suele realizarse mediante el autoexamen de la piel, el cual se recomienda realizar de forma rutinaria con base en los criterios ABCDE o los "signos del patito feo" (Grob J. The 'Ugly Duckling' Sign: Identification of the Common Characteristics of Nevus in an Individual as a Basis for Melanoma Screening. Arch Dermatol. 1998; 134:103-104). Cuando posteriormente un médico diagnostica formalmente el melanoma, es importante determinar el subtipo específico, ya que se conocen múltiples variedades clínicas y patológicas. La forma más común es el melanoma cutáneo - un melanoma de extensión superficial que representa alrededor del 70% de los casos - y que ocurre especialmente con frecuencia en personas de piel clara. La gravedad de la afección depende en gran medida de la capacidad de las células del melanoma para migrar fuera de la región primaria. Por esta razón, es de suma importancia evaluar si el tumor está localizado o se ha diseminado a los ganglios u órganos linfáticos.

25 La estadificación del melanoma es crucial para el pronóstico del paciente y para decidir estrategias adicionales de monitoreo y tratamiento. Esto también se refleja en el sorprendente contraste en las cifras de supervivencia general a 5-años reportadas en una revisión sistemática de la literatura de 9 países europeos: 95%-100% (estadio I), 65%-92.8% (estadio II), 41%-71% (estadio III), y 9%-28% (estadio IV). Estas diferencias dependen en gran medida de la capacidad del melanoma para metastatizar, a diferencia de las lesiones de melanoma localizadas. La discriminación precisa entre las diferentes clases de estadios es importante y se basa más comúnmente en el sistema TNM, que se refiere al grosor del tumor primario (T), la presencia y/o extensión de las células tumorales a los ganglios linfáticos (N) y la presencia de metástasis a distancia a otros órganos (M).

30 Cuando se evalúa la extensión del estadio-tumoral primario, los médicos tienen en cuenta el grosor del tumor, pero también características adicionales tales como la presencia de ulceración y la tasa mitótica de las células del tumor primario. Normalmente, sólo en pacientes con un estadio-tumoral alto, debido a un melanoma grueso y/u otras variables como ulceración, se evaluarán los ganglios linfáticos y la diseminación metastásica. Como se entiende en la técnica, el pronóstico se refiere a una predicción del resultado médico del paciente. Por ejemplo, un individuo puede clasificarse como que tiene un pronóstico pobre o un buen pronóstico. El pronóstico de un paciente que padece melanoma indica, por ejemplo, probabilidad de supervivencia a largo-plazo, supervivencia global, supervivencia libre de progresión, predicción de reincidencia versus remisión de la enfermedad, y progresión de la enfermedad.

35 Actualmente, un procedimiento ampliamente aplicado para estratificar con precisión a los pacientes y predecir el pronóstico es determinar la presencia de metástasis en los SLNs mediante el método SLNB (Biopsia del Ganglio Linfático Centinela). El manejo de los pacientes con melanoma cutáneo ha progresado significativamente desde la inclusión del mapeo linfático mediante el procedimiento de SLNB a principios de los años 1990s. Morton DL, Wen DR, Wong JH, et al. Technical details of intraoperative lymphatic mapping for early stage melanoma. Arch Surg. 1992;127(4):392-399). La técnica quirúrgica se ha mejorado mediante la implementación de una aproximación intraoperatoria de doble modalidad que utiliza tinte azul y un radiotrazador con detección por sonda gamma. Además, la evaluación patológica se mejoró mediante el empleo de cortes seccionados del SLN e inmunohistoquímica. Esto ha dado como resultado una mejor identificación del primer ganglio linfático de drenado o grupo de ganglios que está situado muy cerca del tumor (es decir, el SLN) y, por lo tanto, es un sitio probable de enfermedad metastásica. Este procedimiento también se conoce como "mapeo de ganglio linfático centinela".

50 El impacto significativo de la positividad de SLN también se demostró en el estudio MSLT-1, el cual indicó una diferencia en la tasa de supervivencia a 5-años para pacientes con SLNs positivo para tumor del 72.3% y pacientes con SLNs negativo para tumor del 90.2%, respectivamente (Morton DL, Thompson JF, Cochran AJ, et al. Sentinel-Node Biopsy or Nodal Observation in Melanoma. new engl J Med. 2006;355(13):1307-1317). De acuerdo con la 8va. edición de las directrices sobre melanoma del American Joint Committee on Cancer (AJCC), el procedimiento SLNB se recomienda para pacientes con un melanoma cutáneo de ≥ 0.8 mm (Gershenwald JE, Scolyer RA, Hess KR, et al. Melanoma Staging: Evidence-Based Changes in the American Joint Committee on Cancer Eighth Edition Cancer

Staging Manual. CA Cancer J Clin. 2017;67(6):472-492). Para este grupo de pacientes, generalmente se realiza un procedimiento SLNB y el tratamiento adicional depende del nivel de metástasis. Dentro del grupo límite, se puede considerar SLNB, particularmente si el melanoma se presenta con parámetros pronósticos adversos adicionales. Para pacientes con un grosor de melanoma de <0.8 mm, el tratamiento estándar generalmente se considera que es suficiente y no se recomienda SLNB. El tratamiento estándar implica la escisión local del melanoma primario con márgenes amplios, es decir, resección quirúrgica del tumor. Como se utiliza en la presente, se entiende por "resección" la extirpación quirúrgica del tejido maligno característico del melanoma de un paciente humano. De acuerdo con una realización, se entenderá que resección significa la eliminación del tejido maligno de manera que la presencia de tejido maligno restante dentro de dicho paciente sea indetectable con los métodos disponibles.

La tasa a la que las SLNBs se clasifican como positivas es muy variable y depende en gran medida de los factores pronósticos conocidos del tumor primario. Para los pacientes en estadio clínico I o II, el porcentaje de metástasis de SLNB es 15-30%, mientras que en los melanomas delgados se muestra que es 5.2%. La edición actual de las directrices de estadificación del Melanoma Expert Panel destacó la relevancia clínica de la subcategorización de los melanomas T1 en 0.8 mm. Esto se basa en la tendencia detectada en varios estudios de supervivencia de melanomas T1 de que existe una intercepción clínica potencial en la región de 0.7 a 0.8 mm. Sin embargo, basándose en el seguimiento a largo plazo de los pacientes después de una SLNB, se demostró que la recurrencia ganglionar regional tiene lugar en pacientes con ganglios centinela que inicialmente estaban libres de tumor. Esta información permitió calcular el desempeño de la prueba SLNB, que se estableció que tenía una tasa general de falsos-negativos de 12.5%. Más recientemente, Morton et al. reportaron que en los melanomas de grosor intermedio con una SLNB positivo de 16.0%, 4.8% tuvo un resultado de prueba falso negativo basado en la recurrencia dentro de un período de seguimiento de 10 años. Para el 32.9% de SLNB positivo de los melanomas gruesos, la tasa de falsos negativos fue 10.3%.

SLNB no es sólo un método para estadificar potencialmente el melanoma cutáneo, sino que también forma parte de un tratamiento el cual, dependiendo de la clasificación metastásica del SLN, puede ser necesario o no. El procedimiento SLNB puede causar complicaciones al paciente y el coste es alto. Por lo tanto, este procedimiento sólo se realiza en un grupo selecto de pacientes que se considera que tienen un mayor riesgo de diseminación metastásica - de la gran mayoría de las lesiones de bajo riesgo. El riesgo de metástasis puede evaluarse mediante la evaluación de factores clínico-patológicos que incluyen la profundidad de la invasión del tumor, conocida como profundidad de Breslow, y la ulceración de la superficie del tumor. Los tumores ulcerosos y melanomas que crecen más profundamente verticalmente hacia la piel se asocian con un mayor riesgo de un resultado adverso. Por ejemplo, las biopsias del SLN no se recomiendan para melanomas delgados T1a, 'pueden recomendarse' para pacientes con melanoma delgado T1b, se recomiendan para pacientes con melanoma de grosor intermedio T2 y T3 y 'pueden recomendarse' para pacientes con melanoma-grueso T4.

Aunque este concepto de uso de variables clínico-patológicas generalmente permite identificar pacientes de alto riesgo en los extremos del espectro tumoral, este método no es preciso para lesiones intermedias. Además, también existen excepciones dentro de los grupos de alto o -bajo riesgo. Por ejemplo, se sabe que el 5% de los melanomas "delgados" (profundidad de invasión <0.8 mm) metastatizan regionalmente, a pesar del hecho de que comúnmente se clasifican como de bajo riesgo según las variables clínico-patológicas estándar. Para poder distinguir mejor las lesiones de mayor riesgo del 95% de las lesiones biológicamente indolentes, se necesitan variables histológicas adicionales como la tasa mitótica del tumor (mitosis/mm²) y se han introducido métodos moleculares adicionales como la hibridación fluorescente in-situ (FISH). Desgraciadamente, estas técnicas sólo tuvieron un éxito parcial, resultando en el hecho de que hasta el 95% de las SLNB son negativas. Como resultado, existe un tratamiento excesivo con los criterios de diagnóstico actuales que lleva a que la mayoría de los pacientes se sometan a cirugías SLNB innecesarias y desarrollen efectos secundarios asociados que podrían haberse evitado. Viceversa, una fracción de los pacientes que actualmente no son remitidos a SLNB en función de sus factores clínico-patológicos, aún pueden presentar metástasis en los SLNs o metástasis a distancia en una etapa posterior. Por lo tanto, aunque el procedimiento SLNB es preciso para identificar ganglios positivos SLN en tejido resecado, la selección de pacientes elegibles para este procedimiento SLNB sigue siendo un desafío. Es un objetivo de la presente invención clasificar a los individuos según su riesgo de tener un SLN positivo para metástasis y, por lo tanto, necesitar el procedimiento SLNB. Un objetivo adicional de la presente invención es predecir el pronóstico de un individuo que padece melanoma cutáneo primario. Esta información es útil, por ejemplo, para poder determinar una estrategia de tratamiento óptima.

US 2016/222457 A1 divulga un método para diferenciar entre melanoma con alto y bajo riesgo de metástasis regional utilizando una firma de expresión que comprende ocho genes marcadores, incluidos ITGB3 y PLAT.

Sumario de la invención

La invención proporciona un método para clasificar a un individuo que padece melanoma cutáneo primario, que comprende determinar en una muestra de tejido de melanoma de dicho individuo una firma de expresión génica, en el que la firma de expresión génica comprende ITGB3, PLAT, GDF15 e IL8 y en el que el individuo se clasifica como que tiene SLN positivo para metástasis o se clasifica como que tiene SLN negativo para metástasis.

La invención también proporciona un método para determinar una programación de tratamiento y/o diagnóstico para un individuo que padece melanoma cutáneo, que comprende determinar en una muestra de tejido de melanoma de

dicho individuo una firma de expresión génica que comprende determinar el nivel de expresión de ITGB3, PLAT, GDF15 e IL8 y determinar una programación de tratamiento y/o diagnóstico basado en los niveles de expresión.

5 La invención también proporciona un método para predecir el pronóstico de un individuo que padece melanoma cutáneo primario, que comprende determinar en una muestra de tejido de melanoma de dicho individuo una firma de expresión génica, en el que la firma de expresión génica comprende ITGB3, PLAT, GDF15 e IL8 y el pronóstico del individuo se determina basado en los niveles de expresión génica.

La invención también proporciona un tratamiento contra el cáncer para utilizar en el tratamiento de un individuo que padece melanoma cutáneo primario, que comprende

10 - determinar en una muestra de tejido de melanoma de dicho individuo una firma de expresión génica, en el que la firma de expresión génica comprende ITGB3, PLAT, GDF15 e IL8,

- clasificar dicho individuo como que tiene SLN positivo para metástasis y/o un pronóstico pobre basado en la firma de expresión génica, y

- tratar a dicho individuo realizando una SLNB y/o proporcionar un tratamiento contra el cáncer a dicho individuo, en el que dicho tratamiento contra el cáncer se selecciona de

15 - quimioterapia, preferiblemente en el que dicha quimioterapia se selecciona de dacarbazina, temozolomida, carboplatino, paclitaxel, cisplatino, y vinblastina;

- radioterapia;

- inmunoterapia, preferiblemente en el que dicha inmunoterapia se selecciona de citocinas, inhibidores de puntos de control inmunes, e inmunoterapia oncolítica; y

20 - fármacos de terapia dirigida seleccionados de inhibidores de BRAF, preferiblemente seleccionados de vemurafenib o dabrafenib, e inhibidores de MEK, preferiblemente seleccionados de cobimetinib y trametinib.

Descripción breve de las figuras

25 Figura 1: Curvas ROC promedio para los clasificadores de regresión logística entrenados en DLCV en: 1) firma génica (modelo molecular) ITGB3, PLAT, SPP1, GDF15 e IL8, 2) variables clínico-patológicas (edad y profundidad de Breslow), 3) firma génica ITGB3, PLAT, SPP1, GDF15 e IL8 y variables clínico-patológicas combinadas. El eje-x representa la tasa de descubrimiento de falsos positivos (es decir, 1-especificidad), el eje-y la tasa de descubrimiento verdadera (es decir, sensibilidad).

30 Figura 2: Curvas ROC para la puntuación ITLP y para la firma génica SLN (denominada como el "modelo de regresión logística"), en toda la cohorte de 770 pacientes. El eje-x representa la tasa de descubrimiento de falsos positivos (es decir, 1-especificidad), el eje-y, la tasa de descubrimiento verdadera (es decir, sensibilidad).

Figura 3: Diagramas de caja del área bajo las curvas ROC para los diferentes subconjuntos de genes ITGB3, PLAT, SPP1, GDF15 e IL8, para el conjunto completo de 5 genes ITGB3, PLAT, SPP1, GDF15 e IL8 y para la firma ITLP.

35 Figura 4: Curvas ROC promedio para los clasificadores de regresión logística entrenados en DLCV en: 1) expresión génica, 2) variables clínico-patológicas, 3) expresión génica y variables clínico-patológicas combinadas. El eje-x representa la tasa de descubrimiento de falsos positivos (es decir, 1-especificidad), el eje-y, la tasa de descubrimiento verdadera (es decir, sensibilidad).

Figura 5: Diagramas de caja del área bajo las curvas ROC para los diferentes subconjuntos de genes y para el conjunto completo de 4 genes.

40 Figura 6: COMPARACIÓN DEL DESEMPEÑO GENERAL: CL versus GE versus GECL. Curvas ROC para los clasificadores de regresión logística entrenados en DLCV en: 1) expresión génica, 2) variables clínico-patológicas, 3) expresión génica y variables clínico-patológicas combinadas.

Figura 7: NPV versus SLNBRR. Valor predictivo negativo (NPV) versus la Tasa de Reducción del Ganglio Linfático centinela (SLNB RR) para los clasificadores de regresión logística entrenados en DLCV en: 1) expresión génica, 2) variables clínico-patológicas, 3) expresión génica y variables clínico-patológicas combinadas.

45 Figura 8: SUBCONJUNTOS DE GENES – DIAGRAMAS DE CAJA AUC. Diagramas de caja del área bajo la curva (AUC) de la curva ROC para clasificadores de regresión logística con subconjuntos de 2,3,4,5,6,7,8 genes entrenados en toda la cohorte.

Descripción detallada de las realizaciones divulgadas

50 La divulgación proporciona, en parte, métodos, kits, firmas génicas y medios para detectar dichas firmas génicas para realizar un análisis de una muestra de tejido tumoral de un melanoma cutáneo primario. En un aspecto, la divulgación

proporciona una "firma génica SLN". La firma génica SLN clasifica a un individuo que padece melanoma cutáneo primario, en particular, la firma génica clasifica el riesgo de que el individuo tenga un SLN positivo para metástasis y/o un pronóstico pobre. Esta evaluación de riesgo es útil para médicos y pacientes al momento de decidir si se justifica un procedimiento de SLNB y/o una estrategia de tratamiento alternativa. Esta evaluación también es útil a la hora de seleccionar pacientes para su inclusión en ensayos clínicos.

Como se utiliza en la presente, el SLN es el primer ganglio linfático (o conjunto de primeros ganglios linfáticos) que recibe drenaje linfático de un tumor y es el primer ganglio linfático (o conjunto de primeros ganglios linfáticos) donde es probable que se propague el cáncer. Un N-SLN es un ganglio linfático que no es el primer ganglio linfático que recibe drenaje linfático de un tumor. Dicho N-SLN suele ser un ganglio en la misma cuenca nodal o muy cerca del SLN.

En algunas realizaciones, la firma génica clasifica el riesgo de SLN positivo para metástasis. En algunas realizaciones, los métodos divulgados en la presente clasifican a un individuo como que tiene un SLN positivo para metástasis o como que tiene un SLN negativo para metástasis. En algunas realizaciones, las firmas génicas clasifican el pronóstico del individuo. Como se utiliza en la presente, pronóstico se refiere a la predicción de un resultado médico y puede basarse en medidas tales como supervivencia general, supervivencia específica de melanoma, supervivencia libre de recurrencia, supervivencia libre de reincidencia y supervivencia libre de reincidencia a distancia.

Una ventaja de la firma génica SLN es que puede reducir el número de procedimientos quirúrgicos para pacientes clasificados como SLN negativos para metástasis (y/o clasificados como que tienen un buen pronóstico). En particular, los pacientes con lesiones intermedias probablemente se habrían sometido a un procedimiento SLNB con una buena probabilidad de que el SLN fuera, de hecho, negativo para metástasis. La clasificación precisa de estos pacientes con la firma génica SLN evita la necesidad de un procedimiento SLNB y puede utilizarse para reemplazar la SLNB como el estándar actual de atención para lesiones intermedias. La reducción de SLNBs innecesarias reduce los costes generales de atención médica y reduce el número de pacientes que sufren complicaciones causadas por la extracción de SLNs. Además, la clasificación de un individuo como SLN positivo o negativo para metástasis también proporciona información de pronóstico y puede utilizarse para determinar una programación de tratamiento o diagnóstico.

Sorprendentemente, la firma génica SLN es capaz de predecir con mayor precisión el pronóstico que las biopsias de SLN de atención estándar (véase Ejemplo 7). Si bien no deseando quedar limitados por la teoría, una posible explicación en cuanto al poder pronóstico mejorado de la firma génica sobre las biopsias de SLN se relaciona con la limitación técnica de realizar tales biopsias (por ejemplo, identificar el ganglio linfático correcto para la biopsia, limitaciones de la detección de células tumorales, error humano en el procesamiento/clasificación de las muestras). Además de o como en la alternativa, la firma génica divulgada puede ser capaz de predecir la metástasis de SLN en un estadio antes de que pueda detectarse en una biopsia (por ejemplo, el tumor ha hecho metástasis y las células tumorales están en ruta hacia el SLN). En este sentido, la firma SLN se puede utilizar para reemplazar a la SLNB como criterio de inclusión en ensayos clínicos y/o tratamientos adicionales.

Una ventaja adicional de la firma génica SLN es que puede identificar pacientes con melanoma de grosor delgado que, según el estándar de atención actual, pueden no ser elegibles para SLNB basado los parámetros clínicos, pero basado en la firma génica tienen un alto riesgo de SLN positivo para metástasis. En particular, la firma génica aumentará en gran medida la identificación de SLN positivo para metástasis en pacientes con melanoma delgado (<0.8mm), que actualmente no son elegibles para procedimientos SLNB de acuerdo con las directrices. La detección y el tratamiento tempranos de estos pacientes aumentará la supervivencia general y libre de progresión para esta subpoblación de pacientes.

Los ejemplos divulgados en la presente demuestran que una firma de expresión génica (es decir, la firma génica SLN) comprende uno o más de los siguientes genes: ITGB3, PLAT, GDF15 e IL8, es útil para clasificar individuos y predecir el pronóstico y, en particular, para clasificar el SLN positivo o negativo para metástasis. Por consiguiente, en un aspecto la divulgación se proporciona una firma génica que comprende ITGB3, PLAT, GDF15 e IL8. En algunas realizaciones, la firma génica SLN comprende todos los genes siguientes: ITGB3, PLAT, SPP1, GDF15 e IL8. Preferiblemente, en el que la firma de expresión génica comprende GDF15, MLANA, PLAT, IL8, ITGB3, LOXL4, SERPINE2, y TGFBR1, más preferiblemente en el que la firma de expresión génica comprende GDF15, MLANA, PLAT, IL8, ITGB3, LOXL4, ADIPOQ, SERPINE2, y TGFBR1, más preferiblemente en el que la firma de expresión génica comprende GDF15, MLANA, PLAT, IL8, ITGB3, LOXL4, ADIPOQ, PRKCB, SERPINE2, ADAM12, LGALS1 y TGFBR1.

En algunas realizaciones, la firma génica comprende todo lo siguiente: GDF15, MLANA, PLAT, IL8, ITGB3, LOXL4, ADIPOQ, SERPINE2, y TGFBR1. Los inventores también demuestran que una firma génica que carece de ADIPOQ funciona de manera similar. En alguna realización, la firma génica comprende:

ITGB3, PLAT, GDF15, IL8, MLANA y uno o ambos de LOXL4 y SERPINE2;

ITGB3, PLAT, GDF15, IL8, MLANA y uno o ambos de SERPINE2 y TGFBR1;

ITGB3, PLAT, GDF15, IL8, y uno o ambos de MLANA y TGFBR1;

ITGB3, PLAT, GDF15, IL8, y uno o ambos de TGFBR1 y SERPINE2;

ITGB3, PLAT, GDF15, IL8, SERPINE2, y uno o ambos de LOXL4 y TGFBR1

ITGB3, PLAT, GDF15, IL8, LOXL4;

ITGB3, PLAT, GDF15, IL8, SERPINE2:

ITGB3, PLAT, GDF15, IL8, TGFBR1; o

5 ITGB3, PLAT, GDF15, IL8, MLANA.

En algunas realizaciones, la firma génica comprende todo lo siguiente: GDF15, MLANA, PLAT, IL8, ITGB3, LOXL4, ADIPOQ, PRKCB, SERPINE2, ADAM12, LGALS1 y TGFBR1. En algunas realizaciones, la firma génica comprende o todo lo siguiente: GDF15, MLANA, PLAT, IL8, ITGB3, LOXL4, PRKCB, SERPINE2, ADAM12, LGALS1 y TGFBR1.

10 En algunas realizaciones, la firma génica consta de los genes mencionados anteriormente. Como entenderá un experto, cuando la firma génica SLN consta de los genes mencionados anteriormente, los métodos para realizar un análisis pueden comprender medir la expresión de genes adicionales (por ejemplo, para normalización), pero sólo se utiliza la firma génica para clasificar a un individuo.

El gen ITGB3 codifica para la Integrina beta-3. Se puede encontrar una secuencia de ARNm de Homo sapiens ejemplar en la base de datos NCBI bajo NM_000212.2 (17 de junio de 2018).

15 El gen SPP1 codifica para la fosfoproteína 1 secretada. Se puede encontrar una secuencia de ARNm de Homo sapiens ejemplar en la base de datos NCBI bajo NM_001040058.1 (24 de junio de 2018). El gen GDF15 codifica para el factor de diferenciación del crecimiento 15. Se puede encontrar una secuencia de ARNm de Homo sapiens ejemplar en la base de datos NCBI bajo NM_004864.3 (17 de junio de 2018).

20 El gen IL8 codifica para interleucina 8. Se puede encontrar una secuencia de ARNm de homo sapiens ejemplar en la base de datos NCBI bajo AF043337.1 (1 de febrero de 2001).

El gen MLANA codifica para melan-A. Se puede encontrar una secuencia de ARNm de homo sapiens ejemplar en la base de datos NCBI bajo NM_005511 (20 de octubre de 2018).

El gen LOXL4 codifica para lisil oxidasa tipo 4. Se puede encontrar una secuencia de ARNm de homo sapiens ejemplar en la base de datos NCBI bajo NM_032211 (22 de noviembre de 2018).

25 El gen ADIPOQ codifica para adiponectina, C1Q y un dominio que contiene colágeno. Se puede encontrar una secuencia de ARNm de homo sapiens ejemplar en la base de datos NCBI bajo NM_004797 (2 de diciembre de 2018).

El gen PRKCB codifica para la proteína cinasa C beta. Se puede encontrar una secuencia de ARNm de homo sapiens ejemplar en la base de datos NCBI bajo NM_212535 (12 de noviembre de 2018).

30 El gen SERPINE2 codifica para serpina miembro 2 de la familia E. Se puede encontrar una secuencia de ARNm de homo sapiens ejemplar en la base de datos NCBI bajo NM_006216 (17 de noviembre de 2018). El gen ADAM12 codifica para el dominio 12 de metalopeptidasa ADAM. Se puede encontrar una secuencia de ARNm de homo sapiens ejemplar en la base de datos NCBI bajo NM_003474 (5 de agosto de 2018).

El gen LGALS1 codifica para galectina 1. Se puede encontrar una secuencia de ARNm de homo sapiens ejemplar en la base de datos NCBI bajo NM_002305 (22 de noviembre de 2018).

35 El gen TGFBR1 codifica para el receptor 1 del factor de crecimiento transformante beta. Se puede encontrar una secuencia de ARNm de homo sapiens ejemplar en la base de datos NCBI bajo NM_004612 (28 de octubre de 2018).

40 La divulgación proporciona además métodos para clasificar a un individuo que comprenden determinar en una muestra la firma génica SLN. En algunas realizaciones, el individuo puede clasificarse como que tiene SLN positivo para metástasis o SLN negativo para metástasis. En otra realización, el individuo puede clasificarse como de buen o pobre pronóstico. Anteriormente se reportó una firma génica asociada con metástasis del SLN (Meves et al. J Clinical Oncology 2015 33:2509-2516). El algoritmo utiliza las variables clínico-patológicas edad, profundidad de Breslow y ulceración, combinadas con la expresión génica del melanoma primario de cuatro genes ITGB3, LAMB1, PLAT y TP53, para predecir la metástasis del SLN. Como se muestra en la Figura 3, la firma génica SLN presente supera a la firma reportada anteriormente.

45 En un aspecto, la divulgación proporciona una "firma génica N-SLN". La firma génica N-SLN clasifica a un individuo que padece melanoma cutáneo primario, en particular, la firma génica clasifica el riesgo de que el individuo tenga un ganglio linfático no-centinela (N-SLN) positivo para metástasis. Esta evaluación de riesgo es útil para médicos y pacientes al momento de decidir las opciones de tratamiento y determinar el pronóstico del paciente.

50 En algunas realizaciones, la firma génica N-SLN clasifica el riesgo de N-SLN positivo para metástasis. Un individuo puede clasificarse como que tiene N-SLN positivo para metástasis o N-SLN negativo para metástasis. La invasión de

células tumorales a los ganglios linfáticos ubicados a distancia es un indicador de pronóstico pobre y sugiere el uso de formas de tratamiento más agresivas. Se espera que la detección y el tratamiento tempranos mejoren los resultados del paciente.

5 Los ejemplos divulgados en la presente demuestran que una firma de expresión génica (es decir, la firma génica N-SLN) que comprende uno o más de los siguientes genes: KRT14, SPP1, FN1, LOXL3, es útil para clasificar individuos y, en particular, para determinar el riesgo de N-SLN positivo para metástasis.

El gen KRT14 codifica para queratina 14. Se puede encontrar una secuencia de ARNm de Homo sapiens ejemplar en la base de datos NCBI bajo NM_000526.4 (17 de junio de 2018).

10 El gen FN1 codifica para fibronectina 1. Se puede encontrar una secuencia de ARNm de Homo sapiens ejemplar en la base de datos NCBI bajo NM_001306129.1 (3 de junio de 2018).

El gen LOXL3 codifica para lisil oxidasa tipo 3. Se puede encontrar una secuencia de ARNm de homo sapiens ejemplar en la base de datos NCBI bajo NM_001289165.1 (30 de junio de 2018).

15 La divulgación proporciona además métodos para clasificar a un individuo que comprenden determinar en una muestra la firma génica N-SLN. En algunas realizaciones, el individuo puede clasificarse como que tiene N-SLN positivo para metástasis o N-SLN negativo para metástasis. En algunas realizaciones, se proporcionan métodos que determinan tanto la firma génica SLN como la firma génica N-SLN.

20 El análisis de las firmas génicas divulgadas en la presente se puede realizar en cualquier individuo, incluyendo mamíferos y humanos, aunque se prefieren humanos. En algunas realizaciones, al individuo se le ha diagnosticado un melanoma cutáneo de T1-T3. En algunas realizaciones, el individuo aún no se ha sometido a una biopsia del SLN del melanoma primario, en particular cuando la firma génica es la firma génica SLN. La firma génica es particularmente útil para clasificar individuos con una edad joven, con una tasa mitótica alta (por ejemplo, superior a 2/mm²), antecedentes de enfermedad, antecedentes familiares de enfermedad con pronóstico pobre, y/o con invasión vascular linfática.

25 Las firmas de expresión génica son útiles para predecir el riesgo o la probabilidad de que las células tumorales hayan metastatizado en un SLN o N-SLN. Como está claro para un experto, la clasificación de un individuo se refiere a una probabilidad o "riesgo de" y no a que el 100% de todos los pacientes que se predice que estarán en riesgo tendrán de hecho metástasis detectables (denominadas sensibilidad o Acuerdo Porcentual Positivo) ni que el 0% de todos los pacientes que se predice que no tendrán metástasis en realidad no tendrán metástasis (denominadas como especificidad o Acuerdo Porcentual Negativo). Como se describe en los ejemplos, las firmas de expresión génica SLN y N-SLN exhiben altos niveles de desempeño tanto de sensibilidad como de especificidad. Como se divulga en los ejemplos, la firma génica SLN es capaz de predecir mejor el pronóstico de un individuo afectado con melanoma que la biopsia de SLN de atención estándar. La divulgación demuestra así que la firma de expresión génica es útil para predecir el pronóstico de un individuo.

35 Como lo sabe un experto, la carga de metástasis medida por el volumen de enfermedad metastásica puede diferir entre individuos. En algunas realizaciones, metástasis se refiere a la presencia de agrupaciones de células tumorales y no incluye ganglios linfáticos que solo contienen células tumorales aisladas o raras. En algunas realizaciones, metástasis se refiere a la presencia de agrupaciones de células que tienen al menos 0.1mm de diámetro con o sin extensión extra-capsular.

40 Está dentro del alcance de un experto en la técnica obtener una muestra adecuada para determinar la expresión génica. Las muestras adecuadas incluyen biopsias de lesiones de melanoma cutáneo primario. Dichas biopsias incluyen una lesión reseca (por ejemplo, extirpación por escisión-amplia de un tumor). Las muestras pueden procesarse o conservarse mediante cualquier medio conocido en la técnica que sea compatible con el perfilamiento de expresión génica. Por ejemplo, la muestra puede ser una biopsia de lesión de melanoma cutáneo primario embebida en parafina fijada con formalina, así como una muestra congelada.

45 Preferiblemente, la muestra es una muestra que contiene ARN. Los métodos generales para la extracción de ARNm son bien conocidos en la técnica y se describen en libros de texto estándar de biología molecular, incluyendo Ausubel et al. (1997) Current Protocols of Molecular Biology, John Wiley and Sons. Los métodos para la extracción de ARN de tejidos embebidos en parafina se divulgan, por ejemplo, en Rupp & Locker (1987) Lab Invest. 56:A67, y De Andrés et al., BioTechniques 18:42044 (1995). En particular, el aislamiento de ARN se puede realizar utilizando un kit de purificación, un conjunto de tampones y proteasa de fabricantes comerciales, tal como Qiagen, de acuerdo con las instrucciones del fabricante (QIAGEN Inc., Valencia, Calif.). Por ejemplo, el ARN total de células en cultivo se puede aislar utilizando mini-columnas Qiagen RNeasy. Hay varios kits de aislamiento de ARN disponibles comercialmente y pueden utilizarse en los métodos de la invención.

55 Los métodos divulgados en la presente comprenden determinar una firma de expresión génica. En particular, los métodos comprenden determinar un nivel de expresión génica. Los niveles de expresión génica se pueden determinar midiendo el nivel de expresión de ácido nucleico o proteína. Preferiblemente, se determina el nivel de expresión de ARNm. En algunas realizaciones, el ácido nucleico o la proteína se purifican de la muestra y la expresión génica se

mide mediante análisis de expresión de ácidos nucleicos o proteínas. El nivel de expresión de proteínas puede determinarse mediante cualquier método conocido en la técnica, incluyendo ELISAs, inmunocitoquímica, citometría de flujo, transferencia tipo Western, proteómica, y espectrometría de masas.

5 Preferiblemente, se determinan los niveles de expresión de ácidos nucleicos. El nivel de expresión de ácido nucleico puede determinarse mediante cualquier método conocido en la técnica, incluyendo RT-PCR, PCR cuantitativa, transferencia tipo Northern, secuenciación de genes, en particular secuenciación de ARN, y técnicas de perfilamiento de expresión génica. Los métodos representativos para el análisis de la expresión génica basado en secuenciación incluyen el Análisis en Serie de la Expresión Génica (SAGE), y el análisis de la expresión génica mediante secuenciación masiva de firmas paralelas (MPSS).

10 Preferiblemente, el ácido nucleico es ARN, tal como ARNm o pre-ARNm. Como entiende un experto, el nivel de expresión de ARN determinado puede detectarse directamente o puede determinarse indirectamente, por ejemplo, generando primero ADNc y/o amplificando el ARN/ADNc. En algunas realizaciones, se obtiene una muestra de melanoma primario; se extrae ARN de la muestra de tejido; le sigue la transcripción reversa de un transcrito de ARN de los genes de interés (por ejemplo, biomarcadores y genes constitutivos) para producir ADNc de los transcritos de ARN; y amplificar los ADNc para producir amplicones a partir de los ADNc para determinar los niveles de expresión de los transcritos de ARN.

En algunas realizaciones, la expresión génica puede determinarse mediante análisis de expresión génica NanoString. NanoString es un método multiplexado para detectar la expresión génica y proporciona un método para la medición directa de ARNm sin el uso de transcripción o amplificación. NanoString y aspectos del mismo se describen en Geiss et al., "Direct multiplexed measurement of gene expression with color-coded probe pairs" *Nature Biotechnology* 26, 317 - 325 (2008);

20 No es necesario que el nivel de expresión sea un valor absoluto, sino que puede ser un valor de expresión normalizado o un valor relativo. Por ejemplo, los niveles de expresión pueden normalizarse contra la expresión de genes constitutivos o referencia. Dichos genes incluyen ABCF1, ACTB, ALAS1, CLTC, G6PD, GAPDH, GUSB, HPRT1, LDHA, PGK1, POLR1B, POLR2A, RPL19, RPLPO, SDHA, TBP, y TUBB.

La normalización también es útil cuando la expresión se determina en función de datos de microarreglos. La normalización permite corregir la variación dentro de los microarreglos y a través de las muestras para que los datos de diferentes chips se puedan analizar simultáneamente. El algoritmo del análisis robusto de multi-arreglos (RMA) se puede utilizar para pre-procesar los datos del conjunto de sondas en los niveles de expresión génica para todas las muestras. (Irizarry R A, et al., *Biostatistics* (2003) e Irizarry R A, et al., *Nucleic Acids Res.* (2003)). Además, también se puede emplear el algoritmo de preprocesamiento predeterminado de Affymetrix (MAS 5.0). Los métodos adicionales para normalizar los datos de expresión se describen en US20060136145.

En algunas realizaciones, los niveles de expresión se determinan utilizando PCR en tiempo-real (es decir, PCR cuantitativa o qPCR). En la PCR en tiempo-real (qPCR), las reacciones se caracterizan por el punto en el tiempo durante la ciclación en el que se detecta por primera vez la amplificación de una diana, en lugar de la cantidad de diana acumulada después de un número fijo de ciclos. Este punto en el que se detecta la señal por primera vez se denomina ciclo de umbral (Ct).

En algunas realizaciones, la expresión de la firma génica se cuantifica relativa entre sí mediante la normalización contra la expresión de genes constitutivos restando el Ct de la firma génica del Ct promediado de los genes constitutivos. En algunas realizaciones, estos valores de ΔCt se combinan luego con la edad del paciente y la profundidad de Breslow de la lesión del melanoma en un algoritmo para calcular la predicción de metástasis del SLN. En algunas realizaciones, los genes constitutivos utilizados para la normalización son ACTB, RPLP0 y RPL8. Sin embargo, se pueden utilizar otros genes constitutivos. Las proporciones de las señales de expresión génica se pueden combinar posteriormente con variables clínicas en un algoritmo para calcular la predicción del resultado de un paciente SLNB. Los resultados se expresan como una clasificación binaria (negativa o positiva). Un resultado "negativo" indicaría que el individuo tiene un riesgo bajo de SLN positivo para metástasis o más bien que el individuo tiene un buen pronóstico, mientras que un resultado "positivo" indicaría que el individuo tiene un alto riesgo de SLN positivo para metástasis, o más bien un pronóstico pobre.

Los métodos descritos en la presente clasifican a un individuo basándose en firmas de expresión génica. En algunas realizaciones, la expresión diferencial de uno o más genes de la firma en un individuo indica que el individuo está en riesgo de metástasis, o más bien, indica el pronóstico del individuo. Como se utiliza en la presente, "expresado-diferencialmente" significa que el nivel de expresión medido en un sujeto difiere significativamente de una referencia. La referencia puede ser un valor único o un valor numérico. Está dentro del alcance de una persona experta determinar el valor de referencia apropiado. En algunas realizaciones, el valor de referencia es un valor predeterminado. En algunas realizaciones, el valor de referencia es el promedio del valor de expresión en una clase de paciente particular. Por ejemplo, el valor de referencia puede ser el promedio del valor de expresión en una clase de pacientes que tienen metástasis de SLN clínicamente confirmada (o para la firma N-SLN, pacientes que tienen metástasis de N-SLN clínicamente confirmada). Un valor de referencia también puede tener la forma de una ecuación o derivarse de ella.

Está dentro del alcance de un experto en la técnica determinar si el nivel de expresión en el paciente difiere "significativamente" de una referencia.

5 En una realización ejemplar, el valor de referencia se determina a partir de una cohorte de pacientes con melanoma que se sometieron a una SLNB como se describe en los ejemplos. Está claro para un experto que también se pueden utilizar datos de estudios similares.

10 La fuerza de la correlación entre el nivel de expresión de un gen expresado diferencialmente y una clase de respuesta específica del paciente puede determinarse mediante una prueba estadística de significancia. Por ejemplo, se puede usar una prueba de chi cuadrada para asignar un valor de chi cuadrada a cada marcador expresado diferencialmente, indicando la fuerza de la correlación de la expresión de ese marcador con una clase de respuesta específica del paciente. De manera similar, la métrica de la estadística-T y la métrica de Wilkins proporcionan un valor o puntuación indicativo de la fuerza de la correlación entre la expresión del marcador y su clase de respuesta específica del paciente. Además, se pueden utilizar herramientas de análisis SAM o PAM para determinar la fuerza de las correlaciones.

15 En algunas realizaciones, la firma de expresión génica de un individuo se compara con la firma de expresión de referencia para determinar si la firma de expresión génica de un individuo es lo suficientemente similar al perfil de referencia. Alternativamente, la firma de expresión génica de un individuo se compara con una pluralidad de firmas de expresión de referencia para seleccionar el perfil de expresión de referencia que sea más similar al perfil de expresión génica de un individuo. Se puede utilizar cualquier método conocido en la técnica para comparar dos o más conjuntos de datos para detectar similitudes entre ellos para comparar la firma de expresión génica de un individuo con los perfiles de expresión de referencia.

20 En aprendizaje automático y estadísticas, la clasificación se refiere a identificar a qué conjunto de categorías pertenece una nueva observación, sobre la base de un conjunto de datos de entrenamiento que contienen observaciones (o instancias) cuya membresía de categoría se conoce. Un algoritmo que implementa la clasificación, especialmente en una implementación concreta, se conoce como clasificador. Se conocen muchos clasificadores en la técnica, con límites de clasificador lineales o no-lineales, tales como, pero no limitados a: ClaNC, clasificador de media más cercana, método de votación ponderada, clasificador de Bayes simple, análisis discriminante lineal (LDA), análisis discriminante cuadrático (QDA), máquinas de vectores de soporte (SVM), o clasificador k- del vecino más cercano (k-nn). En una realización preferida, se utiliza un clasificador de regresión logística. En los ejemplos se describen realizaciones ejemplares que implementan un clasificador de regresión logística.

30 Como entiende un experto, el entrenamiento de una firma de expresión génica se puede realizar para favorecer la sensibilidad o la especificidad. La sensibilidad se refiere a la proporción de positivos reales que se identifican correctamente como tales y se desea una alta sensibilidad para evitar falsos negativos (por ejemplo, pacientes clasificados como negativos para metástasis que en realidad son positivos).

35 La especificidad se refiere a la proporción de negativos reales que se identifican correctamente como tales y se desea una alta especificidad para evitar falsos positivos (por ejemplo, pacientes clasificados como positivos para metástasis que en realidad son negativos). Preferiblemente, los clasificadores están entrenados para una alta sensibilidad con la finalidad de identificar individuos con metástasis.

40 En algunas realizaciones, los métodos para clasificar a un individuo utilizan además la edad del individuo y/o la profundidad de Breslow del tumor. Opcionalmente, se puede determinar la ulceración y/o la tasa de mitosis. La profundidad de Breslow se mide desde la parte superior de la capa granular de la epidermis (o, si la superficie está ulcerada, desde la base de la úlcera) hasta la célula invasiva más profunda a lo largo de la base amplia del tumor (dermal/subcutánea). La ulceración se refiere al desprendimiento de tejido muerto y se cree que refleja un rápido crecimiento del tumor, lo que lleva a la muerte de las células en el centro del melanoma. La tasa mitótica se puede medir examinando el tumor extirpado y contando el número de células que presentan mitosis. Cuanto mayor sea el recuento mitótico, más probable es que el tumor haya metastatizado. En realizaciones particulares, se utiliza un modelo combinado que comprende la firma génica que comprende GDF15, MLANA, PLAT, IL8, ITGB3, LOXL4, SERPINE2 y TGFBR1 y las variables clínicas de edad y profundidad de Breslow. En algunas realizaciones, la firma génica comprende adicionalmente AIDPOQ.

45 La divulgación también proporciona kits para determinar las firmas de expresión génica divulgadas en la presente. En algunas realizaciones, los kits comprenden pares de cebadores para realizar qPCR en las firmas génicas divulgadas en la presente. En algunas realizaciones, los kits comprenden un ácido nucleico de control para uno o más, preferiblemente para cada, par de cebadores. Preferiblemente, el ácido nucleico de control es ADNc y más preferiblemente el ADNc corresponde a una secuencia que abarca al menos un límite intrón/exón del respectivo gen. Dicho ADNc es útil para distinguir la expresión génica de la contaminación genómica. En algunas realizaciones, uno o más cebadores del par de cebadores están modificados químicamente. Dichos cebadores modificados incluyen cebadores marcados de forma fluorescente o radiactiva.

50 Los resultados de los análisis de expresión génica divulgados en la presente son útiles para determinar una programación de diagnóstico. Por ejemplo, un individuo clasificado como que tiene SLN positivo para metástasis o un pronóstico pobre puede someterse a una SLNB. En algunas realizaciones, la inmunoterapia se administra a un

individuo que se predice que será positivo para SLN o más bien se predice que tendrá un pronóstico pobre. Una lectura posterior de SLNB puede ser una medida de la respuesta a la inmunoterapia.

5 Basado en las firmas de expresión génica, se puede determinar un régimen de tratamiento adecuado. Tal como se utilizan en la presente, los términos "tratamiento", "tratar" y "tratando" se refieren a revertir, aliviar, retrasar la aparición de, o inhibir el progreso del melanoma, o uno o más síntomas del mismo. En algunas realizaciones, los individuos clasificados como que tienen SLN metastásico positivo o más bien con un pronóstico pobre pueden ser tratados con SLNB. La ubicación del SLN se puede determinar basándose en la ubicación del melanoma y/o con métodos tales como "mapeo de SLN", como lo conoce un experto y se describe en la presente.

10 Un N-SLN positivo para metástasis se puede tratar realizando un procedimiento quirúrgico, por ejemplo, una disección quirúrgica de los ganglios linfáticos. N-SLN metastásico se puede tratar con una disección completa del ganglio linfático y/u otras terapias para tratar el melanoma. En algunas realizaciones se administra un tratamiento contra el cáncer al individuo. En algunas realizaciones, se administra un "tratamiento adyuvante" a un individuo. El tratamiento adyuvante, como se utiliza en la presente, se refiere a la administración de uno o más fármacos a un paciente después de la resección quirúrgica de uno o más tumores cancerosos, donde toda la enfermedad detectable y resecable (por ejemplo, cáncer) ha sido eliminada del paciente, pero donde permanece un riesgo estadístico de reincidencia. El tratamiento adyuvante es útil para disminuir la probabilidad o la gravedad de la recurrencia de la enfermedad.

15 Las terapias conocidas contra el melanoma que pueden estar indicadas en función de las firmas de expresión génica incluyen:

20 - Quimioterapia: por ejemplo, dacarbazina (DTIC), temozolomida (Temodal), carboplatino (Paraplatin, Paraplatin AQ), paclitaxel (Taxol), cisplatino (Platinol AQ), y vinblastina y (Velbe);

- Medicamentos de terapia dirigida: por ejemplo, inhibidores de BRAF (vemurafenib (Zelboraf) y dabrafenib (Tafinlar)) e inhibidores de MEK (cobimetinib (Cotellic) y trametinib (Mekinist));

- Radioterapia;

25 - Inmunoterapia: por ejemplo, citocinas (por ejemplo, interferón alfa-2b o interleucina-2), inhibidores de puntos de control inmune (por ejemplo, ipilimumab (Yervoy), nivolumab (Opdivo), Pembrolizumab (Keytruda)), o inmunoterapia oncolítica.

La administración de una terapia farmacológica adecuada puede realizarse por cualquier vía apropiada. Las vías adecuadas incluyen oral, rectal, nasal, tópica (incluyendo bucal y sublingual), vaginal, y parenteral (incluyendo subcutánea, intramuscular, intravenosa, intradérmica, intratecal, y epidural).

30 Como se utiliza en la presente, "comprender" y sus conjugaciones se utiliza en su sentido no-limitante para significar que los elementos que siguen a la palabra están incluidos, pero los elementos que no se mencionan específicamente no se excluyen. Además, el verbo "consistir" puede sustituirse por "consistir esencialmente en", lo que significa que un compuesto o compuesto adjunto tal como se define en la presente puede comprender componente(s) adicionales a los específicamente identificados, sin que dichos componente(s) adicionales alteren la característica única de la invención.

35 Los artículos "un" y "uno, una" se utilizan en la presente para referirse a uno o a más de uno (es decir, a al menos uno) del objetivo gramatical del artículo. A modo de ejemplo, "un elemento" significa un elemento o más de un elemento.

40 La palabra "aproximadamente" o "alrededor de" cuando se utiliza en asociación con un valor numérico (aproximadamente 10, alrededor de 10) significa preferiblemente que el valor puede ser el valor dado de 10 más o menos el 1% del valor.

La invención se explica con más detalle en los siguientes ejemplos. Estos ejemplos no limitan el alcance de la invención, sino que simplemente sirven para aclarar la invención.

Ejemplos

Ejemplo 1: Firma génica para predecir el estatus de metástasis del SLN

45 Ensamblamos una cohorte de 813 pacientes consecutivos con melanoma que se sometieron a una biopsia SLN en centros de atención terciaria. Los resultados de interés fueron SLN y las metástasis identificadas histológicamente en la biopsia del SLN. La expresión de 29 genes de respuesta estromal que promueven metástasis en tejido de biopsia de diagnóstico de melanoma primario se realizó mediante reacción en cadena de la polimerasa. Se aplicó regresión logística con regularización a variables clínico-patológicas y datos moleculares en un esquema de entrenamiento-
50 validación cruzada-de-doble-bucle (DLCV).

Conjunto de datos clínicos

De las variables clínicas anteriores, solo 6 muestran una diferencia significativa entre SLNB positivo y SLNB negativo:

- Edad,
- Profundidad de Breslow
- Ulceración
- Tasa mitótica
- 5 • Nivel de Clark
- Invasión angio-linfática

De las 6 variables, decidimos considerar solamente edad, profundidad de Breslow, ulceración y tasa mitótica, pero no el nivel de Clark ni la invasión angio-linfática, ya que estas últimas variables no siempre están disponibles y su calidad puede diferir dependiendo del proveedor de atención médica que realiza la SLNB.

10 Cuantificamos la expresión génica por delta Ct y no por número de copias como en Meves et al. 2015, ya que no hay diferencia significativa en el desempeño al utilizar el número de copia con nuestra firma, sino solo un aumento en la carga experimental. La corrección de fondo KRT14 también se ha eliminado en comparación con Meves et al. 2015. La llamada normalización de ITLP también se ha eliminado, ya que se basa en una anulación ad-hoc de la expresión medida y requiere muchos parámetros arbitrarios, principalmente en forma de umbrales.

15 Entre los pacientes con biopsia positiva del ganglio linfático centinela, la carga de metástasis, medida por el volumen de enfermedad metastásica, puede diferir significativamente:

Volumen	Definición
1	Células tumorales aisladas o células tumorales raras solamente
2	Agrupaciones de células < 0.1 mm de diámetro
3	Agrupaciones de células >= 0.1 mm w/o sin extensión extra-capsular
4	Agrupaciones de células >= 0.1 mm con extensión extra-capsular
9	Desconocido

20 Sigue siendo controversial si las muestras que tienen agrupaciones de células de menos de 0.1 mm de diámetro deben considerarse positivas para metástasis y, desde un punto de vista clínico, generalmente se consideran negativas. Por lo tanto, decidimos excluir del conjunto de entrenamiento del clasificador a 43 pacientes de 813 con un volumen de enfermedad metastásica igual a 1 o 2, y evaluamos por separado el desempeño del clasificador en este conjunto. Se midieron los siguientes 29 genes:

KRT14, MLANA, MITF, ITGB3, PLAT, LAMB1, TP53, AGRN, THBS2, PTK2, SPP1, COL4A1, CDKN1A, CDKN2A, PLOD3, GDF15, FN1, TNC, THBS1, CTGF, LOXL1, LOXL3, ITGA5, ITGA3, ITGA2, CSRC, CXCL1, IL8, LAMB.

25 **Medidas de desempeño**

Para evaluar el desempeño de un clasificador se han construido tablas de contingencia. De estas tablas de contingencia se han derivado dos criterios, es decir, el PPA (Acuerdo Porcentual Positivo) y el NPA (Acuerdo Porcentual

30 Negativo): que se definen como: $PPA = 100 \frac{TP}{TP+FN}$ y $NPA = 100 \frac{TN}{TN+FP}$ donde TP representa el número de Verdaderos Positivos, FN el número de Falsos Negativos, TN el número de Verdaderos Negativos, y FP el número de Falsos Positivos. Las fórmulas para los criterios PPA y NPA son equivalentes a la sensibilidad y especificidad, respectivamente. Sin embargo, dado que se realiza una comparación relativa a una referencia que no es el estándar de oro (datos FISH), se utilizarán los términos PPA y NPA.

Para optimizar el entrenamiento de un clasificador, se requiere un único criterio de desempeño que debe maximizarse (o minimizarse en caso de un criterio de error). Con esta finalidad se utilizará el promedio del PPA y NPA:

35
$$p = \frac{PPA + NPA}{2} = 100 \frac{\frac{FN}{TP + FN} + \frac{FP}{FP + TN}}{2}$$

Donde p indica el desempeño. P=50 implica desempeño aleatorio, y p=100 implica clasificación perfecta.

Otras medidas de desempeño:

- Valor predictivo negativo
- Valor predictivo positivo

- Exactitud
 - Exactitud equilibrada
 - Proporción de probabilidad logarítmica para un resultado negativo
 - Proporción de probabilidad logarítmica para un resultado positivo
- 5 • Área bajo la curva ROC.

Clasificadores

Todos los clasificadores están entrenados en los niveles de expresión de los 29 genes, en las variables clínico-patológicas y en ambas.

Clasificador: Regresión logística con máxima probabilidad penalizada

10 Utilizamos clasificadores de regresión logística implementados en el paquete R glmnet. Los parámetros se estiman mediante estimación de máxima probabilidad con un término de penalización de norma-L1 (regularización LASSO) para obtener una representación parsimoniosa.

Validación cruzada de doble bucle del clasificador.

15 Wessels et al. [Wessels et al. Bioinformatics, Volume 21, Issue 19, 2005, Pages 3755-3762] describieron un marco de referencia de aplicación general para crear clasificadores de diagnóstico a partir de datos de alto rendimiento mediante validación cruzada de doble bucle (DLCV). Un ejercicio DLCV permite al desarrollador estimar/predecir el desempeño (en términos de error generalizado) del clasificador para su futura aplicabilidad en datos independientes del conjunto de datos de entrenamiento. Esta metodología se adoptó combinada con filtrado directo como selector de características, la estadística t como criterio para evaluar los genes individuales, y diferentes clasificadores.

20 El procedimiento de entrenamiento y validación se realizó empleando 100 repeticiones de validación cruzada de 3 veces en el bucle externo (validación), y validación cruzada de 10 veces en el bucle interno. En el bucle interno, el algoritmo aprende el parámetro lambda óptimo para la regularización LASSO. En todos los puntos, las divisiones de datos se estratificaron con respecto a las probabilidades de clase anteriores.

El método de validación cruzada de doble bucle se puede describir en unos pocos pasos:

- 25 1. Para cada repetición, los datos se dividen (estratifican) en 3 partes (divisiones diferentes para cada repetición).
2. Para cada vez, se utilizan 2 partes para el bucle interno (conjunto de entrenamiento); el tercero se utiliza en el bucle externo para la validación (conjunto de validación).
3. En los datos del conjunto de entrenamiento, se realiza una validación cruzada de 10 veces para estimar la lambda óptima utilizada en el término de penalización LASSO (construcción de una curva de aprendizaje).
- 30 4. A continuación, se entrena un clasificador en el conjunto de entrenamiento completo, utilizando la lambda óptima superior.
5. Finalmente, el desempeño de ese clasificador se evalúa en el conjunto de validación.
- 35 6. Una vez completadas todas las repeticiones, se crea un clasificador final utilizando todas las muestras con la lambda óptima promedio. El número promedio de n que se obtuvo se utiliza para entrenar al clasificador. Este clasificador luego se aplicará en un conjunto de validación externo. Normalmente, los conjuntos de datos están desequilibrados con respecto a las clases anteriores. Por lo tanto, la precisión equilibrada es un mejor desempeño de clasificación que la precisión, ya que toma en cuenta las clases anteriores. En cada iteración, en el bucle interno, utilizamos la puntuación de Brier como criterio de desempeño.

Clasificador basado en Expresión génica (GE)

40 Los parámetros del modelo clasificador logístico basado en la expresión génica fueron los siguientes:

(Intercepción)	0.78199268
ITGB3	-0.19497568
PLAT	-0.12305874
SPP1	-0.00831690
GDF15	-0.04554275
IL8	-0.03306434

5 La Tabla 1 representa el desempeño del clasificador final entrenado en todo el clasificador de cohorte de 770 pacientes para cuatro puntos operativos diferentes: 1) max bACC: precisión equilibrada máxima, 2) SEeqSP, sensibilidad igual a especificidad, 3) NPV97, NPV establecido en 0.97 en el entrenamiento, 4) LRNn025, proporción de probabilidad logarítmica para el resultado negativo de la prueba establecido en 0.25 en el entrenamiento. Si el coeficiente es positivo, un valor más alto implica un mayor riesgo. Si el coeficiente es negativo, entonces el valor reducido implica una disminución del riesgo. Las variables con coeficientes mayores (absolutos) tienen una contribución mayor.

10 La Tabla 2 representa el desempeño de los clasificadores entrenados en DLCV, promediado durante más de 100 repeticiones, para cuatro puntos operativos diferentes: 1) max bACC: precisión equilibrada máxima, 2) SEeqSP, sensibilidad igual a especificidad, 3) NPV97, NPV establecido en 0.97 en el entrenamiento, 4) LRNn025, proporción de probabilidad logarítmica para el resultado negativo de la prueba establecido en 0.25 en el entrenamiento.

Clasificador basado en Variables Clínicas (CL)

Los parámetros del modelo clasificador logístico basado en variables clínicas fueron los siguientes:

(Intercepción)	-1.86716485
Edad	-0.01919991
Profundidad de Breslow	0.72079160
Ulceración-sí	0.14301462

15 En la Tabla 3. El parámetro "Edad" se introduce en años y la "Profundidad de Breslow" en milímetros. La ulceración es una variable Booleana (sí/no). La tabla representa el desempeño del clasificador final entrenado en todo el clasificador de cohorte de 770 pacientes para cuatro puntos operativos diferentes: 1) max bACC: precisión equilibrada máxima, 2) SEeqSP, sensibilidad igual a especificidad, 3) NPV97, NPV establecido en 0.97 en el entrenamiento, 4) LRNn025, proporción de probabilidad logarítmica para el resultado negativo de la prueba establecido en 0.25 en el entrenamiento. La Tabla 4 representa el desempeño de los clasificadores entrenados en DLCV, promediado durante más de 100 repeticiones, para cuatro puntos operativos diferentes: 1) max bACC: precisión equilibrada máxima, 2) SEeqSP, sensibilidad igual a especificidad, 3) NPV97, NPV establecido en 0.97 en el entrenamiento, 4) LRNn025, proporción de probabilidad logarítmica para el resultado negativo de la prueba establecido en 0.25 en el entrenamiento.

Clasificador basado en Expresión Génica y Variables Clínicas (GECL)

Los parámetros del modelo clasificador logístico basado en variables clínicas y expresión génica fueron los siguientes:

(Intercepción)	0.59189841
Edad	-0.01356918
Profundidad de Breslow	0.4722686
ITGB3	-0.15443565
PLAT	-0.13580062
SPP1	-0.00778917
GDF15	-0.05924340
IL8	-.003781148

25 La Tabla 5 representa el desempeño del clasificador final entrenado en todo el clasificador de cohorte de 770 pacientes para cuatro puntos operativos diferentes: 1) max bACC: precisión equilibrada máxima, 2) SEeqSP, sensibilidad igual a especificidad, 3) NPV97, NPV establecido en 0.97 en el entrenamiento, 4) LRNn025, proporción de probabilidad logarítmica para el resultado negativo de la prueba establecido en 0.25 en el entrenamiento.

30 La Tabla 6 representa el desempeño de los clasificadores entrenados en DLCV, promediado durante más de 100 repeticiones, para cuatro puntos operativos diferentes: 1) max bACC: precisión equilibrada máxima, 2) SEeqSP, sensibilidad igual a especificidad, 3) NPV97, NPV establecido en 0.97 en el entrenamiento, 4) LRNn025, proporción de probabilidad logarítmica para el resultado negativo de la prueba establecido en 0.25 en el entrenamiento.

La Figura 1 representa las curvas ROC para los clasificadores de regresión logística entrenados en DLCV en: 1) expresión génica, 2) variables clínico-patológicas, 3) expresión génica y variables clínico-patológicas combinadas.

La Tabla 7 representa el desempeño promedio de los clasificadores entrenados en DLCV en: 1) expresión génica ("GE"; es decir, firma génica ITGB3, PLAT, SPP1, GDF15 e IL8, 2) variables clínico-patológicas ("CL"; es decir, edad y profundidad de Breslow), 3) expresión génica y variables clínico-patológicas combinadas ("GECL"). Se han considerado tres puntos operativos diferentes: 1) max bACC: precisión equilibrada máxima, 2) SEeqSP, sensibilidad igual a especificidad, 3) NPV97, NPV establecido en 0.97 en el entrenamiento.

Ejemplo 2. Ejemplo comparativo

Clasificador basado en puntuación ITLP (ITGB3, LAMB1, PLAT y TP53)

La Tabla 8 representa el desempeño de la puntuación ITLP en toda la cohorte de 770 pacientes.

Clasificador basado en puntuación ITLP y Variables Clínicas

Los parámetros del modelo clasificador logístico basado en variables clínicas y expresión génica fueron los siguientes:

(Intercepción)	-2.07660948
Edad	-0.0111771
Profundidad de Breslow	0.52425831
ITLP	0.50016404

La Tabla 9 representa el desempeño del clasificador final entrenado en todo el clasificador de cohorte de 770 pacientes para cuatro puntos operativos diferentes: 1) max bACC: precisión equilibrada máxima, 2) SEeqSP, sensibilidad igual a especificidad, 3) NPV97, NPV establecido en 0.97 en el entrenamiento, 4) LRNn025, proporción de probabilidad logarítmica para el resultado negativo de la prueba establecido en 0.25 en el entrenamiento.

La Tabla 10 representa el desempeño de los clasificadores entrenados en DLCV, promediado durante más de 100 repeticiones, para cuatro puntos operativos diferentes: 1) max bACC: precisión equilibrada máxima, 2) SEeqSP, sensibilidad igual a especificidad, 3) NPV97, NPV establecido en 0.97 en el entrenamiento, 4) LRNn025, proporción de probabilidad logarítmica para el resultado negativo de la prueba establecido en 0.25 en el entrenamiento.

Ejemplo 3. Análisis comparativo

Firma génica ITLP versus ITGB3, PLAT, GDF15, SPP1 e IL8

La Figura 2 representa las curvas ROC para la puntuación ITLP y para la firma génica ITGB3, PLAT, GDF15, SPP1 e IL8 ("denominada regresión logística" en la figura). La firma génica ITGB3, PLAT, GDF15, SPP1 e IL8 supera claramente a la firma ITLP.

Ejemplo 4: Desempeño de los subconjuntos de genes.

Los ejemplos anteriores utilizaron 5 genes: ITGB3, PLAT, GDF15, SPP1 e IL8 para la firma genética. Investigamos el desempeño de todos los subconjuntos posibles de 2, 3 y 4 genes. El número de subconjuntos de una dimensión específica que se pueden seleccionar del número total de genes para las siguientes firmas es el siguiente: 10 subconjuntos de firmas con dos genes, 10 subconjuntos de firmas con tres genes, 5 subconjuntos de firmas con cuatro genes y una firma que comprende a los 5. Evaluamos el desempeño en términos del área bajo la curva ROC y lo comparamos con la firma ITLP. El AUC (o el intervalo de la misma) fue 0.68 para ITLP, 0.72-0.75 para todos los subconjuntos de 2, 0.74-0.77 para todos los subconjuntos de 3, 0.76-0.77 para todos los subconjuntos de 4, y 0.77 para la firma de 5 genes. Esto también se representa en la Figura 3. En consecuencia, todas las firmas génicas que comprenden al menos dos de los siguientes genes: ITGB3, PLAT, GDF15, SPP1 e IL8 superan a la firma ITLP.

Ejemplo 5: Desempeño de las 43 muestras con bajo volumen de enfermedad metastásica. Los pacientes con bajo volumen de enfermedad metastásica (volumen 1 y 2) han sido excluidos, en primera instancia, de la cohorte utilizada para entrenar al clasificador. Sigue siendo controversial si las muestras que tienen agrupaciones de células de menos de 0.1 mm de diámetro deben considerarse positivas para metástasis y, desde un punto de vista clínico, generalmente se consideran negativas. En el presente estudio, 43 pacientes fueron inicialmente excluidos del análisis dado que tienen un volumen 1 o 2. La aplicación del clasificador ITGB3, PLAT, GDF15, SPP1 e IL8 en estos pacientes resultó en la clasificación de 29 como positivos y 14 como negativos.

Ejemplo 6: Análisis de clasificación errónea

Falso negativo. Las muestras positivas clasificadas erróneamente proceden en su mayoría de pacientes con melanomas delgados (menos de 2 mm), sin ulceración, y sin invasión angio-linfática. En otras palabras, se trata de pacientes que presentan a priori un riesgo muy bajo de desarrollar metástasis. Hay algunas muestras clasificadas erróneamente en las 100 repeticiones del algoritmo.

- 5 Falso positivo. Las muestras negativas clasificadas erróneamente proceden en su mayoría de pacientes con melanomas gruesos (más de 2 mm), con ulceración, y con invasión angio-linfática. En otras palabras, se trata de pacientes que presentan a priori un alto riesgo de desarrollar metástasis. Hay algunas muestras clasificadas erróneamente en las 100 repeticiones del algoritmo.

Distribución de salida del clasificador

- 10 La distribución de la probabilidad predicha es uni-modal, no Gaussiana con una cola derecha larga. El umbral utilizado para elegir el punto operativo está cerca de la media de la distribución. Las probabilidades estimadas no superan 0.6.

Ejemplo 7: Asociación pronóstica de la firma génica SLN.

- 15 Se generaron estimaciones de supervivencia de Kaplan-Meier para tres tipos de supervivencia para el clasificador SLN que comprende los genes ITGB3, PLAT, SPP1, GDF15 e IL8 (denominado en el ejemplo como "GECL") (Tablas 19-21 (a)), estatus de SLNB (Tablas 19-21 (a)), y la combinación de estos dos (Tabla 22-24 (a)). Como se sabe en la literatura, un estatus positivo de SLNB se asocia con una supervivencia inferior. Sorprendentemente, el modelo de GECL proporcionó una separación mucho más fuerte en las estimaciones de supervivencia, lo que también es evidente a partir de las proporciones de riesgo más grandes (Tabla 25 (a)). Además, en un análisis multivariado, el clasificador GECL también tiene una proporción de riesgo mucho mayor y un valor p más significativo (véase la Tabla 26 (a)). De particular interés es la muy buena supervivencia específica de melanoma del grupo GECL negativo, con estimaciones de supervivencia de 0.966 a los 160 meses (Tabla 19 (a)).

- 20 De manera similar, se generaron estimaciones de supervivencia de Kaplan-Meier para tres tipos de supervivencia para el clasificador SLN que comprende los genes GDF15, MLANA, PLAT, IL8, ITGB3, LOXL4, ADIPOQ, PRKCB, SERPINE2, ADAM12, LGALS1, y TGFBR1 (mencionados en el ejemplo como "GECL") (Tablas 19-21), estatus de SLNB (Tablas 19-21), y la combinación de estos dos (Tabla 22-24). Las Tablas 19-26(b) describen los resultados con un NPV establecido en 0.97 en el entrenamiento, mientras que las Tablas 19-26(c) describen los resultados con un NPV establecido en 0.98 en el entrenamiento. Véase el ejemplo 8 para obtener más información sobre este clasificador.

- 25 La combinación de los resultados de la salida del clasificador de GECL con el estatus de SLNB proporcionó cuatro grupos: Verdaderos Positivos (TP), Falsos Positivos (FP), Falsos Negativos (FN), y Verdaderos Negativos (TN). Como se indicó anteriormente, hay muy pocos casos de FN, y un porcentaje sustancial de casos de FP. Sorprendentemente, el grupo de falsos positivos (positivo por GECL, estatus de SLNB negativo), tiene estimaciones de supervivencia muy similares a las del grupo de Verdaderos Negativos. Esto indica que la salida del clasificador de GECL es superior al estatus de SLNB como medio para proporcionar un pronóstico a los pacientes con Melanoma.

- 30 Se utiliza el método de Kaplan-Meier para estimar la probabilidad de supervivencia en varios intervalos de tiempo. La prueba de rangos logarítmicos es una prueba no paramétrica que se utiliza para comparar curvas de supervivencia entre dos o más grupos.

- 35 La proporción de riesgo (HR) se ha definido como la proporción de (riesgo del resultado en un grupo)/(riesgo del resultado en otro grupo), que ocurre en un intervalo de tiempo determinado. La proporción de riesgo de 1 significa falta de asociación, una proporción de riesgo mayor que 1 sugiere un mayor riesgo, y una proporción de riesgo inferior a 1 sugiere un riesgo menor. La proporción de riesgo se utiliza para representar la diferencia relativa entre solamente dos grupos.

Ejemplo 8: Refinamiento del clasificador de SLN

- 40 Se midieron cuatro paneles de genes en toda la cohorte de 855 para un total de 109 genes únicos. Sin embargo, no se pudo medir la expresión génica de algunas de las muestras porque no se pudieron recuperar en el momento del análisis o no había suficiente ARN. Por tanto, el descubrimiento se ha realizado en una cohorte de 754 pacientes en lugar de 770. La cohorte de descubrimiento no incluye pacientes con un volumen bajo de enfermedad metastásica.

Para cada clasificador, reportamos

- 45
- Desempeño del clasificador final entrenado para toda la cohorte
 - Desempeño promedio en validación cruzada (100 repeticiones de validación cruzada de doble bucle con 3 veces para el bucle externo y 10 veces para el bucle interno). El desempeño en las 3 veces se concatena para cubrir toda la cohorte.
- 50

Para cada clasificador, el desempeño se calcula en 4 puntos operativos diferentes:

- MaxbACC: Precisión equilibrada máxima
- SEeqSP: Sensibilidad igual a especificidad
- NPV97: Valor predictivo negativo igual al 97% en el entrenamiento • NPV98: Valor predictivo negativo igual al 98% en el entrenamiento

5

Modelo clínico-patológico (CL)

Los parámetros del modelo clasificador logístico basado en variables clínico-patológicas son los siguientes.

características	parámetros
(Intercepción)	-2.0547083
edad	-0.0112913
breslow_profundidad	0.6116335
invasión_angiolinfáticasí	0.1205238

10 La Tabla 27 representa el desempeño del clasificador final entrenado en toda la cohorte de 754 pacientes para cuatro puntos operativos diferentes: 1) max bACC: precisión equilibrada máxima, 2) SEeqSP, sensibilidad igual a especificidad, 3) NPV97, NPV establecido en 0.97 en el entrenamiento, 4) NPV98, NPV establecido en 0.98 en el entrenamiento. Si el coeficiente es positivo, entonces un valor más alto implica un mayor riesgo. Si el coeficiente es negativo, entonces el valor reducido implica una disminución del riesgo. Las variables con coeficientes más grandes (absolutos) tienen una contribución mayor.

15 La Tabla 28 representa el desempeño de los clasificadores entrenados en DLCV, promediado durante más de 100 repeticiones, para cuatro puntos operativos diferentes: 1) max bACC: precisión equilibrada máxima, 2) SEeqSP, sensibilidad igual a especificidad, 3) NPV97, NPV establecido en 0.97 en el entrenamiento, 4) NPV98, NPV establecido en 0.98 en el entrenamiento.

Modelo de expresión génica (GE)

Los parámetros del modelo clasificador logístico basado en la expresión génica son los siguientes.			
características	parámetros	características	parámetros
(Intercepción)	0.2163637	ADIPOQ	0.0011397
GDF15	-0.0043387	PRKCB	0.0070037
MLAN	-0.0075059	SERPINE2	-0.2954317
PLAT	-0.1181693	ADAM12	-0.0088591
IL8	-0.0413252	LGALS1	-0.0291427
ITGB3	-0.0725009	TGFBR1	-0.0003869
LOXL4	-0.0211895		

20 La Tabla 29 representa el desempeño del clasificador final entrenado en toda la cohorte de 754 pacientes para cuatro puntos operativos diferentes: 1) max bACC: precisión equilibrada máxima, 2) SEeqSP, sensibilidad igual a especificidad, 3) NPV97, NPV establecido en 0.97 en el entrenamiento, 4) NPV98, NPV establecido en 0.98 en el entrenamiento. Si el coeficiente es positivo, entonces un valor más alto implica un mayor riesgo. Si el coeficiente es negativo, entonces el valor reducido implica una disminución del riesgo. Las variables con coeficientes más grandes (absolutos) tienen una contribución mayor.

25

La Tabla 30 representa el desempeño de los clasificadores entrenados en DLCV, promediado durante más de 100 repeticiones, para cuatro puntos operativos diferentes: 1) max bACC: precisión equilibrada máxima, 2) SEeqSP, sensibilidad igual a especificidad, 3) NPV97, NPV establecido en 0.97 en el entrenamiento, 4) NPV98, NPV establecido en 0.98 en el entrenamiento.

30 Modelo combinado con clínico-patológica y expresión génica (GECL)

Los parámetros del modelo clasificador logístico basado en la expresión génica y variables clínico-patológicas son los siguientes.			
características	parámetros	características	parámetros
(Intercepción)	-0.1712118	IL8	-0.0299491
edad	-0.0119619	ITGB3	-0.0650425
breslow_profundidad	0.4152354	LOXL4	-0.0090683
GDF15	-0.0034702	ADIPOQ	0.0042436
MLAN	-0.0203507	SERPINE2	-0.2287815
PLAT	-0.1176641	TGFBR 1	-0.0095219

5 La Tabla 31 representa el desempeño del clasificador final entrenado en toda la cohorte de 754 pacientes para cuatro puntos operativos diferentes: 1) max bACC: precisión equilibrada máxima, 2) SEeqSP, sensibilidad igual a especificidad, 3) NPV97, NPV establecido en 0.97 en el entrenamiento, 4) NPV98, NPV establecido en 0.98 en el entrenamiento. Si el coeficiente es positivo, entonces un valor más alto implica un mayor riesgo. Si el coeficiente es negativo, entonces el valor reducido implica una disminución del riesgo. Las variables con coeficientes más grandes (absolutos) tienen una contribución mayor.

10 La Tabla 32 representa el desempeño de los clasificadores entrenados en DLCV, promediado durante más de 100 repeticiones, para cuatro puntos operativos diferentes: 1) max bACC: precisión equilibrada máxima, 2) SEeqSP, sensibilidad igual a especificidad, 3) NPV97, NPV establecido en 0.97 en el entrenamiento, 4) NPV98, NPV establecido en 0.98 en el entrenamiento.

15 La Figura 6 representa las curvas ROC para los clasificadores de regresión logística entrenados en DLCV en: 1) expresión génica, 2) variables clínico-patológicas, 3) expresión génica y variables clínico-patológicas combinadas, y la Figura 7 representa el Valor Predictivo Negativo (NPV) versus la Tasa de Reducción del Ganglio Linfático Centinela (SLNB RR) para los clasificadores de regresión logística entrenados en DLCV en: 1) expresión génica, 2) variables clínico-patológicas, 3) expresión génica y variables clínico-patológicas combinadas.

Comparaciones en diferentes puntos operativos (OP): CL versus CE versus GECL

20 La Tabla 33 representa el desempeño promedio de los clasificadores entrenados en DLCV en: 1) expresión génica ("GE"; es decir, firma génica GDF15, MLANA, PLAT, IL8, ITGB3, LOXL4, ADIPOQ, PRKCB, SERPINE2, ADAM12, LGALS1 y TGFBR1), 2) variables clínico-patológicas ("CL"; es decir, edad, profundidad de Breslow y presencia de invasión_angiolinfática), 3) expresión génica y variables clínico-patológicas combinadas ("GECL": es decir, edad, profundidad de Breslow, GDF15, MLANA, PLAT, IL8, ITGB3, LOXL4, ADIPOQ, SERPINE2 y TGFBR1). Para precisión equilibrada máxima del punto operativo (max bACC). La Tabla 34 representa el desempeño promedio de los clasificadores entrenados en DLCV en: 1) expresión génica ("GE"; es decir, firma génica GDF15, MLANA, PLAT, IL8, ITGB3, LOXL4, ADIPOQ, PRKCB, SERPINE2, ADAM12, LGALS1 y TGFBR1), 2) variables clínico-patológicas ("CL"; es decir, edad, profundidad de Breslow y presencia de invasión_angiolinfática), 3) expresión génica y variables clínico-patológicas combinadas ("GECL": es decir, edad, profundidad de Breslow, GDF15, MLANA, PLAT, IL8, ITGB3, LOXL4, ADIPOQ, SERPINE2 y TGFBR1). Para punto operativo con sensibilidad igual a especificidad (SEeqSP).

30 La Tabla 35 representa el desempeño promedio de los clasificadores entrenados en DLCV en: 1) expresión génica ("GE"; es decir, firma génica GDF15, MLANA, PLAT, IL8, ITGB3, LOXL4, ADIPOQ, PRKCB, SERPINE2, ADAM12, LGALS1 y TGFBR1), 2) variables clínico-patológicas ("CL"; es decir, edad, profundidad de Breslow y presencia de invasión_angiolinfática), 3) expresión génica y variables clínico-patológicas combinadas ("GECL": es decir, edad, profundidad de Breslow, GDF15, MLANA, PLAT, IL8, ITGB3, LOXL4, ADIPOQ, SERPINE2 y TGFBR1). Para puntos operativos con NPV establecido en 0.97 en el entrenamiento (NPV97).

35 La Tabla 36 representa el desempeño promedio de los clasificadores entrenados en DLCV en: 1) expresión génica ("GE"; es decir, firma génica GDF15, MLANA, PLAT, IL8, ITGB3, LOXL4, ADIPOQ, PRKCB, SERPINE2, ADAM12, LGALS1 y TGFBR1), 2) variables clínico-patológicas ("CL"; es decir, edad, profundidad de Breslow y presencia de invasión_angiolinfática), 3) expresión génica y variables clínico-patológicas combinadas ("GECL": es decir, edad, profundidad de Breslow, GDF15, MLANA, PLAT, IL8, ITGB3, LOXL4, ADIPOQ, SERPINE2 y TGFBR1). Para puntos operativos con NPV establecido en 0.98 en el entrenamiento (NPV98).

40 Desempeño estratificado por estadificación T

La Tabla 37 representa el desempeño promedio estratificado por estadio T de los clasificadores entrenados en DLCV en variables clínico-patológicas ("CL"; es decir, edad, profundidad de Breslow y presencia de invasión_angiolinfática). Para un punto operativo con NPV establecido en 0.97 en el entrenamiento (NPV97).

5 La Tabla 38 representa el desempeño promedio estratificado por estadio T de los clasificadores entrenados en DLCV en expresión génica ("GE"; es decir, GDF15, MLANA, PLAT, IL8, ITGB3, LOXL4, ADIPOQ, PRKCB, SERPINE2, ADAM12, LGALS1 y TGFBR1). Para un punto operativo con NPV establecido en 0.97 en el entrenamiento (NPV97).

10 La Tabla 39 representa el desempeño promedio estratificado por estadio T de los clasificadores entrenados en DLCV en expresión génica y variables clínico-patológicas combinadas ("GECL": es decir, edad, profundidad de Breslow, GDF15, MLANA, PLAT, IL8, ITGB3, LOXL4, ADIPOQ, SERPINE2 y TGFBR1). Para puntos operativos con NPV establecido en 0.97 en el entrenamiento (NPV97).

La Tabla 40 representa el desempeño promedio estratificado por estadio T de los clasificadores entrenados en DLCV en variables clínico-patológicas ("CL"; es decir, edad, profundidad de Breslow y presencia de invasión_angiolinfática). Para un punto operativo con NPV establecido en 0.98 en el entrenamiento (NPV98).

15 La Tabla 41 representa el desempeño promedio estratificado por estadio T de los clasificadores entrenados en DLCV en expresión génica ("GE"; es decir, GDF15, MLANA, PLAT, IL8, ITGB3, LOXL4, ADIPOQ, PRKCB, SERPINE2, ADAM12, LGALS1 y TGFBR1). Para un punto operativo con NPV establecido en 0.98 en el entrenamiento (NPV98).

20 La Tabla 42 representa el desempeño promedio estratificado por estadio T de los clasificadores entrenados en DLCV en expresión génica y variables clínico-patológicas combinadas ("GECL": es decir, edad, profundidad de Breslow, GDF15, MLANA, PLAT, IL8, ITGB3, LOXL4, ADIPOQ, SERPINE2 y TGFBR1). Para puntos operativos con NPV establecido en 0.98 en el entrenamiento (NPV98).

Desempeño estratificado por estadificación clínica

La Tabla 43 representa el desempeño promedio estratificado por estadio clínico de los clasificadores entrenados en DLCV en variables clínico-patológicas ("CL"; es decir, edad, profundidad de Breslow y presencia de invasión_angiolinfática). Para un punto operativo con NPV establecido en 0.97 en el entrenamiento (NPV97).

25 La Tabla 44 representa el desempeño promedio estratificado por estadio clínico de los clasificadores entrenados en DLCV en expresión génica ("GE"; es decir, GDF15, MLANA, PLAT, IL8, ITGB3, LOXL4, ADIPOQ, PRKCB, SERPINE2, ADAM12, LGALS1 y TGFBR1). Para un punto operativo con NPV establecido en 0.97 en el entrenamiento (NPV97).

30 La Tabla 45 representa el desempeño promedio estratificado por estadio clínico de los clasificadores entrenados en DLCV en expresión génica y variables clínico-patológicas combinadas ("GECL": es decir, edad, profundidad de Breslow, GDF15, MLANA, PLAT, IL8, ITGB3, LOXL4, ADIPOQ, SERPINE2 y TGFBR1). Para puntos operativos con NPV establecido en 0.97 en el entrenamiento (NPV97).

La Tabla 46 representa el desempeño promedio estratificado por estadio clínico de los clasificadores entrenados en DLCV en variables clínico-patológicas ("CL"; es decir, edad, profundidad de Breslow y presencia de invasión_angiolinfática). Para un punto operativo con NPV establecido en 0.98 en el entrenamiento (NPV98).

35 La Tabla 47 representa el desempeño promedio estratificado por estadio clínico de los clasificadores entrenados en DLCV en expresión génica ("GE"; es decir, GDF15, MLANA, PLAT, IL8, ITGB3, LOXL4, ADIPOQ, PRKCB, SERPINE2, ADAM12, LGALS1 y TGFBR1). Para un punto operativo con NPV establecido en 0.98 en el entrenamiento (NPV98).

40 La Tabla 48 representa el desempeño promedio estratificado por estadio clínico de los clasificadores entrenados en DLCV en expresión génica y variables clínico-patológicas combinadas ("GECL": es decir, edad, profundidad de Breslow, GDF15, MLANA, PLAT, IL8, ITGB3, LOXL4, ADIPOQ, SERPINE2 y TGFBR1). Para puntos operativos con NPV establecido en 0.98 en el entrenamiento (NPV98).

Subconjuntos de genes

45 La Figura 8 representa los diagramas de caja del área bajo la curva (AUC) de la curva ROC para clasificadores de regresión logística con subconjuntos de 2, 3, 4, 5, 6, 7, 8 genes seleccionados de GDF15, MLANA, PLAT, IL8, ITGB3, LOXL4, ADIPOQ, SERPINE2, y TGFBR1 y entrenados en toda la cohorte.

La Tabla 49 representa el Número de subconjuntos de una dimensión específica que se pueden seleccionar del número total de genes en cada firma, y el desempeño en términos de área mínima y máxima bajo la curva ROC.

Ejemplo 9: Perfilador del ganglio linfático no centinela (N-SLN).

50 La disección completa de los ganglios linfáticos (CLND), con la extirpación del ganglio linfático no-centinela (N-SLN), ha sido el estándar para pacientes clínicamente con melanoma de ganglios negativos con ganglios linfáticos centinela (SLN) positivos. La biopsia del SLN seguida de CLND inmediata mejora el control regional de la enfermedad, y los ensayos clínicos aleatorizados han demostrado que la cirugía temprana para la enfermedad con un volumen menor

de SLN-positivo produce menos secuelas a largo plazo (por ejemplo, linfedema) que la cirugía en caso de reincidencia ganglionar. Además, la metástasis de SLN y N-SLN es un factor de pronóstico adverso utilizado para seleccionar pacientes para terapia adyuvante. Sin embargo, para los pacientes matriculados en MSLT-II, la CLND no proporcionó ningún beneficio de supervivencia, la cual tiene una tasa de complicaciones más alta que solo la cirugía del SLN. Se necesitan nuevos métodos para identificar a los pacientes que probablemente se beneficiarán de la CLND, es decir, aquellos con riesgo de metástasis regional en N-SLN, para mejorar la selección de pacientes para CLND. Aquí diseñamos tres clasificadores basados en la expresión génica (KRT14, SPP1, FN1 y LOXL3), variables clínico-patológicas (edad, número de SLN positivos, dimensión más grande del SLN, profundidad máx de Breslow y tasa mitótica máx) y en ambos, para predecir el estatus de N-SLN, es decir, presencia o ausencia de metástasis. Estos clasificadores se pueden utilizar para seleccionar qué pacientes deben someterse al procedimiento CLND.

Los métodos utilizados fueron en gran medida los mismos que los indicados en el ejemplo 1. Se evaluaron los mismos 29 genes. Los parámetros del modelo clasificador de regresión logística basado en la expresión génica son los siguientes:

Intercepción	-0.94837538
KRT14_3	0,06687021
SPP1_3	-0.02696531
FN1_3	-0,20982959
LOXL_3.3	-0,01722991

15 Desempeño del modelo

El número de pacientes en toda la cohorte fue 140.

La Tabla 11 representa el desempeño de los clasificadores entrenados en DLCV, promediado durante más de 100 repeticiones, para cuatro puntos operativos diferentes: 1) max bACC: precisión equilibrada máxima, 2) SEeqSP, sensibilidad igual a especificidad, 3) NPV97, NPV establecido en 0.97 en el entrenamiento, 4) LRNn025, proporción de probabilidad logarítmica para el resultado negativo de la prueba establecido en 0.25 en el entrenamiento.

La Tabla 12 representa el desempeño de los clasificadores entrenados en DLCV, promediado durante 100 repeticiones, para cuatro puntos operativos diferentes: 1) max bACC: precisión equilibrada máxima, 2) SEeqSP, sensibilidad igual a especificidad, 3) NPV97, NPV establecido en 0.97 en el entrenamiento, 4) LRNn025, proporción de probabilidad logarítmica para el resultado negativo de la prueba establecido en 0.25 en el entrenamiento.

25 Clasificador basado en Variables Clínicas (CL)

Los parámetros del modelo clasificador logístico basado en variables clínicas fueron los siguientes:

(Intercepción)	-4,60115005
Edad	0,03787092
Profundidad máx de Breslow	0,04057410
slnb_num_positivo.cat2+	1,14121008
Ganglio_linfático_más grande_	0,90937183
dim.cat2.5+	
Max_mitótica.cat>6	0,35332099

La Tabla 13 representa el desempeño del clasificador final entrenado en todo el clasificador de cohorte de 140 pacientes para cuatro puntos operativos diferentes: 1) max bACC: precisión equilibrada máxima, 2) SEeqSP, sensibilidad igual a especificidad, 3) NPV97, NPV establecido en 0.97 en el entrenamiento, 4) LRNn025, proporción de probabilidad logarítmica para el resultado negativo de la prueba establecido en 0.25 en el entrenamiento.

La Tabla 14 representa el desempeño de los clasificadores entrenados en DLCV, promediado durante más de 100 repeticiones, para cuatro puntos operativos diferentes: 1) max bACC: precisión equilibrada máxima, 2) SEeqSP,

sensibilidad igual a especificidad, 3) NPV97, NPV establecido en 0.97 en el entrenamiento, 4) LRNn025, proporción de probabilidad logarítmica para el resultado negativo de la prueba establecido en 0.25 en el entrenamiento.

Clasificador basado en Expresión Génica y Variables Clínicas (GECL)

Parámetros de regresión logística

Intercepción	-2.34454707
Edad	0.02009938
slnb_num_positivo.cat2+	0.70415832
Ganglio_linfático_más grande_dim .cat2.5+	0.67873684
KRT14_3	0.06115660
SPP1_3	-0.01572158
FN1_3	-0.22148054
LOXL3_3	-0.04907296

5

Parámetros del modelo clasificador logístico basado en expresión génica.

La Tabla 15 representa el desempeño del clasificador final entrenado en todo el clasificador de cohorte de 140 pacientes para cuatro puntos operativos diferentes: 1) max bACC: precisión equilibrada máxima, 2) SEeqSP, sensibilidad igual a especificidad, 3) NPV97, NPV establecido en 0.97 en el entrenamiento, 4) LRNn025, proporción de probabilidad logarítmica para el resultado negativo de la prueba establecido en 0.25 en el entrenamiento.

10

La Tabla 16 representa el desempeño de los clasificadores entrenados en DLCV, promediado durante 100 repeticiones, para cuatro puntos operativos diferentes: 1) max bACC: precisión equilibrada máxima, 2) SEeqSP, sensibilidad igual a especificidad, 3) NPV97, NPV establecido en 0.97 en el entrenamiento, 4) LRNn025, proporción de probabilidad logarítmica para el resultado negativo de la prueba establecido en 0.25 en el entrenamiento.

15

La Figura 4 representa las curvas ROC promedio para los clasificadores de regresión logística entrenados en DLCV en: 1) expresión génica, 2) variables clínico-patológicas, 3) expresión génica y variables clínico-patológicas combinadas. El eje-x representa la tasa de descubrimiento de falsos positivos (es decir, 1-especificidad), el eje-y, la tasa de descubrimiento real (es decir, sensibilidad).

20

Figura 5: Diagramas de caja del área bajo las curvas ROC para los diferentes subconjuntos de genes y para el conjunto completo de 4 genes.

La Tabla 17 representa el desempeño promedio de los clasificadores entrenados en DLCV en: 1) expresión génica, 2) variables clínico-patológicas, 3) expresión génica y variables clínico-patológicas combinadas. Se han considerado tres puntos operativos diferentes: 1) max bACC: precisión equilibrada máxima, 2) SEeqSP, sensibilidad igual a especificidad, 3) NPV97, NPV establecido en 0.97 en el entrenamiento.

25

La firma del perfilador, perfiladorN-SLN comprende 4 genes: KRT14, SPP1, FN1 y LOXL3. También hemos analizado a todos los subconjuntos posibles de 2, 3 y 4 genes.

Marcador	Subconjuntos de 2	Subconjuntos de 3	Genes en firma
NSLNB	6	4	4

Número de subconjuntos de una dimensión específica que se pueden seleccionar del número total de genes en cada firma.

30

Hemos evaluado el desempeño en términos del área bajo la curva ROC (véase tabla y figura)

Medida del desempeño	Subconjunto de 2	Subconjunto de 3	Genes en la firma
Intervalo del AUC	0.72-0.76	0.77-0.80	0.81

El intervalo del área bajo la curva (AUC) para los diferentes subconjuntos de genes, para el conjunto completo de 4 genes.

La Figura 5 representa diagramas de caja del área bajo las curvas ROC para los diferentes subconjuntos de genes y para el conjunto completo de 4 genes.

5 Describimos el descubrimiento, diseño y desarrollo del perfiladorN-SLN. Hemos demostrado que el perfiladorN-SLN se puede utilizar para seleccionar qué pacientes deben someterse al procedimiento CLND.

El desempeño del clasificador basado en expresión génica es de interés, dado que por el momento (i) no hay métodos disponibles para seleccionar a los pacientes que se beneficiarían del procedimiento CLND, y (ii) las variables clínicopatológicas utilizadas por el clasificador podrían no siempre estar disponibles en la clínica.

10 Ejemplo 10

Preselección basada en la profundidad de Breslow (BD<=2 versus BD>2)

15 La profundidad de Breslow (BD) es una variable clínico-patológica importante que caracteriza el melanoma cutáneo primario, y se sabe que los melanomas delgados (BD <= 2 mm) pueden tener características moleculares y fisiológicas diferentes en comparación con los melanomas gruesos (BD>2 mm). Por lo tanto, tendría sentido elegir dos puntos operativos diferentes para melanomas delgados y gruesos. Para melanomas delgados con una profundidad de Breslow <= 2 mm (cohorte de 561 muestras), seguimos utilizando el clasificador como se describe en la sección anterior (7.8), con el punto operativo elegido de modo que el NPV sea 0.97 en el entrenamiento:

20 La Tabla 18a representa el desempeño promedio de los clasificadores entrenados en DLCV en: 1) expresión génica, 2) variables clínico-patológicas, 3) expresión génica y variables clínico-patológicas combinadas. El punto operativo elegido es NPV97, es decir, NPV establecido en 0.97 en el entrenamiento.

Para melanomas gruesos con BD>2 mm (cohorte de 209 muestras), utilizamos el mismo clasificador, pero con un punto operativo diferente, es decir, un punto elegido para maximizar la exactitud equilibrada:

25 La Tabla 18b representa un desempeño promedio de los clasificadores entrenados en DLCV en: 1) expresión génica, 2) variables clínico-patológicas, 3) expresión génica y variables clínico-patológicas combinadas. El punto operativo elegido se ha seleccionado de forma que maximice la exactitud equilibrada.

El desempeño para una profundidad de Breslow >2, aunque inferior al desempeño para una profundidad de Breslow <=2 mm, sigue siendo aceptable dado que el clasificador alcanza un NPV del 90% en una subpoblación con una probabilidad previa de positividad de SLNB del 35%: al tomar la prueba y tener un resultado negativo, la probabilidad de tener una SLNB Positivo baja del 35% al 10%.

30 Método para determinar si un sujeto se clasifica como ganglio linfático positivo o ganglio negativo

35 Datos ficticios (véase la Tabla) se utilizan como un ejemplo para el método de clasificación, utilizando 2 genes por simplicidad, para predecir si el clasificador va a marcar una muestra como ganglio linfático positivo o ganglio linfático negativo (el método/modelo es el mismo para tanto el perfilador SLN como el perfilador N-SLN, solamente los parámetros y las identidades de los genes y las variables clínicas son diferentes). La tabla describe los parámetros del modelo $\beta_0, \beta_1, \beta_2$ para dos genes (x, y) modelo de juguete, los datos de expresión génica ficticios ΔC_i , la proporción

40 impar logarítmica estimada $\log \frac{p}{1-p}$, la probabilidad estimada p , y los marcadores de clase de salida estimados basadas en el corte $\theta = 0.19$. La proporción impar logarítmica y la probabilidad se calculan con base en la ecuación 1, y la probabilidad se calcula utilizando la ecuación 2. El marcador de salida se asigna comparando la probabilidad estimada con el corte θ . si la probabilidad estimada es mayor o igual a θ , entonces la muestra se clasifica como ganglio positivo; si la probabilidad estimada es menor que θ entonces la muestra se clasifica como ganglio negativo.

ID del gen	Parámetros	Expresión génica (ΔC_i)	$\log \frac{p}{1-p}$	p	$p \geq \theta$
	$\beta_0 = 0.78$				
x	$\beta_1 = -0.19$	7.2	-1.272	0.21	Ganglio +
y	$\beta_2 = -0.12$	5.7			

Tabla. Parámetros del modelo $\beta_0, \beta_1, \beta_2$, datos de expresión génica ΔC_i para genes x, y, proporción-impar logarítmica estimada $\log \frac{p}{1-p}$, probabilidad estimada p , y la clase de salida estimada con base en un corte $\theta = 0.19$.

$$\log \frac{p}{1-p} = \beta_0 + \beta_1 x + \beta_2 y \text{ (eq. 1)}$$

$$p = \frac{e^{\beta_0 + \beta_1 x + \beta_2 y}}{1 + e^{\beta_0 + \beta_1 x + \beta_2 y}} \text{ (eq. 2)}$$

Ejemplo 11

5 Equivalente al análisis realizado en el Ejemplo 8, se encontró una realización preferida de la invención que emplea 8 genes, y 2 variables clínico-patológicas (Edad y Profundidad de Breslow), que comprende el siguiente conjunto de genes: GDF15, MLANA, PLAT, IL8, ITGB3, LOXL4, SERPINE2, y TGFBR1. Los parámetros del modelo de regresión logística se proporcionan a continuación:

características	parámetros
(Intercepción)	-0.1736287
edad	-0.0118062
breslow_profundidad	0.4126846
GDF15	-0.0027222
MLANA	-0.0185367
PLAT	-0.1170421
IL8	-0.0299650
ITGB3	-0.0636119
LOXL4	-0.0054420
SERPINE2	-0.2342833
TGFBR1	-0.0067781

10 Se evaluaron diferentes puntos operativos, revelando que múltiples puntos operativos proporcionan clasificadores clínicamente relevantes, con un NPV alto, y una reducción sustancial del número de procedimientos de SLNB (SLNB.RR = Tasa de reducción de SLNB). Se prefiere específicamente el punto operativo NPV975 de 0.116; véase la tabla a continuación.

Tabla con desempeños del modelo GECL con 8 genes y 2 variables clínico-patológicas

Entrada	P%	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	SLNB.RR	ER	AUC	Corte	LRn	LRp
maxbACC	0.17	128	626	102	156	471	26	0.797	0.752	0.948	0.397	0.76	0.775	0.656	0.0345	0.84	0.166	0.27	3.22
SEeqSP	0.17	128	626	99	146	466	29	0.773	0.776	0.944	0.414	0.776	0.775	0.659	0.0366	0.84	0.192	0.292	3.46
NPV97	0.17	128	626	115	271	355	19	0.898	0.567	0.965	0.298	0.623	0.733	0.488	0.0172	0.84	0.193	0.175	2.08
NPV975	0.17	128	626	119	319	307	8	0.93	0.49	0.972	0.272	0.585	0.71	0.418	0.0119	0.84	0.116	0.143	1.82
NPV98	0.17	128	626	122	364	242	6	0.953	0.387	0.976	0.241	0.483	0.57	0.329	0.00736	0.84	0.0986	0.121	1.55

15 Comparar el desempeño de este modelo GECL con 8 genes y 2 variables clínicopatológicas con los de los Ejemplos 1 y 8, indica que supera (NPV y SLNB.RR) a los modelos de solo expresión (GE), o solo clínicopatológico (CL), y funciona muy similar al modelo GECL del Ejemplo 1.

20 Se realizó un análisis de Kaplan-Meier para evaluar el modelo GECL interceptado con SLNB en relativo a RFS, DRFS, y MSS. Las tasas de supervivencia a 5-años / 60 meses se reportan en la siguiente Tabla. Tanto el GECL como SLNB individualmente tienen una gran separación entre los grupos Pos/Neg. Más importante, aquellos pacientes identificados como Pos por GECL que son SLNB Neg (y omitidos como tales), tienen una supervivencia muy pobre. Esto demuestra la relevancia clínica del modelo GECL como marcador de pronóstico.

	MSS	DRFS	RFS
	5 años	5 años	5 años
GECL = Neg y SLNB = Neg	0.96	0.95	0.88
GECL = Neg y SLNB = Pos	0.83	0.50	0.34
GECL = Pos y SLNB = Neg	0.90	0.79	0.70
GECL = Pos y SLNB = Pos	0.84	0.72	0.50

Tablas:

Tabla 1

OP	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	AUC	Corte	LRn	LRp
max bACC	129	841	83	139	802	49	0.84	0.78	0.91	0.37	0.78	0.71	0.77	0.19	0.47	2.9
SEeqSP	129	841	82	178	483	37	0.71	0.72	0.92	0.34	0.72	0.72	0.77	0.18	0.4	2.6
NPV97	129	841	126	883	23	3	0.98	0.023	0.88	0.17	0.18	0.8	0.77	0.081	0.73	1
LRn025	129	841	126	824	17	3	0.98	0.027	0.85	0.17	0.18	0.8	0.77	0.08	0.89	1

5 Tabla 2

OP	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	AUC	Corte	LRn	LRp
max bACC	129	841	89	156	488	49	0.82	0.76	0.91	0.34	0.74	0.88	0.73	0.19	0.5	2.8
SEeqSP	129	841	87	187	464	42	0.87	0.71	0.92	0.32	0.7	0.89	0.75	0.18	0.46	2.3
NPV97	129	841	129	874	87	5	0.98	0.1	0.93	0.18	0.25	0.83	0.75	0.081	0.37	1.1
LRn025	129	841	124	829	82	4	0.97	0.087	0.94	0.18	0.24	0.83	0.75	0.08	0.32	1.1

Tabla 3

OP	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	AUC	Corte	LRn	LRp
max bACC	129	841	94	164	487	35	0.73	0.71	0.93	0.34	0.72	0.72	0.78	0.18	0.88	0.6
SEeqSP	129	841	92	175	465	37	0.71	0.73	0.93	0.34	0.72	0.72	0.79	0.17	0.4	0.6
NPV97	129	841	115	881	310	14	0.99	0.48	0.99	0.28	0.55	0.88	0.78	0.12	0.22	1.7
LRn025	129	841	115	834	307	14	0.98	0.48	0.96	0.28	0.55	0.88	0.79	0.12	0.23	1.7

Tabla 4

OP	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	AUC	Corte	LRn	LRp
max bACC	129	841	92	188	449	37	0.73	0.7	0.92	0.32	0.7	0.71	0.77	0.18	0.41	2.4
SEeqSP	129	841	92	183	459	38	0.7	0.71	0.92	0.29	0.71	0.71	0.77	0.17	0.42	2.4
NPV97	129	841	129	406	238	9	0.99	0.37	0.96	0.23	0.46	0.84	0.77	0.12	0.19	1.8
LRn025	129	841	129	409	232	8	0.94	0.38	0.97	0.23	0.46	0.85	0.77	0.12	0.17	1.5

10

Tabla 5

ES 2 966 015 T3

OP	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	AUC	Corte	LRn	LRp
maxbACC	129	641	105	163	478	24	0.81	0.75	0.95	0.39	0.79	0.79	0.85	0.17	0.25	3.2
SEeqSP	129	641	86	149	482	34	0.74	0.77	0.94	0.39	0.76	0.76	0.83	0.18	0.34	3.2
NPV97	129	641	119	332	306	10	0.92	0.48	0.97	0.28	0.53	0.7	0.83	0.11	0.16	1.8
LRn0.25	129	641	119	341	300	10	0.92	0.47	0.97	0.28	0.54	0.7	0.83	0.11	0.17	1.7

Tabla 6

OP	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	AUC	Corte	LRn	LRp
maxbACC	129	641	97	177	464	32	0.75	0.72	0.94	0.35	0.79	0.74	0.8	0.17	0.34	2.7
SEeqSP	129	641	92	160	481	37	0.71	0.75	0.93	0.37	0.74	0.73	0.8	0.18	0.38	2.9
NPV97	129	641	117	355	285	12	0.91	0.45	0.98	0.25	0.52	0.68	0.8	0.11	0.21	1.8
LRn0.25	129	641	117	362	279	12	0.91	0.44	0.98	0.25	0.51	0.67	0.8	0.11	0.21	1.8

Tabla 7

maxbACC

Entrada	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	AUC	Corte	LRn	LRp
GE	129	641	80	163	482	49	0.62	0.76	0.91	0.34	0.74	0.69	0.75	0.19	0.3	2.5
CL	129	641	92	182	469	37	0.71	0.7	0.92	0.32	0.7	0.71	0.77	0.18	0.41	2.4
GECL	129	641	97	177	464	32	0.75	0.72	0.94	0.35	0.73	0.75	0.8	0.17	0.34	2.7

SEeqSP

Entrada	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	AUC	Corte	LRn	LRp
GE	129	641	87	187	454	42	0.67	0.71	0.92	0.32	0.7	0.69	0.75	0.18	0.35	2.3
CL	129	641	90	183	458	39	0.7	0.71	0.92	0.33	0.71	0.71	0.77	0.17	0.42	2.4
GECL	129	641	92	180	461	37	0.71	0.75	0.93	0.37	0.74	0.73	0.8	0.18	0.38	2.3

NPV97

Entrada	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	AUC	Corte	LRn	LRp
GE	129	641	124	574	67	5	0.96	0.11	0.99	0.18	0.26	0.53	0.75	0.081	0.37	1.3
CL	129	641	129	405	298	9	0.98	0.37	0.96	0.23	0.48	0.65	0.77	0.12	0.19	1.5
GECL	129	641	137	355	285	12	0.91	0.45	0.96	0.25	0.52	0.68	0.8	0.11	0.21	1.8

5

Tabla 8

P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	AUC	Corte	LRn	LRp
129	641	70	142	492	59	0.64	0.78	0.89	0.33	0.74	0.68	NA	NA	0.28	2.4

Tabla 9

OP	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	AUC	Corte	LRn	LRp
maxbACC	129	641	97	166	478	35	0.75	0.74	0.94	0.37	0.75	0.75	0.81	0.17	0.33	2.8
SEeqSP	129	641	94	159	482	38	0.73	0.75	0.93	0.37	0.75	0.74	0.81	0.17	0.38	2.8
NPV97	129	641	119	361	280	10	0.92	0.44	0.97	0.25	0.52	0.68	0.81	0.12	0.19	1.6
LRn0.25	129	641	119	362	279	10	0.92	0.44	0.97	0.25	0.52	0.68	0.81	0.12	0.18	1.6

10 Tabla 10

ES 2 966 015 T3

OP	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	AUC	Corte	LRn	LRp
max bACC	129	841	94	176	465	35	0.73	0.75	0.93	0.95	0.73	0.75	0.78	0.17	0.37	2.7
SEeqSP	129	841	82	167	474	37	0.73	0.74	0.93	0.98	0.74	0.73	0.78	0.17	0.38	2.7
NPV97	129	841	116	354	287	11	0.91	0.45	0.98	0.25	0.33	0.68	0.78	0.12	0.18	1.7
LRn025	129	841	118	360	281	11	0.91	0.44	0.98	0.25	0.32	0.68	0.78	0.12	0.19	1.8

Tabla 11

OP	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	AUC	Corte	LRn	LRp
max bACC	25	115	15	28	89	13	0.6	0.77	0.9	0.37	0.74	0.89	0.78	0.2	0.32	2.7
SEeqSP	25	115	15	25	90	13	0.6	0.78	0.9	0.38	0.75	0.89	0.78	0.21	0.51	2.8
NPV97	25	115	18	44	71	8	0.76	0.88	0.92	0.3	0.64	0.88	0.78	0.18	0.89	2
LRn025	25	115	18	44	71	8	0.76	0.82	0.92	0.3	0.64	0.88	0.78	0.18	0.89	2

Tabla 12

OP	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	AUC	Corte	LRn	LRp
max bACC	25	115	21	34	81	14	0.64	0.7	0.85	0.24	0.66	0.87	0.68	0.2	0.6	1.5
SEeqSP	25	115	20	29	85	15	0.4	0.75	0.85	0.29	0.68	0.87	0.68	0.21	0.8	1.6
NPV97	25	115	16	52	63	9	0.84	0.85	0.88	0.24	0.58	0.89	0.68	0.18	0.86	1.4
LRn025	25	115	16	51	64	9	0.84	0.86	0.88	0.24	0.57	0.8	0.68	0.18	0.86	1.4

5

Tabla 13

OP	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	AUC	Corte	LRn	LRp
max bACC	25	115	18	25	90	7	0.72	0.78	0.93	0.42	0.77	0.75	0.85	0.19	0.38	3.3
SEeqSP	25	115	16	24	91	7	0.72	0.79	0.93	0.43	0.78	0.78	0.85	0.2	0.39	3.4
NPV97	25	115	22	48	68	3	0.88	0.6	0.98	0.38	0.65	0.74	0.85	0.13	0.2	2.2
LRn025	25	115	22	45	70	3	0.88	0.61	0.96	0.39	0.66	0.74	0.85	0.14	0.2	2.2

Tabla 14

OP	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	AUC	Corte	LRn	LRp
max bACC	25	115	16	31	84	9	0.64	0.73	0.9	0.34	0.71	0.89	0.79	0.19	0.48	2.4
SEeqSP	25	115	15	28	89	10	0.6	0.77	0.9	0.37	0.74	0.89	0.79	0.2	0.52	2.7
NPV97	25	115	19	48	67	6	0.78	0.58	0.92	0.28	0.61	0.67	0.78	0.14	0.41	1.8
LRn025	25	115	19	47	68	6	0.79	0.59	0.93	0.29	0.62	0.68	0.79	0.14	0.38	1.8

10

Tabla 15

OP	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	AUC	Corte	LRn	LRp
max bACC	25	115	15	14	101	10	0.6	0.88	0.91	0.52	0.83	0.74	0.88	0.23	0.48	4.9
SEeqSP	25	115	16	30	95	7	0.72	0.83	0.93	0.47	0.81	0.77	0.88	0.22	0.54	4.1
NPV97	25	115	20	22	93	5	0.8	0.81	0.95	0.48	0.81	0.6	0.88	0.21	0.25	4.2
LRn025	25	115	19	21	94	6	0.75	0.82	0.94	0.47	0.81	0.75	0.88	0.21	0.29	4.2

Tabla 16

OP	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	AUC	Corte	LRn	LRp
max bACC	25	115	13	23	92	12	0.52	0.8	0.88	0.36	0.75	0.66	0.79	0.24	0.8	2.6
SEeqSP	25	115	14	25	90	11	0.56	0.78	0.89	0.36	0.74	0.67	0.79	0.22	0.56	2.6
NPV97	25	115	15	31	84	10	0.6	0.73	0.89	0.33	0.71	0.57	0.79	0.21	0.55	2.2
LRn025	25	115	15	30	85	10	0.6	0.74	0.89	0.33	0.71	0.67	0.79	0.21	0.54	2.3

Tabla 17

maxbACC

Entrada	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	AUC	Corte	LRn	LRp
GE	25	115	13	24	87	14	0.49	0.7	0.85	0.34	0.66	0.57	0.68	0.2	0.8	1.5
CL	25	115	16	31	80	9	0.64	0.73	0.8	0.34	0.71	0.69	0.78	0.19	0.49	2.4
GECL	25	115	12	23	92	12	0.52	0.8	0.88	0.36	0.75	0.66	0.79	0.24	0.8	2.6

SEeqSP

Entrada	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	AUC	Corte	LRn	LRp
GE	25	115	14	25	89	15	0.4	0.75	0.85	0.35	0.69	0.57	0.68	0.21	0.8	1.6
CL	25	115	16	29	86	10	0.6	0.77	0.8	0.37	0.74	0.69	0.78	0.2	0.57	2.7
GECL	25	115	14	25	90	11	0.46	0.75	0.89	0.35	0.74	0.67	0.79	0.22	0.56	2.6

NPV97

Entrada	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	AUC	Corte	LRn	LRp
GE	25	115	15	32	83	8	0.64	0.65	0.88	0.24	0.55	0.59	0.66	0.18	0.85	1.7
CL	25	115	18	40	67	6	0.76	0.68	0.88	0.28	0.61	0.67	0.75	0.14	0.71	1.8
GECL	25	115	15	31	84	10	0.6	0.73	0.89	0.33	0.71	0.67	0.79	0.21	0.55	2.2

Tabla 18a

5 Profundidad de Breslow <=2 mm

Entrada	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	AUC	Corte	LRn	LRp
GE	55	505	52	446	59	4	0.93	0.32	0.94	0.1	0.2	0.52	0.75	0.061	0.93	1.1
CL	55	505	47	270	236	8	0.84	0.47	0.68	0.15	0.8	0.63	0.78	0.12	0.35	1.6
GECL	55	505	44	225	280	12	0.79	0.55	0.95	0.16	0.55	0.67	0.81	0.11	0.39	1.9

Tabla 18b

Profundidad de Breslow >2 mm

Entrada	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	AUC	Corte	LRn	LRp
GE	73	138	55	86	50	18	0.75	0.69	0.88	0.5	0.55	0.67	0.76	0.19	0.42	1.8
CL	73	138	72	131	6	1	0.99	0.037	0.69	0.35	0.37	0.61	0.76	0.16	0.37	1
GECL	73	138	70	108	26	3	0.96	0.21	0.9	0.38	0.47	0.58	0.81	0.17	0.2	1.2

10 Tabla 19a: Estimaciones de supervivencia en diferentes puntos de tiempo para grupos basada en el estado del clasificador de estatus GECL o estatus SLNB. Las curvas de supervivencia se compararon utilizando un modelo de riesgos proporcionales de Cox, véase la Tabla 25.

Supervivencia específica del melanoma

ES 2 966 015 T3

Tiempo (Meses)	0	40	80	120	160
Probabilidad de supervivencia: GECL Negativo	1	0.977	0.966	0.966	0.966
Probabilidad de supervivencia: GECL Positivo	1	0.933	0.802	0.717	0.645
Probabilidad de supervivencia: SLNB Negativo	1	0.962	0.906	0.859	0.859
Probabilidad de supervivencia: SLNB Positivo	1	0.903	0.704	0.704	N/A

Tabla 19b:

Supervivencia específica del melanoma					
Tiempo (Meses)	0	40	80	120	160
Probabilidad de supervivencia: GECL Negativo	1	0.969	0.958	0.958	0.958
Probabilidad de supervivencia: GECL Positivo	1	0.934	0.786	0.704	N/A
Probabilidad de supervivencia: SLNB Negativo	1	0.962	0.905	0.866	0.866
Probabilidad de supervivencia: SLNB Positivo	1	0.901	0.695	0.695	N/A

Tabla 19c:

Supervivencia específica del melanoma					
Tiempo (Meses)	0	40	80	120	160
Probabilidad de supervivencia: GECL Negativo	1	0.970	0.962	0.962	0.962
Probabilidad de supervivencia: GECL Positivo	1	0.941	0.816	0.754	0.678
Probabilidad de supervivencia: SLNB Negativo	1	0.962	0.905	0.866	0.866
Probabilidad de supervivencia: SLNB Positivo	1	0.901	0.695	0.695	N/A

5

Tabla 20a: Estimaciones de supervivencia en diferentes puntos de tiempo para grupos basada en el estatus del clasificador GECL o estatus SLNB. Las curvas de supervivencia se compararon utilizando un modelo de riesgos proporcionales de Cox, véase la Tabla 25.

Supervivencia libre de reincidencia a distancia					
Tiempo (Meses)	0	40	80	120	160
Probabilidad de supervivencia: GECL Negativo	1	0.970	0.905	0.862	0.862
Probabilidad de supervivencia: GECL Positivo	1	0.830	0.611	0.503	0.503
Probabilidad de supervivencia: SLNB Negativo	1	0.913	0.796	0.711	0.711
Probabilidad de supervivencia: SLNB Positivo	1	0.769	0.458	0.407	N/A

10

Tabla 20b

Supervivencia libre de reincidencia a distancia					
Tiempo (Meses)	0	40	80	120	160
Probabilidad de supervivencia: GECL Negativo	1	0.955	0.894	0.854	0.854

ES 2 966 015 T3

Probabilidad de supervivencia: GECL Positivo	1	0.819	0.583	0.469	N/A
Probabilidad de supervivencia: SLNB Negativo	1	0.912	0.793	0.716	0.716
Probabilidad de supervivencia: SLNB Positivo	1	0.751	0.464	0.385	N/A

Tabla 20c:

Supervivencia libre de reincidencia a distancia					
Tiempo (Meses)	0	40	80	120	160
Probabilidad de supervivencia: GECL Negativo	1	0.955	0.896	0.842	0.842
Probabilidad de supervivencia: GECL Positivo	1	0.846	0.639	0.544	0.544
Probabilidad de supervivencia: SLNB Negativo	1	0.912	0.793	0.716	0.716
Probabilidad de supervivencia: SLNB Positivo	1	0.751	0.464	0.385	N/A

5 Tabla 21a: Estimaciones de supervivencia en diferentes puntos de tiempo para grupos basada en el estatus del clasificador GECL o estatus SLNB. Las curvas de supervivencia se compararon utilizando un modelo de riesgos proporcionales de Cox, véase la Tabla 25.

Supervivencia libre de reincidencia					
Tiempo (Meses)	0	40	80	120	160
Probabilidad de supervivencia: GECL Negativo	1	0.933	0.893	0.851	0.851
Probabilidad de supervivencia: GECL Positivo	1	0.690	0.601	0.531	0.531
Probabilidad de supervivencia: SLNB Negativo	1	0.834	0.783	0.712	0.712
Probabilidad de supervivencia: SLNB Positivo	1	0.590	0.442	0.442	N/A

Tabla 21b:

Supervivencia libre de reincidencia					
Tiempo (Meses)	0	40	80	120	160
Probabilidad de supervivencia: GECL Negativo	1	0.911	0.870	0.832	0.832
Probabilidad de supervivencia: GECL Positivo	1	0.674	0.584	0.518	N/A
Probabilidad de supervivencia: SLNB Negativo	1	0.832	0.778	0.716	0.716
Probabilidad de supervivencia: SLNB Positivo	1	0.582	0.459	0.459	N/A

10

Tabla 21c:

Supervivencia libre de reincidencia					
Tiempo (Meses)	0	40	80	120	160
Probabilidad de supervivencia: GECL Negativo	1	0.924	0.870	0.818	0.818
Probabilidad de supervivencia: GECL Positivo	1	0.717	0.643	0.591	0.591
Probabilidad de supervivencia: SLNB Negativo	1	0.832	0.778	0.716	0.716

Probabilidad de supervivencia: SLNB Positivo	1	0.582	0.459	0.459	N/A
--	---	-------	-------	-------	-----

Tabla 22a: Estimaciones de supervivencia en diferentes puntos de tiempo para grupos basada en el estatus del clasificador SLNB y GECL. Las curvas de supervivencia se compararon mediante una prueba de rangos logarítmicos, $p < 0.0001$.

Supervivencia específica del melanoma					
Tiempo (Meses)	0	40	80	120	160
Verdadero negativo SLNB (negativo) y GECL (negativo)	1	0.981	0.969	0.969	0.969
Falso negativo SLNB (positivo) y GECL (negativo)	1	0.857	0.857	N/A	N/A
Falso positivo 1 SLNB (negativo) y GECL (positivo)	1	0.942	0.842	0.725	0.7259
Verdadero positivo SLNB (positivo) y GECL (positivo)	1	0.908	0.691	0.691	N/A

5

Tabla 22b:

Supervivencia específica del melanoma					
Tiempo (Meses)	0	40	80	120	160
Verdadero negativo SLNB (negativo) y GECL (negativo)	1	0.973	0.961	0.961	0.961
Falso negativo SLNB (positivo) y GECL (negativo)	1	0.833	N/A	N/A	N/A
Falso positivo SLNB (negativo) y GECL (positivo)	1	0.945	0.825	0.705	N/A
Verdadero positivo SLNB (positivo) y GECL (positivo)	1	0.907	0.697	0.697	N/A

Tabla 22c:

Supervivencia específica del melanoma					
Tiempo (Meses) 0		40	80	120	160
Verdadero negativo SLNB (negativo) y GECL (negativo)	1	0.975	0.967	0.967	0.967
Falso negativo SLNB (positivo) y GECL (negativo)	1	0.750	N/A	N/A	N/A
Falso positivo SLNB (negativo) y GECL (positivo)	1	0.951	0.855	0.774	0.774
Verdadero positivo SLNB (positivo) y GECL (positivo)	1	0.910	0.699	0.699	N/A

10 Tabla 23a: Estimaciones de supervivencia en diferentes puntos de tiempo para grupos basada en el estatus del clasificador SLNB y GECL. Las curvas de supervivencia se compararon mediante una prueba de rangos logarítmicos, $p < 0.0001$.

Supervivencia libre de reincidencia a distancia					
Tiempo (Meses)	0	40	80	120	160

Verdadero negativo SLNB (negativo) y GECL (negativo)	1	0.980	0.922	0.877	0.877
Falso Negativo SLNB (positivo) y GECL (negativo)	1	0.667	0.500	N/A	N/A
Falso Positivo SLNB (negativo) y GECL (positivo)	1	0.850	0.677	0.547	0.547
Verdadero positivo SLNB (positivo) y GECL (positivo)	1	0.778	0.453	0.396	N/A

Tabla 23b:

Supervivencia libre de reincidencia a distancia					
Tiempo (Meses)	0	40	80 120 160	120	160
Verdadero negativo SLNB (negativo) y GECL (negativo)	1	0.963	0.910	0.869	0.869
Falso negativo SLNB (positivo) y GECL (negativo)	1	0.500	N/A	N/A	N/A
Falso positivo SLNB (negativo) y GECL (positivo)	1	0.844	0.637	0.507	N/A
Verdadero positivo SLNB (positivo) y GECL (positivo)	1	0.763	0.479	0.398	N/A

Tabla 23c:

Supervivencia libre de reincidencia a distancia					
Tiempo (Meses)	0	40	80	120	160
Verdadero negativo SLNB (negativo) y GECL (negativo)	1	0.96 7	0.916	0.86 1	0.861
Falso Negat Negativo SLNB (positivo) y GECL (negativo)	1	0.50 0	N/A	N/A	N/A
Falso Positivo SLNB (negativo) y GECL (positivo)	1	0.87 3	0.700	0.60 1	0.601
Verdadero Positivo SLNB (positivo) y GECL (positivo)	1	0.76 5	0.48032 0	0.39 8	N/A

5

Tabla 24a: Estimaciones de supervivencia en diferentes puntos de tiempo para grupos basada en el estatus del clasificador SLNB y GECL. Las curvas de supervivencia se compararon mediante una prueba de rangos logarítmicos, $p < 0.0001$.

Supervivencia libre de reincidencia					
Tiempo (Meses)	0	40	80	120	160
Verdadero negativo SLNB (negativo) y GECL (negativo)	1				
Falso negativo SLNB (positivo) y GECL (negativo)	1	0.438	0.438	N/A	N/A
Falso positivo (activo) y GECL (positivo)	1	0.723	0.663	0.563	0.563
Verdadero Positivo SLNB (posición) y GECL (positivo)	1	0.604	0.436	0.436	N/A

10

Tabla 24b:

Supervivencia libre de reincidencia					
Tiempo (Meses)	0	40	80	120	160
Verdadero negativo SLNB (negativo) y GECL (negativo)	1	0.925	0.883	0.845	0.845
Falso negativo SLNB (positivo) y GECL (negativo)	1	0.233	N/A	N/A	N/A
Falso positivo SLNB (negativo) y GECL (positivo)	1	0.707	0.636	0.537	N/A
Verdadero positivo SLNB (positivo) y GECL (positivo)	1	0.602	0.473	0.473	N/A

Tabla 24c:

Supervivencia libre de reincidencia					
Tiempo (Meses)	0	40	80	120	160
Verdadero negativo SLNB (negativo) y GECL (negativo)	1	0.939	0.884	0.831	0.831
Falso negativo SLNB (positivo) y GECL (negativo)	1	0.266	N/A	N/A	N/A
Falso positivo SLNB (negativo) y GECL (positivo)	1	0.757	0.704	0.632	0.632
Verdadero positivo SLNB (positivo) y GECL (positivo)	1	0.597	0.470	0.470	N/A

5 Tabla 25a: Proporción de riesgo y valores de p para las 2 curvas de salida del clasificador GECL y los resultados de la biopsia SLNB.

	Proporción de riesgo de supervivencia específica del melanoma (valor p)	Proporción de riesgo de supervivencia libre de reincidencia a distancia (valor p)	Proporción de riesgo de supervivencia libre de reincidencia (valor p)
Expresión génica y variable clínica (GECL)	6.03 (p<0.0001)	5.28 (p<0.0001)	4.60 (p<0.0001)
Hallazgo de biopsia de ganglio linfático centinela (SLNB)	3.21 (p<0.0001)	3.27 (p<0.0001)	3.14 (p<0.0001)

Tabla 25b:

	Proporción de riesgo de supervivencia específica del melanoma (valor p)	Proporción de riesgo de supervivencia libre de reincidencia a distancia (valor p)	Proporción de riesgo de supervivencia libre de reincidencia (valor p)
Expresión génica y variable clínica (GECL)	4.69 (p<0.0001)	4.72 (p<0.0001)	3.95 (p<0.0001)
Hallazgo de biopsia de ganglio linfático centinela (SLNB)	3.36 (p<0.0001)	3.46 (p<0.0001)	3.17 (p<0.0001)

Tabla 25c:

	Proporción de riesgo de supervivencia específica del melanoma (valor p)	Proporción de riesgo de supervivencia libre de recidivencia a distancia (valor p)	Proporción de riesgo de supervivencia libre de recidivencia (valor p)
Expresión génica y variable clínica (GECL)	4.48 (p<0.0001)	4.00 (p<0.0001)	3.30 (p<0.0001)
Hallazgo de biopsia de ganglio linfático centinela (SLNB)	3.36 (p<0.0001)	3.46 (p<0.0001)	3.17 (p<0.0001)

5 Tabla 26a: Proporciones de riesgo multivariados y valores p para las 2 curvas de salida del clasificador GECL y la biopsia SLNB

Tabla 26a	Proporción de riesgo de supervivencia específica del melanoma (valor p)	Proporción de riesgo de supervivencia libre de recidivencia a distancia (valor p)	Proporción de riesgo de supervivencia libre de recidivencia (valor p)
Expresión génica y variable clínica (GECL)	4.92 (p=0.0003)	4.26 (p<0.0001)	3.76 (p<0.0001)
Hallazgo de biopsia de ganglio linfático centinela (SLNB)	2.00 (p=0.0229) 2.10	(p=0.0005) 2.04	(p<0.0001)
Tabla 26b:			
Expresión génica y variable clínica (GECL)	3.70 (p=0.0009)	3.68 (p<0.0001)	3.18 (p<0.0001)
Hallazgo de biopsia del ganglio linfático centinela (SLNB)	1.94 (p=0.03641)	2.00 (p=0.0015)	1.90 (p=0.0011)
Tabla 26c:			
Expresión génica y variable clínica (GECL)	3.45 (p= 0.00628)	3.03 (p=0.00021)	2.59 (p=0.00013)
Hallazgo de biopsia de ganglio linfático centinela (SLNB)	2.33 (p= 0.00611)	2.48 (p<0.0001)	2.35 (p<0.0001)

Tabla 27

Entrada	P%	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	SLNB,RR	ER	AUC	Corte	LRn	LRp
maxbACC	0.17	128	526	90	167	459	38	0.7	0.73	0.92	0.35	0.73	0.72	0.66	0.65	0.79	0.17	0.4	2.6
SEeqSP	0.17	128	526	91	176	450	37	0.71	0.72	0.92	0.34	0.72	0.71	0.65	0.049	0.79	0.17	0.4	2.5
NPV97	0.17	128	526	122	420	206	0	0.95	0.33	0.97	0.23	0.44	0.64	0.26	0.038	0.79	0.11	0.14	1.4
NPV98	0.17	128	526	125	506	120	3	0.98	0.19	0.98	0.2	0.32	0.58	0.16	0.034	0.79	0.1	0.12	1.2

Tabla 28

Entrada	P%	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	SLNB,RR	ER	AUC	Corte	LRn	LRp
maxbACC	0.17	128	626	89	170	456	39	0.7	0.73	0.92	0.35	0.72	0.71	0.86	0.051	0.77	0.17	0.41	2.6
SEeqSP	0.17	128	626	90	174	452	36	0.71	0.72	0.92	0.34	0.72	0.71	0.85	0.05	0.77	0.17	0.41	2.5
NPV97	0.17	128	626	120	407	219	3.4	0.93	0.35	0.96	0.23	0.45	0.64	0.3	0.011	0.77	0.11	0.19	1.4
NPV98	0.17	128	626	123	478	148	5	0.96	0.24	0.97	0.2	0.36	0.6	0.2	0.0066	0.77	0.1	0.16	1.3

Tabla 29

Entrada	P%	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	SLMB,RR	ER	AUC	Corre	LRn	LRp
maxACC	0.17	126	626	97	188	438	31	0.73	0.7	0.93	0.34	0.71	0.73	0.52	0.041	0.8	0.18	0.35	2.5
SE95P	0.17	126	626	95	171	455	33	0.74	0.73	0.93	0.36	0.73	0.73	0.55	0.044	0.8	0.19	0.35	2.7
NPV97	0.17	126	626	120	319	307	8	0.91	0.89	0.97	0.27	0.57	0.71	0.42	0.011	0.8	0.13	0.13	1.8
NPV98	0.17	126	626	124	431	195	4	0.97	0.91	0.93	0.22	0.42	0.54	0.26	0.0053	0.8	0.099	0.1	1.4

Tabla 30

Entidad	P%	P	N	TP	FP	TM	FN	SF	SP	NPV	PPV	ACC	hACC	SIMPRR	FR	AIC	Coste	IRn	iRp
maxhACC	0.17	128	626	88	181	445	40	0.69	0.71	0.92	0.33	0.71	0.7	0.64	0.055	0.76	0.18	0.44	2.4
SEagSP	0.17	128	626	88	187	458	42	0.67	0.73	0.92	0.34	0.72	0.7	0.67	0.056	0.78	0.19	0.45	2.5
NPV97	0.17	128	626	112	305	323	16	0.87	0.91	0.95	0.27	0.97	0.69	0.45	0.022	0.76	0.13	0.25	1.8
NPV98	0.17	128	626	123	408	218	7.0	0.94	0.95	0.97	0.23	0.45	0.64	0.3	0.01	0.78	0.099	0.17	1.5

Tabla 31

Enfada	P%	P	N	TP	FP	TN	FN	SF	SP	NPV	PPV	ACC	hACC	SIMP.RR	FR	AIC	Cofie	IRa	IRp
mexhACC	0.17	128	626	182	158	478	26	0.6	0.76	0.95	0.4	0.77	0.78	0.67	0.034	0.84	0.19	0.27	3.2
SFesqSP	0.17	128	626	90	109	487	29	0.77	0.78	0.94	0.47	0.79	0.78	0.68	0.038	0.84	0.19	0.29	3.5
NPV97	0.17	128	626	114	267	359	14	0.89	0.67	0.96	0.3	0.63	0.73	0.49	0.019	0.84	0.13	0.19	2.1
NPV98	0.17	128	626	122	383	243	6	0.95	0.59	0.98	0.24	0.48	0.67	0.33	0.008	0.84	0.098	0.12	1.6

Tabla 32

Fracción	P%	P	N	TP	FP	TN	FN	SF	SP	NPV	PPV	ACC	bACC	SIMLRR	FR	AUIC	Cante	IRp	IRp
max%ACC	0.17	128	626	94	158	475	34	0.73	0.76	0.90	0.39	0.75	0.75	0.58	0.045	0.82	0.19	0.35	3.1
SEeqSP	0.17	128	626	92	142	484	36	0.72	0.77	0.93	0.39	0.73	0.74	0.59	0.048	0.82	0.19	0.37	3.2
NPV97	0.17	128	626	112	255	371	16	0.87	0.59	0.96	0.31	0.64	0.73	0.51	0.022	0.82	0.13	0.22	2.2
NPV96	0.17	128	626	119	369	257	83	0.93	0.41	0.97	0.24	0.5	0.67	0.35	0.012	0.82	0.097	0.17	1.8

Tabla 33

maxbACC		P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	\$\Delta\$NB,RR	ER	AUC	Corte	LRn	LRp
Entrada	P%	128	626	8	170	456	3	0.7	0.73	0.9	0.35	0.7	0.71	0.66	0.051	0.7	0.17	0.41	2.6
				9			9			2		2				7			
GE	0.17	128	626	8	181	445	4	0.69	0.71	0.9	0.33	0.7	0.7	0.64	0.053	0.7	0.18	0.44	2.4
				8			0			2		1				8			
GECL	0.17	128	626	9	150	476	3	0.73	0.76	0.9	0.39	0.7	0.75	0.68	0.045	0.8	0.19	0.35	3.1
				4			4			3		6				2			

Tabla 34

SEeqSP	Entrada	P%	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	SLNB.RR	ER	AUC	Corte	LRn	LRp
	CL	0.17	128	626	9	174	452	3	0.71	0.72	0.92	0.34	0.7	0.71	0.65	0.05	0.7	0.17	0.41	2.5
	GE	0.17	128	626	8	167	459	4	0.67	0.73	0.92	0.34	0.7	0.7	0.67	0.056	0.7	0.19	0.45	2.5
	GECL	0.17	128	626	9	142	484	3	0.72	0.77	0.93	0.39	0.7	0.74	0.69	0.048	0.8	0.19	0.37	3.2

Tabla 35

NPV 97		P%	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	SLNB.RR	ER	AUC	Corfe	LRn	LRp
Entrada	CL	0.17	128	626	120	407	210	8.4	0.93	0.35	0.96	0.23	0.4	0.64	0.3	0.011	0.7	0.11	0.19	1.4
													5				7			
	GE	0.17	128	626	112	305	321	16	0.87	0.51	0.95	0.27	0.5	0.69	0.45	0.022	0.7	0.13	0.25	1.8
													7				8			
	GECL	0.17	128	626	112	255	371	16	0.87	0.59	0.96	0.31	0.6	0.73	0.51	0.022	0.8	0.13	0.22	2.2
													4				2			

Tabla 36 NPV 98

Entrada	P%	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	SLNB.RR	ER	AUC	Corte	LRn	LRp
CL	0.17	128	626	123	478	148	5	0.96	0.24	0.97	0.2	0.3 6	0.6	0.2	0.0068	0.7 7	0.1	0.16	1.3
GE	0.17	128	626	120	408	21.8	7.6	0.94	0.35	0.97	0.23	0.4 5	0.64	0.3	0.01	0.7 8	0.099	0.17	1.5
GECL	0.17	128	626	119	369	257	8.8	0.93	0.41	0.97	0.24	0. 5	0.67	0.35	0.012	0.8 2	0.097	0.17	1.6

Tabla 37 CL

	Entrada	P%	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	SLNE,RR	ER	AUC	Corfe	LRn	LRp
1	T1b	0.003	6	178	3.4	47	131	2.6	0.57	0.74	0.98	0.07	0.73	0.66	0.73	0.014	0.68	0.11	0.57	2.3
	T1	0.13	39	253	34	176	77	5.1	0.87	0.31	0.94	0.16	0.38	0.59	0.28	0.017	0.68	0.11	0.42	1.3
	T2a																			
12	T2b	0.17	11	10	48	7.3	0.79	0.93	0.13	0.89	0.18	0.27	0.53	0.12	0.012	0.7	0.11	0.66	1.1	
	T3a	0.31	38	85	38	85	0.04	0	1	0.00048	1	0.31	0.31	0.5	0.00032	0	0.62	0.11	0	1
14	T3b	0.43	34	45	34	45	0.01	0	1	0.00022	1	0.43	0.43	0.5	0.00013	0	0.6	0.11	0	1

Tabla 38GE

Entrada	P%	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	SLMB,RR	ER	AUC	Corfe	LRn	LRp
F10	0.033	6	108	23	51	127	37	0.38	0.71	0.97	0.032	0.7	0.56	0.71	0.02	0.5	0.13	0.87	1.3
T2a	0.13	39	263	30	119	134	8.9	0.77	0.53	0.94	0.2	0.5	0.66	0.49	0.03	0.7	0.13	0.43	1.0
T2b	0.17	41	55	11	39	16	0.19	0.98	0.78	0.99	0.71	0.4	0.63	0.74	0.029	0.6	0.13	0.059	1.4
T3a	0.31	38	85	35	56	25	3.2	0.91	0.34	0.9	0.38	0.5	0.63	0.26	0.026	0.7	0.13	0.25	1.4
T3b	0.43	34	45	34	36	9.1	0.27	0.99	0.2	0.98	0.49	0.5	0.6	0.12	0.035	0.7	0.13	0.037	1.2

Tabla 39 GECL

Entrada	F%	P	N	TP	FP	FN	SE	SP	NPV	PPV	ACC	bACC	SLMB.RR	ER	AUC	Cofie	LRD	LRD
T1b	0.093	6	178	17	18	159	4.3	0.38	0.99	0.97	0.952	0.87	0.58	0.89	0.69	0.43	0.81	2.7
T2a	0.13	29	253	29	88	165	9.3	0.75	0.55	0.94	0.25	0.68	0.7	0.6	0.75	0.40	0.39	2.2
T2b	0.17	11	55	9.8	35	20	1.2	0.89	0.38	0.94	0.22	0.45	0.92	0.32	0.72	0.43	0.31	1.4
T3a	0.31	38	85	37	78	15	1.1	0.97	0.17	0.94	0.35	0.42	0.57	0.19	0.75	0.43	0.15	1.2
T3b	0.43	24	45	34	48	5.1	0.93	1	0.11	1	0.46	0.49	0.56	0.065	0.74	0.43	0.0857	1.1

Tabla 48CL

Embarada	P%	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	SACC	SNBRR	ER	AUC	CoRR	LRE	LRD
T1b	0.033	6	170	4.2	00	98	1.9	0.66	0.55	0.20	0.051	0.56	0.62	0.54	0.01	0.68	0.1	0.56	1.6
T7a	0.13	39	253	36	209	44	2.8	0.95	0.17	0.94	0.15	0.78	0.55	0.16	0.0094	0.68	0.1	0.4	1.1
T2b	0.17	11	55	11	52	3.3	0.34	0.97	0.06	0.92	0.17	0.21	0.52	0.058	0.0051	0.7	0.1	0.58	1
T3a	0.31	38	85	38	85	0.01	0	1	0.00012	1	0.31	0.31	0.5	8.4e-36	0	0.62	0.1	0	1
T3b	0.43	34	45	34	45	0.01	0	1	0.00072	1	0.43	0.43	0.5	0.00013	0	0.5	0.1	0	1

Tabla 41 CC

Entrada	P%	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	BACC	SLNBERR	ER	AUC	Conto	LRn	LRf
T1E	0.033	6	172	3.6	84	94	2.4	0.61	0.59	0.98	0.042	0.5	0.57	0.52	0.013	0.5	0.039	0.75	1.3
T2e	0.13	39	253	35	166	87	4.3	0.89	0.35	0.95	0.17	0.4	0.62	0.31	0.015	0.7	0.039	0.32	1.4
T2E	0.17	11	55	11	45	9.5	0.02	1	0.17	1	0.19	0.3	0.39	0.15	3e-04	0.6	0.039	0.0683	1.2

Tabla 41GE

Entrada	P%	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	SLNB:RR	ER	AUC	Corte	LRn	LRP
F3a	0.31	36	85	3	6	16	0.86	0.98	0.19	0.96	1.35	0.4	0.58	0.14	0.0062	0.7	0.099	0.11	1.2
				7	9							3				3			
F3b	0.43	34		3	4	5	0.05	1	0.11	0.99	0.46	0.4	0.55	0.083	0.00065	0.7	0.099	0.014	1.1
		45		4	0							9				3			

Tabla 42

Entrada	P%	P	N	TPFP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	SLNBRR	ER	AUC	Corte	LRn	LRp	GECL		
																			34	45	
T1b	0.033	6	178	2.8	52	126	3.2	0.47	0.71	0.98	0.653	0.7	0.59	0.7	0.017	0.59	0.097	0.75	1.7		
T2a	0.13	39	253	34	147	106	5.2	0.87	0.42	0.95	0.19	0.48	0.64	0.38	0.018	0.75	0.097	0.32	1.5		
T2b	0.17	11	55	11	44	11	0.26	0.98	0.2	0.98	0.2	0.33	0.59	0.17	0.0039	0.72	0.097	0.1	1.2		
T3a	0.31	38	85	38	79	5.9	0.13	1	0.069	0.99	0.32	0.36	0.53	0.049	0.0011	0.75	0.097	0.037	1.1		
T3b	0.43			34	43	2.4	0	1	0.054	1	0.44	0.46	0.52	0.031	0	0.74	0.097	0	1		

Tabla 43

CL

Entrada	P%	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	SLNB,RR	ER	AUC	Corte	LRn	LRp
estadio IB	0.095	45	431	37	222	209	7.6	0.83	0.48	0.96	0.14	0.52	0.66	0.45	0.016	0.73	0.11	0.35	1.6
estadio IIA	0.26	49	140	48	133	7.4	0.79	0.98	0.053	0.89	0.27	0.3	0.52	0.043	0.0042	0.67	0.11	0.37	1
estadio IIB	0.43	34	34	45	45	0.01	0	1	0.00022	1	0.43	0.43	0.5	0.00013	0	0.6	0.11	0	1

45

Tabla 44

GE

Entrada	P%	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	sLNB.RR	ER	AUC	Corte	LRn	LRp
estadio IB	0.095	45	431	32	171	260	13	0.72	0.6	0.95	0.16	0.62	0.68	0.57	0.026	0.72	0.13	0.46	1.8
estadio IIA	0.26	40	140	46	96	44	3.4	0.93	0.32	0.93	0.32	0.47	0.62	0.25	0.018	0.71	0.13	0.22	1.4
estadio IIB	0.43	34	45	34	36	9.1	0.27	0.99	0.2	0.98	0.49	0.54	0.6	0.12	0.0035	0.73	0.13	0.037	1.2

Tabla 45

GECL

Entrada	P%	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	SLNBRR	ER	AUC	Corte	LRn	LRp
estadio IB	0.085	45	431	31	108	323	14	0.69	0.75	0.96	0.23	0.74	0.72	0.71	0.03	0.77	0.13	0.42	2.8
estadio IIA	0.26	49	140	47	105	35	2.3	0.95	0.25	0.94	0.31	0.43	0.6	0.19	0.012	0.76	0.13	0.18	1.3
estadio IIB	0.43	34		34	40	5.1	0.03	1	0.11	1	0.46	0.49	0.56	0.065	0.00039	0.74	0.13	0.0067	1.1

45

Tabla 46

CL

Entrada	P%	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	SLNBRR	ER	AUC	Corte	LRn	LRp
estadio IB	0.095	45	431	40	289	142	4.6	0.9	0.33	0.97	0.12	0.38	0.61	0.31	0.0099	0.73	0.1	0.31	1.3
estadio IIA	0.26	49	140	49	137	3.3	0.34	0.99	0.024	0.92	0.26	0.28	0.51	0.019	0.0018	0.67	0.1	0.31	1
estadio IIB	0.43	34	45	34	45	0.01	0	1	0.00022	1	0.43	0.43	0.5	0.00013	0	0.6	0.1	0	1

Tabla 47

GE

Entrada	P%	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	bACC	SLNB,RR	ER	AUC	Corte	LRn	LRp
estadio IB	0.095	45	431	38	250	181	6.7	0.85	0.42	0.96	0.13	0.46	0.64	0.4	0.014	0.72	0.099	0.35	1.5
estadio IIA	0.26	49	140	48	114	26	0.88	0.98	0.18	0.97	0.3	0.39	0.58	0.14	0.0047	0.71	0.099	0.098	1.2
estadio IIB	0.43	34	45	34	40	5	0.05	1	0.11	0.99	0.46	0.49	0.55	0.063	0.00065	0.73	0.099	0.014	1.1

Tabla 40

CLGE

Entrada	P%	P	N	TP	FP	TN	FN	SE	SP	NPV	PPV	ACC	ΔACC	SLIME,RR	ER	AUC	Core	LRn	LRp
estadio B	0.095	45	431	37	200	231	8.4	0.31	0.54	0.67	0.16	0.56	0.56	0.5	0.018	0.77	0.097	0.34	1.8
estadio BA	0.25	49	140	49	123	17	0.39	0.99	0.12	0.50	0.20	0.35	0.56	0.082	0.0321	0.76	0.097	0.055	1.1
estadio BB	0.43	34	45	34	43	24	0	1	0.054	1	0.44	0.46	0.52	0.031	0	0.74	0.097	0	1

Tabla 49

AUC mínima y máxima

	subconjunto de 2	subconjunto de 3	subconjunto de 4	subconjunto de 5	subconjunto de 6	subconjunto de 7	subconjunto de 8
n subconjuntos	36	84	126	126	84	36	9
AUC min	0.68	0.71	0.73	0.76	0.78	0.79	0.81
AUC máx	0.79	0.8	0.8	0.81	0.81	0.82	0.82

REIVINDICACIONES

- 5 1. Un método para clasificar a un individuo que padece melanoma cutáneo primario, que comprende determinar en una muestra de tejido de melanoma de dicho individuo una firma de expresión génica, en el que la firma de expresión génica comprende ITGB3, PLAT, GDF15 e IL8 y en el que el individuo se clasifica como que tiene un SLN positivo para metástasis o se clasifica como que tiene un SLN negativo para metástasis.
- 10 2. Un método para determinar una programación de tratamiento y/o diagnóstico para un individuo que padece melanoma cutáneo, que comprende determinar en una muestra de tejido de melanoma de dicho individuo una firma de expresión génica que comprende determinar el nivel de expresión de ITGB3, PLAT, GDF15 e IL8 y determinar una programación de tratamiento y/o diagnóstico basado en los niveles de expresión.
- 15 3. Un método para predecir el pronóstico de un individuo que padece melanoma cutáneo primario, que comprende determinar en una muestra de tejido de melanoma de dicho individuo una firma de expresión génica, en el que la firma de expresión génica comprende ITGB3, PLAT, GDF15 e IL8 y el pronóstico del individuo es determinado basado en los niveles de expresión génica.
4. El método de la reivindicación 2 o 3, en el que el individuo se clasifica como que tiene un SLN positivo para metástasis o se clasifica como que tiene un SLN negativo para metástasis.
5. El método de cualquiera de las reivindicaciones anteriores, en el que el individuo se selecciona para SLNB basándose en dicha clasificación y/o niveles de expresión.
6. El método de cualquiera de las reivindicaciones anteriores, en el que un individuo clasificado como que tiene un SLN positivo para metástasis se trata realizando una SLNB.
- 20 7. El método de una cualquiera de las reivindicaciones 2-5, en el que un individuo se clasifica como que tiene un pronóstico pobre o un pronóstico bueno.
- 25 8. El método de cualquiera de las reivindicaciones anteriores, que comprende además clasificar a un individuo que padece melanoma cutáneo primario, que comprende determinar en una muestra de dicho individuo una firma de expresión génica, en el que la firma de expresión génica comprende al menos uno de los siguientes genes: KRT14, SPP1, FN1, LOXL3.
9. Un tratamiento contra el cáncer para utilizar en el tratamiento de un individuo que padece melanoma cutáneo primario, que comprende:
 - determinar en una muestra de tejido de melanoma de dicho individuo una firma de expresión génica, en el que la firma de expresión génica comprende ITGB3, PLAT, GDF15 e IL8,
 - 30 - clasificar dicho individuo como que tiene un SLN positivo para metástasis y/o un pronóstico pobre basándose en la firma de expresión génica, y
 - tratar a dicho individuo realizando una SLNB y/o proporcionando un tratamiento contra el cáncer a dicho individuo, en el que dicho tratamiento contra el cáncer se selecciona de
 - 35 - quimioterapia, preferiblemente en el que dicha quimioterapia se selecciona de dacarbazina, temozolomida, carboplatino, paclitaxel, cisplatino, y vinblastina;
 - radioterapia;
 - inmunoterapia, preferiblemente en el que dicha inmunoterapia se selecciona de citocinas, inhibidores de puntos de control inmunes, e inmunoterapia oncolítica; y
 - 40 - fármacos de terapia dirigida seleccionados de inhibidores de BRAF, preferiblemente seleccionados de vemurafenib o dabrafenib, e inhibidores de MEK, preferiblemente seleccionados de cobimetinib y trametinib.
10. El tratamiento contra el cáncer de la reivindicación 9, en el que dicha inmunoterapia se selecciona de citocinas, preferiblemente Interferón alfa-2b o interleucina-2; inhibidores de puntos de control inmune; o inmunoterapia oncolítica.
11. El tratamiento contra el cáncer de la reivindicación 9, en el que dicho tratamiento contra el cáncer se selecciona de ipilimumab, nivolumab, y pembrolizumab.
- 45 12. El método de una cualquiera de las reivindicaciones 1-8 o el tratamiento contra el cáncer para uso de una cualquiera de las reivindicaciones 9-11, en el que la firma de expresión génica comprende además MLANA, LOXL4, SERPINE2, y TGFBR1.
13. El método o tratamiento contra el cáncer para uso de la reivindicación 12, en el que la firma de expresión génica comprende además ADIPOQ.

14. El método o tratamiento contra el cáncer para uso de la reivindicación 12 o 13, en el que la firma de expresión génica comprende además PRKCB, ADAM12, LGALS1.

Figura 1

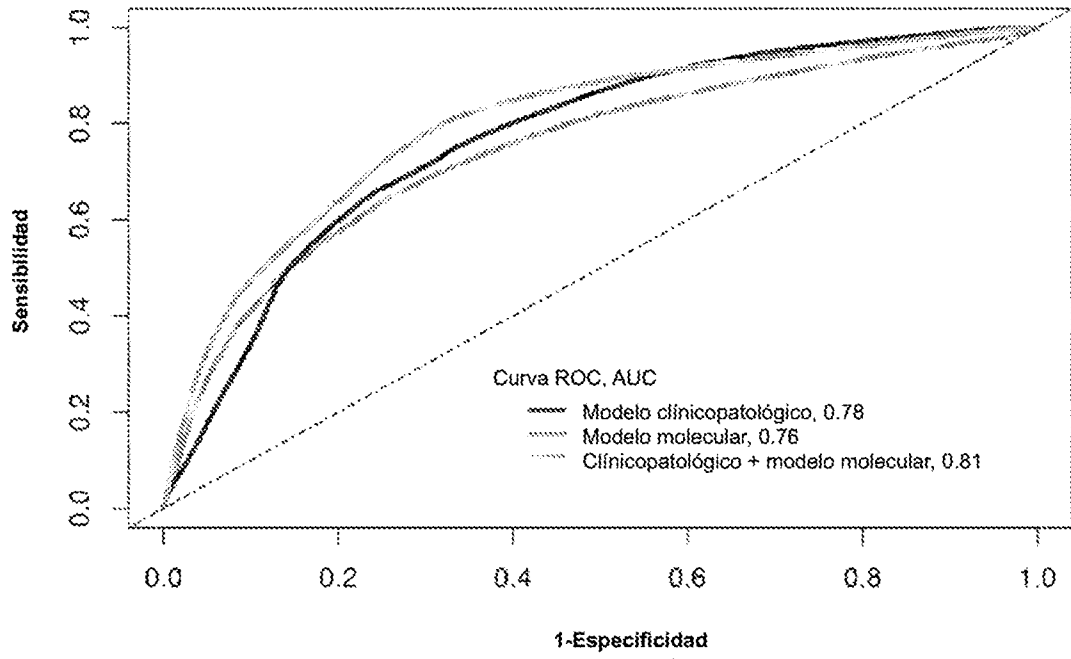


Figura 2

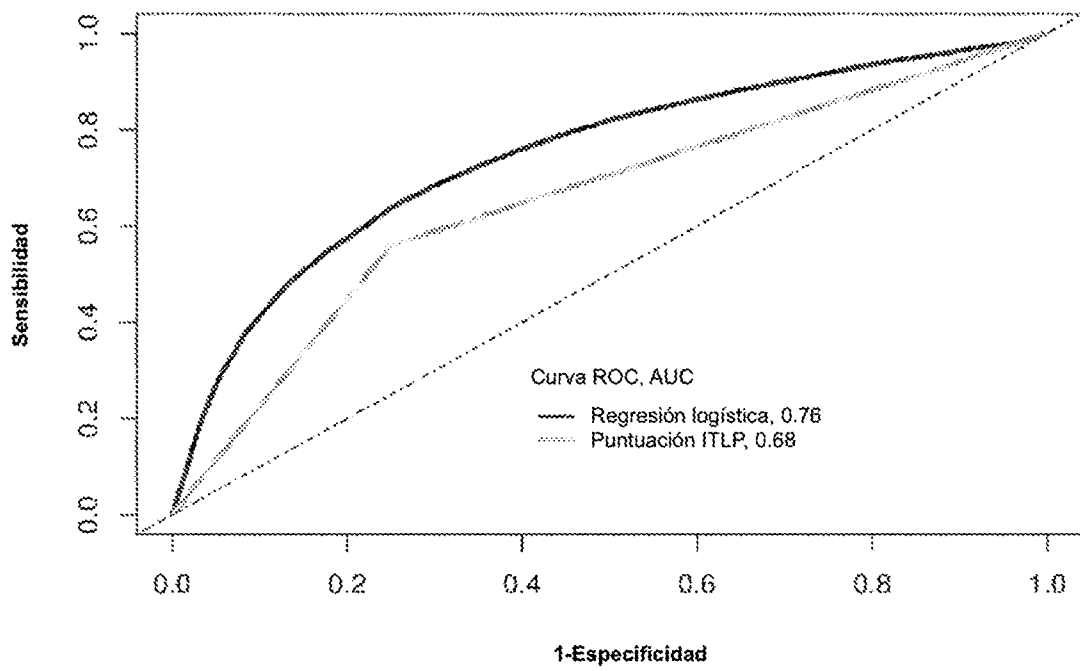


Figura 3

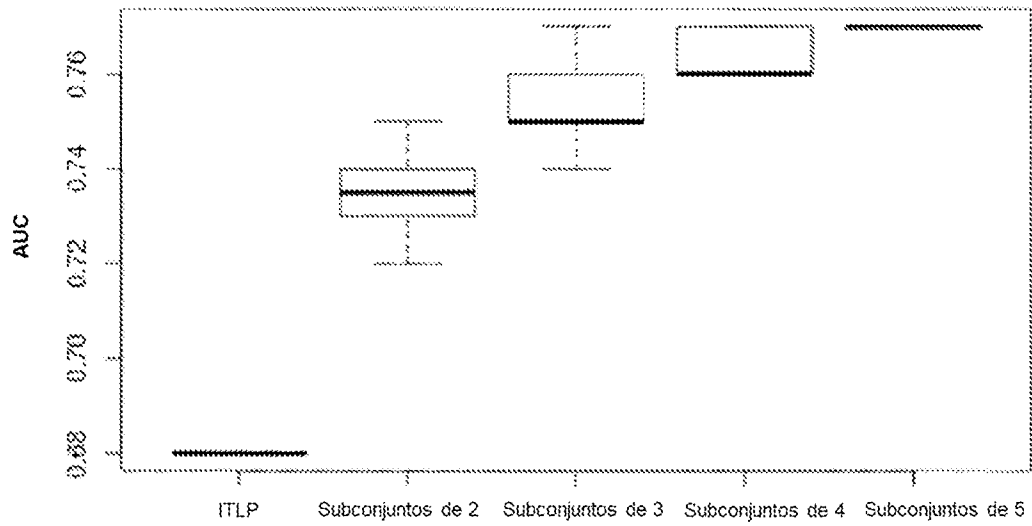


Figura 4

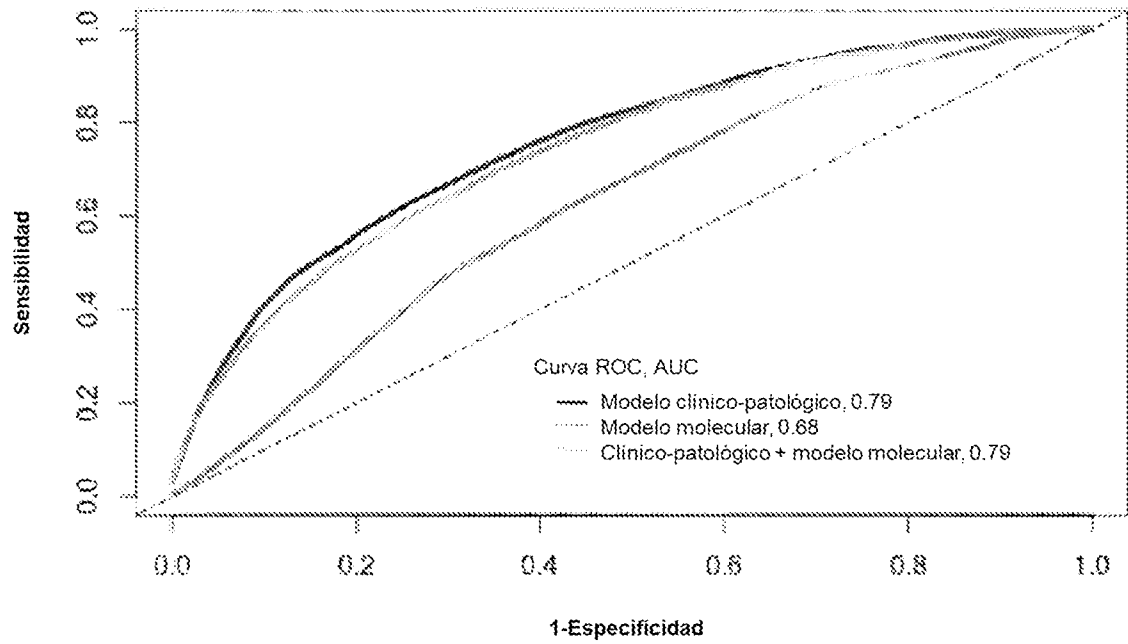


Figura 5

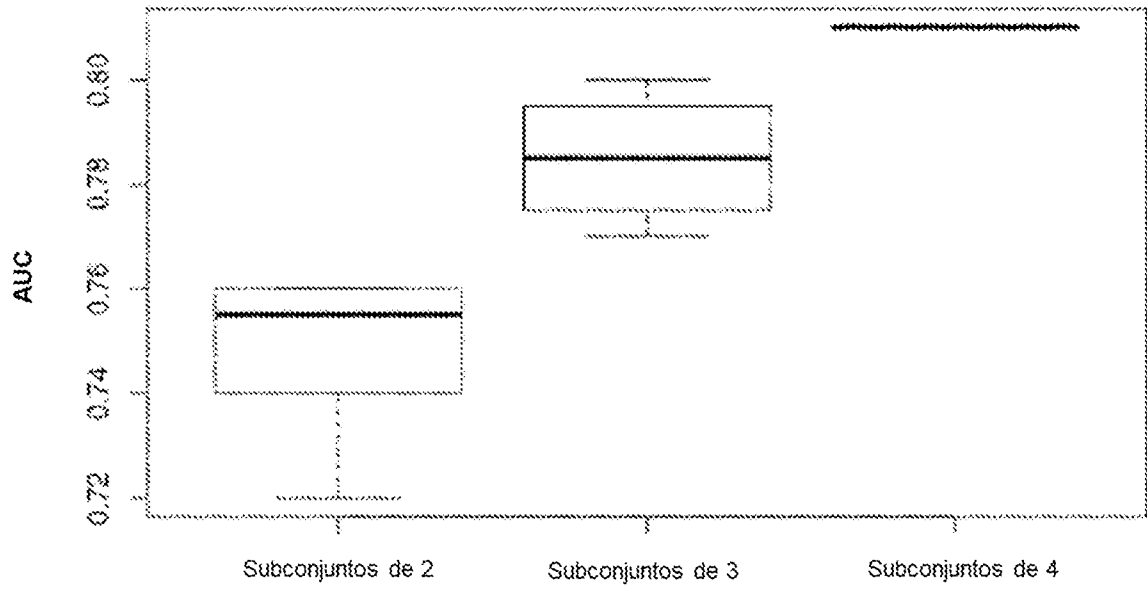


Figura 6

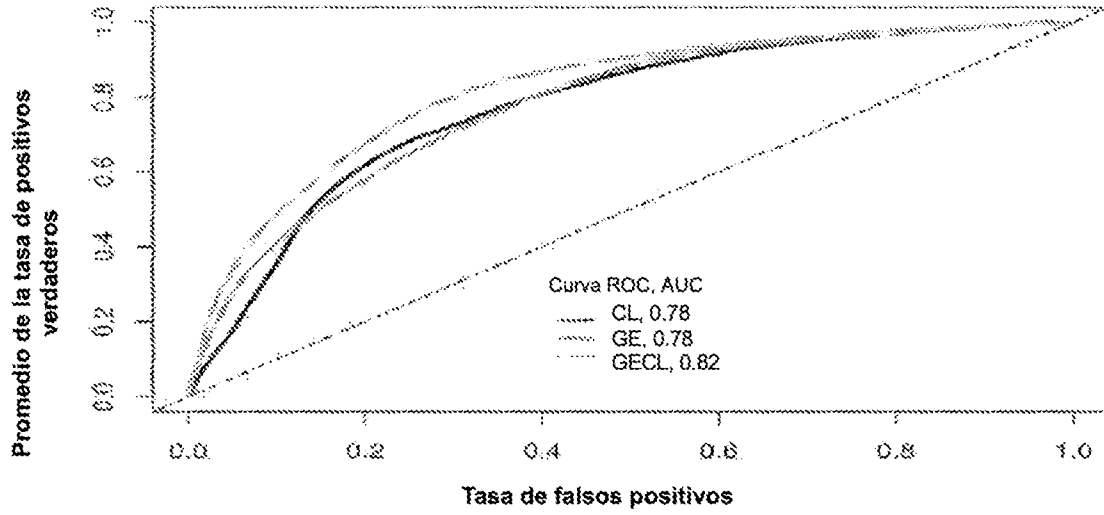


Figura 7

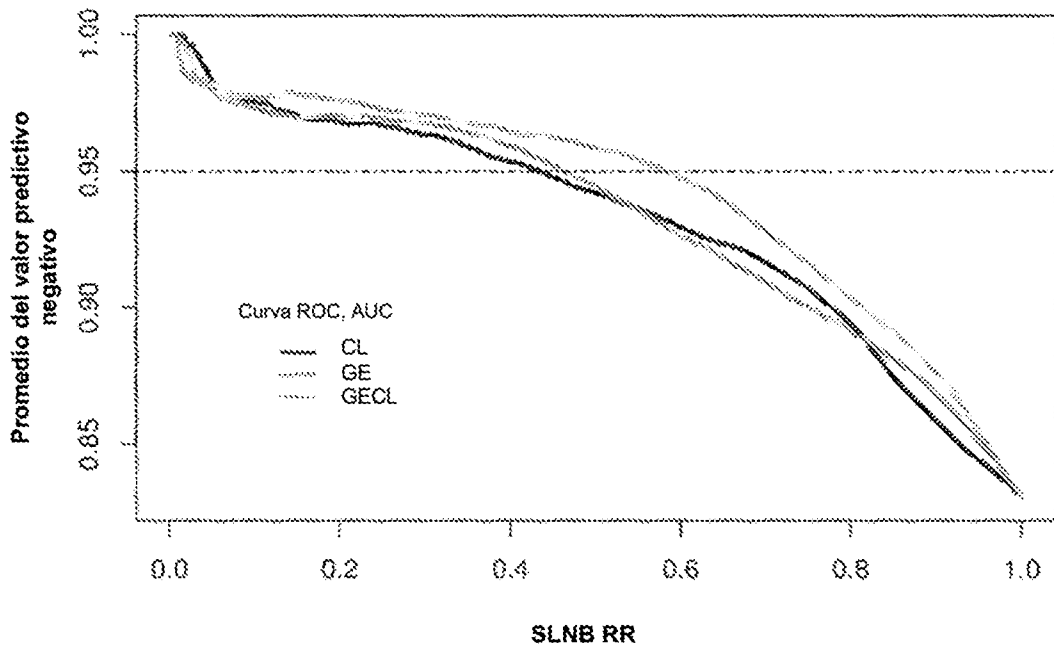


Figura 8

