(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 640 524**

(21) Número de solicitud: 201630400

(51) Int. CI.:

C12Q 1/68 (2006.01)

(12) PATENTE DE INVENCIÓN B1

(22) Fecha de presentación:

01.04.2016

(43) Fecha de publicación de la solicitud:

03.11.2017

Fecha de concesión:

17.09.2018

(45) Fecha de publicación de la concesión:

24.09.2018

(56) Se remite a la solicitud internacional:

PCT/ES2017/070201

(73) Titular/es:

UNIVERSIDAD AUTÓNOMA DE MADRID (90.0%)
Ciudad Universitaria de Cantoblanco, Pabellón C, Calle Einstein, 13
28049 Madrid (Madrid) ES y
FUNDACIÓN PARA LA INVESTIGACIÓN BIOMÉDICA DEL HOSPITAL INFANTIL UNIVERSITARIO NIÑO JESÚS (10.0%)

(72) Inventor/es:

FRESNO ESCUDERO, Manuel;
GIRONÈS PUJOL, Núria y
RAMÍREZ ORELLANA, Manuel

(74) Agente/Representante:

MARTÍN BADAJOZ, Irene

(54) Título: **USO DE TCFL5/CHA COMO NUEVO MARCADOR PARA EL PRONÓSTICO Y/O DIAGNÓSTICO DIFERENCIAL DE LEUCEMIAS LINFOBLÁSTICAS AGUDAS**

Aviso: Se puede realizar consulta prevista por el art. 37.3.8 LP 11/1986.

ES 2 640 524 B1

(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 640 524**

(21) Número de solicitud: 201630400

(57) Resumen:

La presente invención proporciona un método pronóstico y/o para el diagnóstico diferencial de pacientes con leucemia linfoblástica aguda (LLA) que comprende las siguientes etapas:

a) determinar de manera simultánea los niveles de expresión de las isoformas TCFL5 y CHA del gen TCFL5 en una muestra biológica aislada de dicho paciente; y

b) comparar los niveles de expresión de TCFL5/CHA en la muestra del paciente con valores de referencia,

donde una reducción en los valores de la muestra del paciente respecto a los valores de referencia es indicativa de LLA de alto riesgo,

donde preferiblemente dichas muestras de referencia corresponden a una o más muestras de pacientes con LLA o a líneas establecidas de LLA. Proporciona también un kit para la implementación de dicho método. Asimismo, se refiere a un método para el diagnóstico diferencial, la clasificación o estratificación de los pacientes en base a dicho pronóstico; un método para el tratamiento; y un método para la monitorización de la progresión de la enfermedad y/o eficacia del tratamiento en pacientes con LLA.

ES 2 640 524 B1

Aviso: Se puede realizar consulta prevista por el art. 37.3.8 LP 11/1986.

DESCRIPCIÓN

USO DE TCFL5/CHA COMO NUEVO MARCADOR PARA EL PRONÓSTICO Y/O DIAGNÓSTICO DIFERENCIAL DE LEUCEMIAS LINFOBLÁSTICAS AGUDAS

CAMPO DE LA INVENCION

La presente invención pertenece al campo técnico de biotecnología y medicina. En particular, proporciona un método pronóstico de leucemias linfoblásticas agudas (LLA) basado en la determinación de los niveles de expresión de todas y cada una de las isoformas del gen *TCFL5*, incluyendo CHA. Proporciona también un kit para la implementación de dicho método. Asimismo, se refiere a un método para el diagnóstico diferencial, la clasificación o estratificación de los pacientes en base a dicho pronóstico; un método para el tratamiento; y un método para la monitorización de la progresión de la enfermedad y/o eficacia del tratamiento en pacientes con LLA.

ANTECEDENTES DE LA INVENCION

Las leucemias linfoblásticas agudas T y B (T-LLA y B-LLA) son un grupo de neoplasias malignas caracterizadas por la invasión de la médula ósea de precursores hematopoyéticos (*Inaba, H., M. Greaves, and C.G. Mullighan, Acute lymphoblastic leukaemia. Lancet, 2013. 381(9881): p. 1943-55*). Las células leucémicas inmaduras usualmente invaden la sangre con bastante rapidez. Estas células se pueden propagar a otras partes del cuerpo, como a los ganglios linfáticos, el hígado, el bazo, el sistema nervioso central (el cerebro y la médula espinal) y los testículos (en los hombres).

La LLA representa el 12% de las leucemias diagnosticadas en EEUU y Europa y el 60% de todos los casos ocurre en personas menores de 20 años. La LLA es la leucemia más frecuente en niños (3 de cada 4) y el coste del tratamiento se sitúa en aproximadamente 100.000€ por paciente. La LLA supera los 6.000 nuevos casos al año en EEUU y un porcentaje similar en Europa. Alrededor de 20% de los pacientes presenta fracaso al tratamiento, incluyendo pacientes que se consideraron inicialmente como de bajo riesgo.

Las LLA recidivantes pueden considerarse como el quinto cáncer más común en niños y representa un porcentaje importante de las muertes anuales de niños asociadas a cáncer (*Bhojwani, D. and C.H. Pui, Relapsed childhood acute lymphoblastic leukaemia. Lancet Oncol, 2013. 14(6): p. e205-17*). Un aspecto importante en el manejo clínico de B-LLA es la intensidad de la quimioterapia en función del riesgo de recaída (*Zhou, Y., et al., Advances in the molecular pathobiology of B-lymphoblastic leukemia. Hum Pathol, 2012. 43(9): p. 1347-62*). Las T-LLA representan el 10%-15% de las LLA pediátricas y el 25% de los casos de

LLA de adultos y sigue siendo una enfermedad agresiva con supervivencia libre de enfermedad a 5 años entre el 20 y el 50%. El pronóstico sigue siendo pobre en niños y adultos. Las recaídas tempranas son comunes en T-LLA y se asocian con un mal pronóstico. La base del tratamiento actual es el diagnóstico hemopatológico y se pueden agrupar en subtipos, basado en translocaciones cromosómicas y en perfiles de expresión génica (*Inaba, H., et al., Acute lymphoblastic leukaemia. Lancet, 2013. 381(9881): p. 1943-55*).

TCFL5 es un factor de transcripción de la familia basic helix-loop-helix y uno de los 14 genes que caracterizan las leucemias TEL/AML1-positivas (*Gandemer V, Rio AG, de Tayrac M, et al. Five distinct biological processes and 14 differentially expressed genes characterize TEL/AML1-positive leukemia. BMC Genomics. 2007;8:385*). Asimismo, ha sido identificado como uno de los 5 factores genéticos asociados a la resistencia a vincristina en LLA (*Silveira VS, et al. Gene expression pattern contributing to prognostic factors in childhood acute lymphoblastic leukemia. Leuk Lymphoma. 2013;54:310-314, Tabla 1*). *Silveira VS, et al.* también describen una asociación significativa entre niveles más altos de expresión del gen *TCFL5* y el estado de la médula ósea y con un fenotipo B-LLA ETV6/RUNX1 negativo.

TCFL5 ha sido también descrito como un gen diana de NOTCH-1, un importante mediador en el desarrollo de leucemias T-LLA (*Weerkamp, F., et al., Identification of Notch target genes in uncommitted T-cell progenitors: No direct induction of a T-cell specific gene program. Leukemia, 2006. 20(11): p. 1967-77*). El factor de transcripción CHA fue clonado por primera vez por los inventores, siendo descrito como una isoforma del gen *TCFL5* (*Rodriguez, C.I., N. Girones, and M. Fresno, Cha, a basic helix-loop-helix transcription factor involved in the regulation of upstream stimulatory factor activity. J Biol Chem, 2003. 278(44): p. 43135-45*).

La estratificación de los pacientes de LLA permite adaptar la intensidad de la terapia al riesgo de recaída del paciente, contribuyendo así a mejorar la tasa de supervivencia de los pacientes y/o a evitar recaídas. Los perfiles de expresión génica pueden proporcionar una buena herramienta para la subclasificación y la estratificación terapéutica de los pacientes (*Zhou, Y., et al., Advances in the molecular pathobiology of B-lymphoblastic leukemia. Hum Pathol 2012. 43: 1347-1362; Mullighan, C. G., Molecular genetics of B-precursor acute lymphoblastic leukemia. J Clin Invest 2012. 122: 3407-3415; Mullighan, C. G., Genomic profiling of B-progenitor acute lymphoblastic leukemia. Best Pract Res Clin Haematol 2011. 24: 489-503; Harvey, R. C., et al., Identification of novel cluster groups in pediatric high-risk*

B-precursor acute lymphoblastic leukemia with gene expression profiling: correlation with genome-wide DNA copy number alterations, clinical characteristics, and outcome. Blood 2010. 116: 4874-4884;Van Vlierberghe P, Ferrando A. The molecular basis of T cell acute lymphoblastic leukemia. J Clin Invest. 2012;122:3398-3406).

En EEUU se realizó un gran ensayo clínico denominado “high-risk ALL Therapeutically Applicable Research to Generate Effective Treatments” (TARGET) fruto de la colaboración entre el Instituto Nacional del Cáncer (NCI) y el Grupo de Oncología Infantil (COG), para la identificación y validación de nuevas dianas terapéuticas. Se analizaron 207 muestras de 272 (75%) pacientes de LLA de células B precursoras de alto riesgo de los pacientes participantes en el ensayo clínico COG p9906 en un esfuerzo para identificar subgrupos de estos pacientes de alto riesgo caracterizados por presentar perfiles de expresión génica característicos. Los resultados de dicho estudio se describen en la solicitud de patente WO2009/064481. En particular, se identificaron marcadores genéticos asociados con alto riesgo en pacientes de B-LLA y el uso de los mismos para el diagnóstico, pronóstico y/o en la determinación de la eficacia del tratamiento. Más específicamente los métodos descritos comprenden la determinación de los niveles de expresión de genes seleccionados del siguiente grupo: CENTG2 (ArfGAP with GTPase domain, ankyrin repeat and PH domain 1); PTPRM (protein tyrosine phosphatase, receptor type, M); STAPl (signal transducing adaptor family member 1); CCNJ (cyclin J); PCDH17 (procadherin 17); MCAM (melanoma cell adhesion molecule); CAPN3 (calpain 3); CABLESl (Cdk5 and Abl enzyme substrate 1); GPR155 (G protein-coupled receptor 155); MUC4 (mucin 4); GPRl 10 (G protein-coupled receptor 110); IGJ (immunoglobulin J polypeptide); NRXN3 (neurexin 3); CD99 (CD99 molecule); CRLF2 (cytokine receptor-like factor 2); ENAM (enamelin); TP53INP1 (tumor protein p53 inducible nuclear protein 1); IFITMl (interferon induced transmembrane protein 1); IFITM2 (interferon induced transmembrane protein 2); JJTTM3 (interferon induced transmembrane protein 3); TTYH2 (tweety homolog 2); SEMA6A (semaphorin 6A); TNFSF4 (tumor necrosis factor superfamily, member 4); y SLC37A3 (solute carrier family 37, member 3).

WO2015/092755 reporta un método para el diagnóstico o pronóstico de leucemia linfoblástica aguda de alta hiperdiploidia (HeH-ALL en sus siglas en inglés) mediante la determinación simultanea de la presencia/ausencia de los cromosomas 6, 18 y 21 utilizando la tecnología iFISH.

WO2008/019872 se refiere a un método para el diagnóstico de leucemia linfoblástica aguda en niños mediante la determinación de los niveles de expresión de al menos 14 genes de los descritos en las tablas 1-4 o 5-16, en especial aquellos mostrados en la Figura 3 de dicha publicación: DNTT, SBGR101, DEFA3, CAMP, FCER2, DEFA4, BPI, PGLYRP1, LTF, SBGR2, NM_003526, CEACAM8, RNASE3, y ELA2.

Así pues, a pesar de los avances en el tratamiento de la LLA, existe una necesidad urgente de identificar biomarcadores específicos que permitan monitorizar la eficacia del tratamiento y el pronóstico de la enfermedad. La mejora de los métodos pronósticos es esencial para poder predecir de manera fiable el riesgo del paciente y así optimizar la estrategia de tratamiento permitiendo personalizar la terapia, reduciendo la toxicidad en pacientes de bajo riesgo y permitiendo terapias agresivas en pacientes de riesgo elevado.

La adopción de la terapia más adecuada para cada paciente de LLA gracias a su correcta clasificación en función del pronóstico de la enfermedad permitirá aumentar el porcentaje de éxito del tratamiento, consiguiendo un mayor porcentaje de supervivencia y evitando recaídas, presentando también como ventaja la reducción de los costes de tratamiento y hospitalización asociados.

BREVE DESCRIPCION DE LA INVENCION

Los inventores efectuaron un análisis bioinformático partiendo de datos de secuencias de RNAm de TCFL5 publicadas en las bases de datos de secuenciación de RNA “RNAseq” (NCBI, Ensemble, Vega y EC gene) en el que se analizaron las uniones de exones, y se establecieron los mensajeros consenso expresados en humano (ver Figura 1 y Tabla I del Ejemplo 1). En base a dichos resultados, los inventores establecieron que el gen *TCFL5* presenta a nivel de RNAm 6 isoformas mayoritarias distintas: (1) TCFL5R, (2) TCFL5R4b, (3) TCFL5R7, (4) TCFL5R4b6 (5) TCFL5R6 y (6) CHA (Tcfl5R2b o R2b), aumentando el número de isoformas respecto a las 2 isoformas previamente descritas: TCFL5 y CHA.

En el presente documento cuando se hace referencia a la isoforma TCFL5 se refiere a aquella que presenta E1, E2a, E3, E4 (E4a o E4b), E5 y E8; es decir, engloba a TCFL5R y TCFL5R4b.

La isoforma CHA fue descrita previamente por los inventores (*Rodriguez, C.I., N. Girones, and M. Fresno, Cha, a basic helix-loop-helix transcription factor involved in the regulation of upstream stimulatory factor activity. J Biol Chem, 2003. 278(44): p. 43135-45*) y se origina

debido a que CHA utiliza un promotor alternativo situado entre el exón 1 y 2, presentando como exón inicial E2b. A nivel RNAm CHA presenta E2b, E3, E4, E5 y E8. Así pues se diferencia de TCFL5 por la ausencia de E1 y por presentar E2b en lugar de E2a (la porción de la secuencia de E2a distinta de E2b se ha indicado en cursiva en la Figura 3).

Es importante considerar que el exón 2b contiene un 5'UTR y una región codificante coincidente con E2a por lo que a nivel proteico la secuencia correspondiente a E2 es idéntica en TCFL5 y CHA, y por tanto CHA se diferencia únicamente por la ausencia de la secuencia aminoacídica correspondiente a E1.

Asimismo, los inventores detectaron por primera vez que para un mismo tejido, línea celular o muestra de paciente, bajo las mismas condiciones experimentales, las isoformas TCFL5 y CHA (TCFL5R2b) presentan una expresión diferencial, en concreto se observaron distintos grados de expresión tanto a nivel de RNAm (Fig.2 y 7B) como proteico (Fig.7A).

La determinación precisa del riesgo de un paciente a sufrir una recaída constituye el paradigma fundamental en el tratamiento de la leucemia linfoblástica aguda (LLA). Dicha estratificación de los pacientes en función del riesgo permite adaptar la intensidad de la terapia al riesgo de recaída del paciente. Considerando la expresión diferencial de dichas isoformas, los inventores emitieron la hipótesis de que el uso de TCFL5 como marcador pronóstico y/o de diagnóstico diferencial en LLA dependerá de si se determinan los niveles de expresión de la isoforma TCFL5, de la isoforma CHA (TCFL5R2b) o de ambas.

En base a dicha hipótesis, procedieron al diseño de cebadores y sondas que detectan el RNAm de las isoformas TCFL5 y CHA respectivamente de manera específica e independiente, y ambas isoformas simultáneamente (TCFL5/CHA).

Los inventores establecieron una asociación entre niveles bajos de expresión de TCFL5/CHA y un pronóstico de LLA de alto riesgo (Fig.6C y Tabla VI del Ejemplo 5). En particular, observaron que la expresión de RNAm de TCFL5/CHA en el momento del diagnóstico es mucho menor en las LLA de alto riesgo respecto de las LLA de riesgo no alto (medio o bajo) (Tabla VII del Ejemplo 5). En la Fig.6A (Tabla IV del Ejemplo 5) se muestra que en el momento del diagnóstico la expresión de TCFL5/CHA en pacientes de LLA-T es menor que en pacientes de LLA-B, aunque la diferencia no es estadísticamente significativa.

Una asociación entre niveles bajos de expresión de TCFL5/CHA y un pronóstico de LLA de alto riesgo se observó también al analizar exclusivamente muestras de LLA-B (Tablas IX y X del Ejemplo 5) que es el grupo representado de manera mayoritaria en la cohorte analizada.

Asimismo, los inventores determinaron que existe una correlación entre los niveles de expresión de TCFL5/CHA en pacientes con LLA y la progresión de la enfermedad y/o eficacia del tratamiento. Más específicamente, observaron que la expresión de RNAm de TCFL5/CHA en LLA es significativamente menor en la recaída que en el diagnóstico (Fig.6B y Tabla V del Ejemplo 5), dicha asociación se observa especialmente en muestras de B-LLA (Tabla VIII del Ejemplo 5).

Por tanto, de acuerdo con dichos hallazgos la presente invención se refiere en un primer aspecto a un método para el pronóstico y/o diagnóstico diferencial de pacientes con leucemia linfoblástica aguda (LLA) que comprende las siguientes etapas:

a. determinar , preferiblemente de manera simultánea, los niveles de expresión de las isoformas TCFL5 y CHA del gen *TCFL5* en una muestra biológica aislada de dicho paciente; y
b. comparar los niveles de expresión de TCFL5/CHA en la muestra del paciente con valores de referencia,

donde una reducción en los valores de la muestra del paciente respecto a los valores de referencia es indicativa de LLA de alto riesgo.

El término “los niveles de expresión de TCFL5/CHA” tal y como se utiliza en la presente invención hace referencia a los niveles de expresión de TCFL5 y CHA de manera conjunta.

En un aspecto asociado, la presente invención se refiere al uso *in vitro* de los niveles de expresión de las isoformas TCFL5 y CHA del gen *TCFL5* determinados, preferiblemente de manera simultánea, en una muestra biológica aislada de un paciente con leucemia linfoblástica aguda (LLA) para el pronóstico y/o diagnóstico diferencial de pacientes con LLA, donde una reducción en los valores de TCFL5/CHA en la muestra del paciente respecto a valores de referencia es indicativa de LLA de alto riesgo.

En otro aspecto, la presente invención se refiere también a un método para la monitorización de la progresión de la enfermedad y/o eficacia del tratamiento en pacientes con LLA que comprende las siguientes etapas:

a. determinar, preferiblemente de manera simultánea, los niveles de expresión de las isoformas TCFL5 y CHA del gen *TCFL5* en una muestra biológica aislada de dicho paciente; y
b. comparar los niveles de expresión de TCFL5/CHA en la muestra del paciente con valores de referencia,

donde una reducción en los valores de la muestra del paciente respecto a los valores de referencia es indicativo de recaída.

En un aspecto relacionado, la presente invención se refiere al uso *in vitro* de los niveles de expresión de las isoformas TCFL5 y CHA del gen *TCFL5* determinados, preferiblemente de manera simultánea, en una muestra biológica aislada de un paciente con leucemia linfoblástica aguda (LLA) para la monitorización de la progresión de la enfermedad y/o eficacia del tratamiento en pacientes con LLA, donde una reducción en los valores de la muestra del paciente respecto a valores de referencia es indicativa de recaída.

La invención hace referencia también a un método para la determinación del tratamiento más adecuado para un paciente de LLA que comprende la clasificación de dicho paciente de acuerdo con el método para el pronóstico y/o diagnóstico diferencial del primer aspecto de la invención.

En un aspecto adicional, la presente invención se refiere a un método para el tratamiento de pacientes de LLA donde dicho tratamiento se determina en función de la clasificación de dicho paciente de acuerdo con el método para el pronóstico y/o diagnóstico diferencial del primer aspecto de la invención.

En otro aspecto, la invención hace referencia a un kit para el pronóstico, diagnóstico diferencial y/o monitorización de pacientes con leucemia linfoblástica aguda (LLA) que comprende:

a. un cebador y/o sonda que comprende o consiste en una secuencia seleccionada del grupo que consiste en SEQ ID NO: 5 a SEQ ID NO: 70, y secuencias idénticas a cualquiera de las mismas en al menos 75%; y

b. opcionalmente, instrucciones para el uso de dicho reactivo para la determinación de los niveles de expresión de dichas isoformas en una muestra biológica aislada de dicho paciente.

La invención se refiere también al uso de un kit para el pronóstico, diagnóstico diferencial y/o monitorización de pacientes con leucemia linfoblástica aguda (LLA) en un método de acuerdo con el primer aspecto de la invención, donde dicho kit comprende:

a. un reactivo para determinar, preferiblemente de manera simultánea, los niveles de expresión de las isoformas TCFL5 y CHA del gen *TCFL5*; y
b. opcionalmente, instrucciones para el uso de dicho reactivo para la determinación de los niveles de expresión de dichas isoformas en una muestra biológica aislada de dicho paciente.

En un aspecto adicional, la invención está relacionada con una secuencia nucleotídica seleccionada del grupo que consiste en SEQ ID NO: 5 a SEQ ID NO: 70, y secuencias idénticas a cualquiera de las mismas en al menos 75%.

BREVE DESCRIPCION DE LAS FIGURAS

Fig.1: Transcritos RNAm consenso mayoritarios del gen *TCFL5*. Se determinaron mediante análisis de las uniones de exones sobre las bases de datos ya publicadas de secuencias de RNAseq del gen *TCFL5*.

Fig.2: Transcritos de RNAm de *TCFL5* correspondientes al exón E1 vs E2 codificantes presentes en diferentes células/tejidos. Se determinaron mediante análisis de las bases de datos ya publicadas de RNAseq referentes al gen *TCFL5* y se representan por códigos de colores (más oscuro mayor frecuencia), respecto a los 2 exones. Se observa una expresión alta de E2 codificante en células y tejidos donde la expresión de E1 es baja, indicando una expresión diferencial de CHA respecto al resto de isoformas del gen *TCFL5 (*presentan E1).

Fig.3: Secuencia de TCFL5 RNAm (SEQ ID NO:1). Se indican las regiones 5'UTR y 3'UTR. Asimismo, se especifica la secuencia correspondiente a la región codificante en la que los distintos exones E1, E2, E3, E4, E5 y E8 se indican con alternancia de colores (gris y negro). En esta secuencia el exón 2 es E2a, la parte del exón 2a utilizado por TCFL5 pero

no por CHA es representada en cursiva y el ATG iniciador de CHA dentro del exón 2 se ha resaltado en negrita (la parte no traducida de este exón para CHA es distinta y representa el exón 2b, no mostrado). La región común para ambas isoformas se muestra sombreada en gris. Asimismo, se han enmarcado unas de las secuencias utilizadas para el diseño de cebadores donde el cebador forward se une a E5 y el reverse a E8.

Fig.4: Alineamiento de las secuencias codificantes de TCFL5 y CHA (RNAm).

Fig.5: Expresión de TCFL5 y CHA en células Jurkat: efectos de DAPT (inhibidor de Notch1) sobre los niveles de expresión de TCFL5/CHA a nivel proteína (Western Blot).

Fig.6: Expresión de TCFL5/CHA en muestras de LLA obtenidas de pacientes: Los niveles de expresión de TCFL5/CHA a nivel de RNAm se muestran de manera individualizada para cada uno de los grupos de estudio. En las gráficas se comparan los valores obtenidos en muestras (A) de B-ALL vs T-ALL, (B) recogidas en el momento de diagnóstico vs recaídas y (C) clasificadas en función del grado de riesgo bajo, medio o alto de acuerdo con los criterios especificados en materiales y métodos.

Fig.7: Expresión de TCFL5/CHA en muestras de leucemias ALL obtenidas de pacientes en el momento del diagnóstico (LLA012, LL014) y recaída (LLA751, LLA783). (A) Expresión de las proteínas CHA y TCFL5 por Western Blot; (B) Cuantificación del RNAm de TCFL5/CHA por qRT-PCR.

DESCRIPCION DETALLADA DE LA INVENCION

Definiciones

El término “seguimiento o monitorización de la progresión de la enfermedad” como se usa en el presente documento se refiere a determinar la evolución de la enfermedad, por ejemplo determinar si existe una recaída.

El término “eficacia de un tratamiento” como se usa en el presente documento se refiere al grado en el que un tratamiento logra el resultado deseado o previsto, por ejemplo, la capacidad de un fármaco para lograr el efecto deseado.

El término “tratamiento” engloba tanto un tratamiento profiláctico como terapéutico. El

término “tratamiento terapéutico” como se usa en el presente documento se refiere a aquel cuyo objetivo es pasar de un estado de enfermedad o patológico a un estado de salud. El término “tratamiento profiláctico” tal y como se usa en el presente documento se refiere a la prevención de un estado patológico.

El término "cantidad terapéuticamente efectiva" como se usa en el presente documento, se refiere a una cantidad que es efectiva, tras la administración de una dosis única o múltiple a un sujeto (por ejemplo, a un paciente humano) en el tratamiento profiláctico o terapéutico de una enfermedad, trastorno o afección patológica.

El término "sonda" como se usa en el presente documento se refiere a ácidos nucleicos producidos de forma sintética o biológica, de entre 10 y 285 nucleótidos de longitud que contienen secuencias nucleotídicas específicas que permiten una hibridación específica y preferente bajo condiciones predeterminadas a las secuencias de ácido nucleico diana, y opcionalmente se han modificado para la detección o para potenciar el rendimiento del ensayo. En general, es necesario un mínimo de diez nucleótidos para obtener estadísticamente la especificidad y para formar productos de hibridación estables, y un máximo de 285 nucleótidos representa en general un límite superior para la longitud en la que los parámetros de reacción se pueden ajustar fácilmente para determinar secuencias erróneamente emparejadas e hibridación preferente. Opcionalmente, las sondas pueden contener determinados elementos constitutivos que contribuyen a su funcionamiento correcto u óptimo en determinadas condiciones de ensayo. Por ejemplo, las sondas se pueden modificar para mejorar su resistencia a la degradación por nucleasas, para llevar a cabo la detección de ligandos (por ejemplo, marcaje con fluoresceína) o para facilitar su captura sobre un soporte sólido (por ejemplo, cola de polyA).

El término "cebadores" como se usa en el presente documento se refiere a oligonucleótidos o sondas que se pueden usar en un procedimiento de amplificación, tal como una reacción en cadena de la polimerasa ("PCR"), para amplificar una secuencia de nucleótidos. Los cebadores se diseñan en base a la secuencia polinucleotídica de una secuencia diana particular, por ejemplo, una secuencia de RNAm específica. El diseño y la validación de cebadores y sondas es bien conocido en la técnica. Para procedimientos de PCR en tiempo real cuantitativa, véase, por ejemplo, Rodriguez A *et al.* (Methods Mol Biol., 2015, 1275:31-56).

El término “específico" como se usa en el presente documento quiere decir que una

secuencia nucleotídica se hibridará a/amplificará una secuencia diana predeterminada y no se hibridará sustancialmente a/amplificará una secuencia no diana en las condiciones de ensayo, en general se usan condiciones restrictivas.

El término "hibridación" como se usa en el presente documento se refiere a un procedimiento por el que, en condiciones de reacción predeterminadas, dos hebras parcial, sustancial o completamente complementarias de ácido nucleico se deja que entren en contacto de forma antiparalela para formar un ácido nucleico bicatenario con enlaces de hidrógeno específicos y estables, siguiendo reglas explícitas que hacen que las bases de ácidos nucleicos se puedan emparejar entre sí.

El término "parcialmente complementaria" tal y como se usa en el presente documento se refiere a una secuencia nucleotídica que es al menos aproximadamente 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, o 99% complementaria a una secuencia nucleotídica de referencia.

El término "sustancialmente complementaria" tal y como se usa en el presente documento se refiere a una secuencia nucleotídica que es al menos aproximadamente 95%, 96%, 97%, 98%, o 99% complementaria a una secuencia nucleotídica de referencia.

El término "hibridación sustancial" quiere decir que la cantidad de hibridación observada será tal que alguien que observe los resultados considere el resultado positivo con respecto a los datos de hibridación en controles positivos y negativos. Los datos que se consideran "ruido de fondo" no son de hibridación sustancial.

El término "condiciones de hibridación restrictivas" quiere decir de aproximadamente 35 ºC a 65 ºC en una solución salina de NaCl aproximadamente 0,9 molar. La restricción también se puede regir por dichos parámetros de reacción como la concentración y el tipo de especies iónicas presentes en la solución de hibridación, los tipos y concentraciones de agentes desnaturalizantes presentes, y la temperatura de hibridación. En general, como las condiciones de hibridación se vuelven más restrictivas, son preferentes sondas más largas si se van a formar híbridos estables. Como norma, la restricción de las condiciones bajo las que va a tener lugar la hibridación determinará ciertas características de las sondas preferentes que se van a emplear.

El término "anticuerpo" incluye anticuerpos monoclonales y policlonales, así como anticuerpos recombinantes. El término "anticuerpo recombinante" tal y como se utiliza en la

presente invención se refiere a un anticuerpo producido o expresado utilizando un vector de expresión recombinante, donde el vector de expresión comprende un ácido nucleico que codifica el anticuerpo recombinante, tal que la introducción del vector de expresión en una célula huésped apropiada resulta en la producción o expresión del anticuerpo recombinante. Los anticuerpos recombinantes pueden ser anticuerpos quiméricos o humanizados, anticuerpos mono- o multi-específicos. El término "anticuerpo" también se refiere a fragmentos y derivados de todos los anteriores, y puede comprender además cualquier variante de los mismos que retienen la capacidad de unirse específicamente a un epítopo. Los anticuerpos pueden incluir, pero no se limitan a anticuerpos monoclonales (mAbs), anticuerpos de dominio único (sdAb), anticuerpos de cadena sencilla (scFv), fragmentos Fab, F (ab')2, fragmentos Fv (sdFv) disulfuros, anti-idiotipo (anticuerpos anti-Id), intra-cuerpos, anticuerpos sintéticos, y fragmentos de unión a epítopo de cualquiera de los anteriores. El término "anticuerpo" también se refiere a una proteína de fusión que incluye una región equivalente a la región Fc de una inmunoglobulina.

El término "kit" indica un conjunto de reactivos y coadyuvantes requeridos para un análisis. Aunque un kit consiste en la mayoría de los casos de varias unidades, también pueden estar disponibles los varios elementos de análisis presentados en una única unidad, que deben considerarse como kits.

Un método para el pronóstico y/o diagnóstico diferencial de pacientes con leucemia linfoblástica aguda (LLA)

En un primer aspecto la invención se refiere a un método para el pronóstico de pacientes con leucemia linfoblástica aguda (LLA) que comprende las siguientes etapas:

a. determinar, preferiblemente de manera simultánea, los niveles de expresión de las isoformas TCFL5 y CHA del gen *TCFL5* en una muestra biológica aislada de dicho paciente; y
b. comparar los niveles de expresión de TCFL5/CHA en la muestra del paciente con valores de referencia,

donde una reducción en los valores de la muestra del paciente respecto a los valores de referencia es indicativa de LLA de alto riesgo.

El método según el primer aspecto de la invención permite el diagnóstico diferencial, la clasificación o estratificación de los pacientes de LLA en base al pronóstico e.g., LLA de alto riesgo.

En un aspecto relacionado la invención se refiere a un método para el diagnóstico diferencial de pacientes con leucemia linfoblástica aguda (LLA) que comprende las siguientes etapas:

a. determinar, preferiblemente de manera simultánea, los niveles de expresión de las isoformas TCFL5 y CHA del gen *TCFL5* en una muestra biológica aislada de dicho paciente; y
b. comparar los niveles de expresión de TCFL5/CHA en la muestra del paciente con valores de referencia,

donde una reducción en los valores de la muestra del paciente respecto a los valores de referencia es indicativa de LLA de alto riesgo.

Donde la expresión “determinación simultánea” se refiere a la determinación conjunta y mediante una única reacción, por ejemplo mediante el uso de cebadores que amplifican la región común a ambas isoformas (RNAm) y/o el uso de reactivos de afinidad que se unen específicamente a los polipéptidos codificados por las mismas.

El gen *TCFL5* (NCBI Gene ID: 10732) se ha descrito en humanos y codifica para un factor de transcripción de tipo basic hélix-loop-helix. Está localizado en el cromosoma 20 (20q13.33) y presenta 8 exones(E1-E8) que sufren “splicing” alternativos dando lugar a 6 formas mayoritarias según análisis bionformaticos llevados a cabo por los inventores de las bases de datos de secuenciación de RNA “RNAseq” (NCBI, Ensemble, Vega y EC gene). Dichas isoformas del gen *TCFL5* son: (1) TCFL5R, (2) TCFL5R4b, (3) TCFL5R7, (4) TCFL5R4b6 (5) TCFL5R6 y (6) CHA (Tcfl5R2b o R2b), aumentando el número de isoformas respecto a las 2 isoformas previamente descritas: Tcfl5R y CHA (Tcfl5R2b). Ver Figura 1 y Tabla I del Ejemplo 1.

La secuencia canónica de la variante de transcripción o isoforma TCFL5 (RNAm) es TCFL5R y se identifica en la presente invención como SEQ ID NO:1 (NCBI: NM_006602).

La secuencia de la variante de transcripción o isoforma CHA (RNAm) se identifica en la presente invención como SEQ ID NO:2 (NCBI: AJ271337.1).

Los polipéptidos codificados por las isoformas TCFL5 y CHA (RNAm) corresponden a las secuencias de aminoácidos SEQ ID NO:3 y SEQ ID NO:4, respectivamente.

La determinación de las isoformas TCFL5 y CHA de acuerdo con el método de la invención se realiza preferiblemente con reactivos que detectan ambas isoformas del gen *TCFL5* (a nivel de RNAm o proteico) de manera simultánea, más preferiblemente la detección y/o cuantificación de dichas isoformas es específica para TCFL5 y CHA, no detectándose y/o cuantificándose otras isoformas del gen *TCFL5*.

En un modo de realización particular, los niveles de expresión de las isoformas TCFL5 y CHA se determinan a nivel de RNA mensajero (RNAm).

Los procedimientos de biología molecular para cuantificar las secuencias de ácidos nucleicos diana son bien conocidos en la técnica. Estos procedimientos incluyen, pero no se limitan a PCR de punto final, PCR competitiva, transcriptasa inversa asociada a PCR (RT-PCR), PCR cuantitativa (qPCR), transcriptasa inversa asociada a qPCR (RT-qPCR), PCR-pirosecuenciación, PCR-ELISA, micromatrices de ADN, espectrometría de masas, ensayos de hibridación *in situ* tales como inmunotransferencia por puntos (dot-blot) o ensayo de hibridación *in situ* con fluorescencia (FISH), ADN ramificado (bDNA; Nolte, Adv. Clin. Chem. 1998,33:201-235) y versiones múltiplex de dichos procedimientos (véase, por ejemplo, Andoh *et al.*, Current Pharmaceutical Design, 2009;15,2066-2073), así como la próxima generación de cualquiera de las técnicas enumeradas y combinaciones de las mismas, todos los cuales están dentro del alcance de la presente invención. Dichos procedimientos pueden incluir también la pre-conversión del RNAm en cDNA a través de la reacción con una transcriptasa inversa (RT), por ejemplo la reacción de PCR es habitualmente precedida de la conversión del RNAm en cDNA y se refiere como RT-PCR.

Los cebadores y/o sondas, en general, reaccionan ofreciendo una respuesta directa y lineal a cantidades crecientes de las secuencias de ácidos nucleicos dianas. Gracias a la comparación con estándares apropiados, se puede cuantificar fácilmente la cantidad de una secuencia de ácidos nucleicos dados en una muestra. Preferentemente, dicho procedimiento molecular para cuantificación génica se selecciona del grupo que consiste en reacción en cadena de la polimerasa cuantitativa (qPCR), PCR-pirosecuenciación, hibridación *in situ* con fluorescencia (FISH), micromatrices de ADN y PCR-ELISA.

Un procedimiento de cuantificación preferido es FISH, que combina la hibridación de sondas con microscopía óptica fluorescente, microscopía láser confocal o citometría de flujo para la

cuantificación directa de secuencias diana individuales.

Otro procedimiento de cuantificación preferido es la PCR cuantitativa (qPCR) asociada a la transcriptasa inversa. La qPCR o PCR en tiempo real es bien conocidad por un experto en la materia. Se comercializan distintos instrumentos para llevar a cabo dicha reacción, tales como ABI Prism 7700 SDS, GeneAmp 5700 SDS, ABI Prism 7900 HT SDS de Applied Biosystems; iCycler iQ de Bio-Rad; Smart Cycler de Cepheid; Rotor-Gene de Corbett Research; LightCycler de Roche Molecular Biochemicals y Mx4000 Multiplex de Stratagene. El procedimiento de qPCR permite la cuantificación exacta del producto de PCR en tiempo real midiendo la acumulación del producto de PCR muy pronto en la fase exponencial de la reacción, reduciendo así el sesgo en la cuantificación ligado a la eficacia de la amplificación de PCR que se produce en la PCR de punto final. La PCR en tiempo real es bien conocida en la técnica y por tanto, no se describe en detalle en el presente documento. Una visión general de la tecnología y los protocolos para qPCR están disponibles, por ejemplo, de los proveedores mencionados anteriormente, por ejemplo, http://www.sigmaaldrich.com/technical-documents/protocols/biology/sybr-green-qpcr.html o http://www.sigmaaldrich.com/life-science/molecular-biology/pcr/quantitative-pcr/qpcr-technical-guide.html. Una revisión del uso de qPCR en la cuantificación de mRNA se encuentra por ejemplo en Wong ML y Medrano JF, Biotechniques 2005, 39(1):75-85.

Están disponibles diferentes bioquímicas de detección para qPCR. Se pueden usar todas ellas con los instrumentos de qPCR mencionados anteriormente. El término “bioquímica de detección” se refiere a un procedimiento para informar de la amplificación del producto de PCR específico en PCR en tiempo real. Dichas bioquímicas de detección se clasifican en dos grupos principales. El primer grupo comprende moléculas de intercalado de ADN de doble cadena: tales como SYBR Green I y EvaGreen; y el segundo grupo incluye oligonucleótidos marcados típicamente con un fluoróforo. Este último, a su vez, se ha dividido en tres subgrupos: (i) cebadores-sondas (Scorpions, Amplifluor®, LUX™, Cyclicons, Angler®); (ii) sondas de hidrólisis (TaqMan, MGB-TaqMan, Snake assay) y de hibridación (Hybprobe or FRET, Molecular Beacons, HyBeacon™, MGB-Pleiades, MGB-Eclipse, ResonSense®, Yin-Yang or displacing); y (iii) análogos de ácidos nucleicos (PNA, LNA®, ZNA ™, bases no naturales: Plexor™ primer, Tiny-Molecular Beacon), ver E. Navarro eta l. ,Clinica Chimica Acta, Volume 439, 15 January 2015, Pages 231–250.

En un modo de realización preferente, dichas sondas son oligonucleótidos de doble marcaje, tales como sondas de hidrólisis o balizas moleculares. El extremo 5’ del oligonucleótido,

típicamente, se marca con una molécula indicadora (reporter) fluorescente tales como FAM, TET o JOE mientras que el extremo 3' se marca con una molécula desactivadora (quencher), tales como TAM o BHQ1. La secuencia de la sonda es específica para una región de interés en la molécula diana amplificada. En un modo de realización más preferente, dicha sonda es una sonda de hidrólisis que está diseñada de modo que la longitud de la secuencia sitúa el fluoróforo 5' y la molécula desactivadora 3' en proximidad suficientemente estrecha para suprimir la fluorescencia. Varias moléculas indicadoras y extintoras para su uso en sondas de qPCR son bien conocidas en la técnica. Estas están disponibles, por ejemplo, de https://www.eurofinsgenomics.eu/en/dna-rna-oligonucleotides/optimised-application-oligos/qpcr-probes.aspx.

Generalmente, para la cuantificación de secuencias de nucleótidos se utilizan sondas y/o cebadores. El término "un cebador y/o sonda" incluye específicamente "cebadores y/o sondas", englobando por ejemplo a un cebador, una sonda, un cebador y una sonda, una pareja de cebadores, y una pareja de cebadores y una sonda. Ambos términos se utilizan de manera indistinta en la presente invención.

Las sondas y/o cebadores utilizadas en el método de la invención hibridan específicamente con SEQ ID NO: 1 y/o SEQ ID NO: 2. Dichas sondas y/o cebadores son parcialmente complementarias, preferiblemente sustancial o completamente complementarias a SEQ ID NO: 1 y/o SEQ ID NO: 2 o a fragmentos de las mismas tal y como se describe en el presente documento.

Preferiblemente, una sonda y/o cebador es un secuencia de polinucleótidos de entre 10 y 30 nucleótidos, más preferiblemente de entre 15 y 26 nucleótidos, aún más preferiblemente de entre 18 y 22 nucleótidos, y aún mucho más preferiblemente de alrededor de 20 nucleótidos. En una realización particular, dichos cebadores y/o sondas han sido modificados para la detección o para potenciar el rendimiento del ensayo.

En una realización particular, dichos cebadores y/o sondas comprenden o consisten en una secuencia seleccionada del grupo que consiste en SEQ ID NO: 5 a SEQ ID NO: 70 (ver Tabla III), y secuencias idénticas a cualquiera de las mismas en al menos 75%.

Preferiblemente, dichos cebadores y/o sondas comprenden o consisten en una secuencia seleccionada del grupo que consiste en SEQ ID NO: 5 a SEQ ID NO: 70 (ver Tabla III).

En otra realización particular, dichos cebadores y/o sondas son específicos para la cuantificación/amplificación de una región comprendida entre los exones 3 y 8 de TCFL5R (SEQ ID NO: 1) permitiendo cuantificar de manera específica y simultánea la expresión de TCFL5 (independientemente de que el exón 4 presente la variante E4a o E4b) y CHA (TCFL5R2b).

En dicho modo de realización particular, la determinación de los niveles de expresión de las isoformas TCFL5/CHA a nivel RNAm se efectúa mediante la cuantificación de una secuencia comprendida entre el primer nucleótido de E3 y el último nucleótido de E8.

En una realización preferida, dichos cebadores y/o sondas hibridan específicamente con secuencias comprendidas en cualquiera de los exones E3, E4, E5 y/o E8, incluyendo cuando ambos cebadores hibridan en un mismo exón y cualquier combinación de los mismos. Preferiblemente, dichos cebadores y/o sondas comprenden o consisten en una secuencia seleccionada del grupo que consiste en SEQ ID NO: 5 a SEQ ID NO: 42; y SEQ ID NO: 45 a SEQ ID NO: 62 (ver Tabla III), y secuencias idénticas a cualquiera de las mismas en al menos 75%. Más preferiblemente, dichos cebadores y/o sondas comprenden o consisten en una pareja de cebadores seleccionada del grupo que consiste en SEQ ID NO: 5 a SEQ ID NO: 42; y SEQ ID NO: 45 a SEQ ID NO: 62 (ver Tabla III).

En otra realización preferida, dichos cebadores y/o sondas hibridan específicamente con secuencias comprendidas en cualquiera de los exones E3, E4 y/o E5 (amplificando una región común a todas las isoformas), incluyendo cuando ambos cebadores hibridan en un mismo exón y cualquier combinación de los mismos. Preferiblemente, dichos cebadores y/o sondas comprenden o consisten en una secuencia seleccionada del grupo que consiste en SEQ ID NO: 5 a SEQ ID NO: 42 (ver Tabla III), y secuencias idénticas a cualquiera de las mismas en al menos 75%. Más preferiblemente, dichos cebadores y/o sondas comprenden o consisten en una pareja de cebadores seleccionada del grupo que consiste en SEQ ID NO: 5 a SEQ ID NO: 42 (ver Tabla III).

En otra realización preferida, dichos cebadores y/o sondas amplifican la región común y específica de las isoformas TCFL5R y CHA, más específicamente un cebador y/o sonda hibrida en E3 o E5 y el otro en E8, o ambos en E8.

En una realización más preferida, uno de los cebadores y/o sondas hibrida específicamente con una secuencia comprendida en E5 y el otro en E8. Preferiblemente, dichos cebadores y/o sondas comprenden o consisten en una secuencia seleccionada del grupo que consiste en SEQ ID NO: 45 a SEQ ID NO: 54 (ver Tabla III), y secuencias idénticas a cualquiera de las mismas en al menos 75%. Más preferiblemente, dichos cebadores y/o sondas comprenden o consisten en una pareja de cebadores seleccionada del grupo que consiste en SEQ ID NO: 45 a SEQ ID NO: 54 (ver Tabla III).

En otra realización más preferida, uno de los cebadores y/o sondas hibrida específicamente con una secuencia comprendida en E3 y el otro en E8. Preferiblemente, dichos cebadores y/o sondas comprenden o consisten en una secuencia seleccionada del grupo que consiste en SEQ ID NO: 55 a SEQ ID NO: 56 (ver Tabla III), y secuencias idénticas a cualquiera de las mismas en al menos 75%. Más preferiblemente, dichos cebadores y/o sondas comprenden o consisten en una pareja de cebadores seleccionada del grupo que consiste en SEQ ID NO: 55 a SEQ ID NO: 56 (ver Tabla III).

En otra realización más preferida adicional, ambos cebadores y/o sondas hibridan en E8. Preferiblemente, dichos cebadores y/o sondas comprenden o consisten en una secuencia seleccionada del grupo que consiste en SEQ ID NO: 57 a SEQ ID NO: 62 (ver Tabla III), y secuencias idénticas a cualquiera de las mismas en al menos 75%. Más preferiblemente, dichos cebadores y/o sondas comprenden o consisten en una pareja de cebadores seleccionada del grupo que consiste en SEQ ID NO: 57 a SEQ ID NO: 62 (ver Tabla III).

Más preferiblemente, dichos cebadores y/o sondas comprenden o consisten en una secuencia seleccionada del grupo que consiste en SEQ ID NO: 45 (Primer TCFL5/CHA 01 Forward): GAGACTGACAAGGCCACAACT, SEQ ID NO: 46 (Primer TCFL5/CHA 01 Reverse):

CCGCAAAATACGCTCTCAA, y secuencias idénticas a cualquiera de las mismas en al menos 75%.

Las secuencias oligonucleotídicas con una identidad de al menos un 75 % mencionadas en la presente invención presentan preferiblemente una identidad de al menos un 80 %, al menos un 85 %, al menos un 90 %, al menos un 95 %, más preferentemente, 96 %, 97 %, 98 %, 99 % o un 100 % con las respectivas secuencias de referencia. Además, estas secuencias con una identidad de al menos un 75 % pueden tener el mismo número de nucleótidos, o bien presentar más o menos nucleótidos que la secuencia de referencia.

Los niveles de cuantificación pueden ser absolutos o relativos. En general, es preferente que los niveles de expresión se normalicen. La normalización se puede realizar con respecto a diferentes medidas en la muestra, tal como por peso de la muestra, cuantificación de células humanas, cuantificación de ADN total, y/o cuantificación de los niveles de expresión de un gen de expresión constitutiva. Estos procedimientos son bien conocidos para un experto en la técnica.

En un modo de realización particular, la normalización se lleva a cabo con respecto a la cuantificación de los niveles de expresión de un gen de expresión constitutiva. En la presente invención se entiende por "genes que se expresan de forma constitutiva" o "genes de expresión constitutiva", a aquellos genes que se ha descrito que se transcriben de manera constante. Ejemplos de genes que se expresan de forma constitutiva son 2-mioglobulina, ubiquitina, proteína ribosomal 18S, ciclofilina A, GAPDH, proteína de activación de la tirosina 3-monooxigenasa/triptófano 5-monooxigenasa (YWHAZ), beta-actina, β-2- microglobulina o hipoxantina-guanina fosforibosiltransferasa (HPRT).

En un modo de realización preferente, la cuantificación de los niveles de expresión de TCFL5/CHA se realiza por RT-qPCR y comprende la normalización de los niveles de expresión respecto a un gen de expresión constitutiva.

La detección y/o cuantificación de las isoformas TCFL5/CHA se puede llevar a cabo también a nivel proteico. Los polipéptidos codificados por las isoformas TCFL5 y CHA (RNAm)

corresponden a las secuencias de aminoácidos referidas como SEQ ID NO:3 y SEQ ID NO:4 en la presente invención. Existen diversos métodos para la cuantificación de péptidos y proteínas bien conocidos por un experto en la materia, tales como inmunoensayos. Diversos tipos de inmunoensayos son conocidos por un experto en la materia para la cuantificación de manera específica de proteínas de interés, ya sea en solución o utilizando un ensayo en fase sólida. Dichos métodos están basados en el uso de reactivos de afinidad, que pueden ser receptores o ligandos específicos, por ejemplo anticuerpos, preferiblemente marcados. Por ejemplo, el Western Blot o inmunotransferencia permite comparar las abundancias de proteínas separadas mediante un gel electroforético, eg. SDS-PAGE. En esta técnica, las proteínas separadas por electroforesis en gel se transfieren sobre una lámina de material polimérico (generalmente de nitrocelulosa, nylon, o difluoruro de polivinilideno), donde se inmovilizan. Las proteínas diana se revelan mediante el uso de una solución que contiene un anticuerpo específico. El anticuerpo puede conjugarse directamente con un marcador radiactivo, fluorescente o enzimático (método de detección directa) o bien se puede usar un anticuerpo secundario que reconoce el anticuerpo primario y por lo tanto amplifica la señal (método de detección indirecta o ensayo de tipo sándwich).

Tradicionalmente, la cuantificación específica de proteínas en solución se ha efectuado mediante inmunoensayos en un soporte sólido. Típicamente, se inmoviliza un anticuerpo de captura específico para la proteína diana en una superficie polimérica o de plástico y se añade al soporte una solución que contiene la proteína de interés (por ejemplo, suero o lisado celular). Finalmente, se incuba la muestra en el soporte durante un tiempo para permitir que se formen los complejos antígeno-anticuerpo. A continuación, se suelen realizar uno o más lavados para eliminar la solución y la proteína diana se detecta con un segundo anticuerpo que reconoce un epítopo de proteína diferente al reconocido por el anticuerpo de captura. Al igual que en el caso del Western Blot, este anticuerpo de detección puede estar marcado directamente o puede ser reconocido con un anticuerpo secundario. Un inmunoensayo comúnmente utilizado para la cuantificación de proteínas es el ensayo de inmunoabsorción ligado a enzimas (ELISA) en el que el anticuerpo de detección lleva una enzima que convierte un sustrato comúnmente incoloro en un compuesto coloreado o un sustrato no fluorescente en un compuesto fluorescente. Asimismo, en otros inmunoensayos en fase sólida, el anticuerpo puede estar marcado con un isotopo radioactivo o con fluorescencia.

Otros métodos que pueden ser utilizados para la cuantificación de las isoformas TCFL5 y CHA de acuerdo con la invención son técnicas basadas en espectrometría de masas (MS)

tales como la cromatografía liquida acoplada a la espectrometría de masas (LC/MS), descrita por ejemplo en US2010/0173786, o el tandem LC-MS/MS (WO2012/155019, US2011/0039287, Rauh M., J Chromatogr B Analyt Technol Biomed Life Sci. 2012 Feb 1;883-884:59-67), así como el uso de arrays de péptidos, proteínas o anticuerpos y versiones múltiplex de las técnicas mencionadas, así como la próxima generación de dichas técnicas y combinaciones de las mismas.

En un modo de realización particular de la invención, opcionalmente en combinación con una o más de las características descritas anteriormente, la cuantificación de TCFL5/CHA se realiza mediante un inmunoensayo que comprende el uso de cualquier anticuerpo que detecte ambas isoformas, tales como HPA055223 (Sigma) o SAB4500152 (Sigma), o bien anticuerpos específicos de cada una de ellas.

El grado de variación en los valores de expresión de TCFL5/CHA en la muestra del paciente respecto a los valores de referencia puede ser, por ejemplo, de al menos el 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 75%, 90%, 100%, 110%, 120%, 130%, 140%, 150% o más. Preferiblemente, dicha variación es estadísticamente significativa. Como se usa en este documento, "estadísticamente significativa" se refiere a un valor de p de menos de 0,05, por ejemplo, un valor de p de menos de 0,025, un valor de p de menos de 0,01 o un valor de p de menos de 0,005, utilizando una prueba estadística apropiada. Un experto en la técnica sabrá definir aquellas pruebas estadísticas más apropiadas. Preferentemente, en aquellos casos en los que haya una distribución normal y homocedasticidad, se usa un modelo paramétrico tal como la prueba de la t de student o la prueba de ANOVA; y cuando no se logre al menos uno de estos dos requisitos, entonces se usa, en general, un modelo no paramétrico tal como la prueba de la U de Mann-Whitney o la prueba de Kruskal-Wallis.

Las muestras biológicas para el uso en los métodos de la invención pueden obtenerse a partir de una variedad de tejidos o fluidos biológicos, en particular de sangre, pero también pueden ser utilizadas muestras de médula ósea, linfa, líquido cefalorraquídeo, líquido sinovial, y similares. Tales muestras se pueden separar por centrifugación, decantación, separación en gradiente de densidad, aféresis, selección por afinidad, FACS, etc. antes del análisis. Las muestras biológicas utilizadas en el método de la presente invención son preferiblemente muestras de médula ósea, de sangre periférica o de líquido cefalorraquídeo. Dichos tipos de muestra se utilizan de manera habitual en la práctica clínica y un experto en la materia sabrá identificar los medios adecuados para su obtención y conservación (Coustan-Smith E et al., Blood. 2002 Oct 1;100(7):2399-402; Martinez-Laperche C et al.,

American journal of hematology 88: 359-64). Una vez obtenida la muestra biológica, se puede utilizar directamente, congelar, o mantener en un medio de cultivo apropiado. Varios medios de cultivo pueden ser empleados para mantener las células en cultivo. Las muestras se pueden obtener por cualquier procedimiento adecuado, tal como la extracción de sangre, punción venosa, biopsia, o similares.

Las células mononucleares (MC) son generalmente aisladas de dicha muestra biológica, mediante métodos conocidos en el estado del arte. Las células mononucleares se aíslan típicamente mediante centrifugación por gradiente de densidad, por ejemplo con Ficoll® o Percoll®.

En una realización particular, la determinación de los niveles de expresión de las isoformas TCFL5 y CHA se efectúa en una población celular que se ha aislado de una muestra biológica del paciente mediante un procedimiento que comprende una centrifugación por gradiente de densidad.

Tras el aislamiento de MCs de las muestras biológicas, una selección específica del tipo celular de interés puede llevarse a cabo usando uno de los muchos métodos descritos en la literatura, sobre la base de la expresión de marcadores de superficie celular específicos y, si procede, de otras proteínas, así como la actividad de proliferación, el metabolismo y / o el estado morfológico de las células. Típicamente, la purificación de una población celular de interés en una muestra biológica comprende una selección positiva y/o negativa en base a la expresión de marcadores de superficie celular característicos.

En otra realización particular, la determinación de los niveles de expresión de las isoformas TCFL5 y CHA se efectúa en una población de células de LLA que se ha aislado de una muestra biológica del paciente mediante un procedimiento que comprende además la selección positiva mediante el uso de marcadores de superficie celular asociados a leucemia.

Típicamente se utilizan combinaciones de marcadores celulares que se encuentran en LLA y no en líneas celulares normales que se encuentren habitualmente también en dicha muestra biológica. Por ejemplo, las células LLA-B pueden ser seleccionadas por eliminación de otras MCs tales como los linfocitos T (células CD2+, o CD3+), y/o por la presencia de marcadores de superficie específicos, como por ejemplo CD19 (específico de linfocitos B), y/o CD10 (antígeno asociado a la leucemia linfoblástica). En una realización particular, las células de LLA-B se aíslan mediante una selección que comprende la expresión de CD19, y/o CD10.

La población de células de LLA se puede seleccionar sobre la base de la expresión de al menos un marcador de superficie celular. La selección normalmente se lleva a cabo utilizando proteínas que se unen específicamente a una de dichas proteínas de superficie celular, típicamente anticuerpos, y que se puede vincular a soportes sólidos (por ejemplo, partículas o superficies de plástico) o pueden ser conjugadas con moléculas de marcaje (por ejemplo, un fluorocromo) que pueden ser detectadas por ejemplo mediante citometría de flujo.

Preferiblemente, la muestra biológica del paciente y la muestra biológica de referencia son del mismo tipo, es decir, tienen un mismo origen biológico y han sido aisladas utilizando los mismos procedimientos. Dichas muestras biológicas pueden ser tomadas alrededor del momento del diagnóstico, antes, durante o después del tratamiento, preferiblemente son tomadas alrededor del momento del diagnóstico.

Los valores de referencia pueden ser los niveles de expresión determinados para el producto de expresión en una o varias muestras de referencia (por ejemplo, valores medios +/- s.m.e.) o bien valores predeterminados. Típicamente dicho valor de referencia se refiere como valor umbral o de corte (cut-off).

Una variedad de métodos estadísticos y matemáticos para establecer el valor umbral o de corte son conocidos en el estado de la técnica. Un valor de expresión umbral o de corte para un biomarcador particular puede ser seleccionado, por ejemplo, utilizando un análisis de Receiver Operating Characteristic (ROC). Un experto en la técnica apreciará que dicho valor

de corte se puede variar, por ejemplo, moviéndolo a lo largo del gráfico ROC, para obtener diferentes valores de sensibilidad o especificidad y por tanto afectando el rendimiento general del ensayo. El mejor punto de corte ("best cut-off") se refiere al valor obtenido del gráfico ROC para un biomarcador que produce la mejor sensibilidad y especificidad. Los valores de sensibilidad y especificidad se calculan sobre el rango de umbrales (cut-offs). Así pues, los valores de umbral o de corte pueden ser seleccionados de manera que los valores de sensibilidad y / o especificidad son al menos aproximadamente 70%, y pueden ser, por ejemplo, al menos 75%, al menos 80%, al menos 85%, al menos 90 %, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99% o al menos 100% en al menos 60% de la población de estudio, o en al menos 65%, 70 %, 75% o 80% de la población de estudio.

En una realización particular, las muestras de referencia o control para la obtención de los valores de referencia corresponden a una o más muestras de linfocitos aislados de donantes sanos (por ejemplo, de sangre periférica o de médula ósea) o bien a líneas establecidas de linfocitos B (Raji, ATCC® CCL-86™) o T humanos (Jurkat, Clone E6-1 (ATCC® TIB-152™).

En otra realización particular, las muestras de referencia para la obtención de los valores de referencia o control corresponden a una o más muestras celulares de pacientes con LLA o a líneas celulares establecidas de LLA. Típicamente, las líneas celulares y/o muestras celulares de donantes (sanos o pacientes) utilizadas han sido previamente caracterizadas, por ejemplo, de acuerdo a su fenotipo (e.g., LLA-B o LLA-T), inmunofenotipo (un panel standard incluye: CD10, CD19, CD20, CD34, CD38 y CD45), marcadores genéticos de LLA-B (e.g. PAX5, IKZF1, o EBF1) y citogenéticos (e.g., presencia de translocaciones MLL-AF4, ETV6-RUNX1 o BCR-ABL1. TEL-AML-1), o LLA-T (e.g. HOX11L2, LYL1 más LMO2, TAL1 más LMO1 o LMO2, HOX11, y MLL-EN, NOTCH1, c-MYC, etc) características demográficas y/o clínicas del donante (e.g., edad, sexo, origen étnico, contaje de glóbulos blancos (WBC) etc.), y/o respuesta al tratamiento. Líneas celulares características de LLA-B incluyen pero no se limitan a CCL-120 (ATCC). Asimismo, líneas celulares características de LLA-T incluyen pero no se limitan a CCL-119, CCL-120.1, CRL-1552, CRL-2264, CRL-2265, PTS-CCL-119, CRM-CCL-119, CRM-CCL-119D, CRL-11386, TIB-195 (todas de ATCC).

Preferiblemente, dichas muestras de pacientes han sido asociadas a pacientes de riesgo medio y/o bajo tanto LLA-B, como LLA-T .(*Zhou, Y., You, M. J., Young, K. H., Lin, P., Lu, G., Medeiros, L. J., and Bueso-Ramos, C. E. (2012) Advances in the molecular pathobiology of B-lymphoblastic leukemia. Human pathology 43, 1347-136 Van Vlierberghe, P., and Ferrando, A. (2012) The molecular basis of T cell acute lymphoblastic leukemia. J Clin Invest 122, 3398-3406*)

En un modo de realización particular del método de la invención se utiliza para el pronóstico y/o diagnóstico diferencial de pacientes de LLA-B o LLA-T. Preferiblemente, en pacientes de LLA-B. Independientemente del tipo de LLA diagnosticado, en un modo de realización preferido, dichos pacientes son niños. La población infantil en LLA se define típicamente como aquellos pacientes que tienen menos de 20 años. En una realización, preferida, dichos pacientes tienen entre 0 y 15 años, e incluso entre 0-12 meses. En otra realización preferida, dichos pacientes tienen entre 1 y 19 años.

En el tratamiento de pacientes con LLA, es práctica habitual el determinar el tratamiento más adecuado en función del diagnóstico, estratificación o clasificación del paciente como perteneciente a un grupo de riesgo, es decir se escoge o personaliza el tratamiento según el grupo de riesgo al que pertenecen, siendo menos intenso en los de riesgo medio o bajo y mayor en los de riesgo alto.

La clasificación de la enfermedad puede ser, por ejemplo, una clasificación de preferencia basada en el riesgo de recaída (remisión vs fracaso terapéutico); pero que también puede basarse en las características clínicas de los pacientes, en los datos citogenéticos; en el subtipo de leucemia; la respuesta al tratamiento y/o en la etiología de la enfermedad.

Típicamente, se definen como pacientes de alto riesgo aquellos que presentan mayores probabilidades de recaída. Dichos criterios para la clasificación y/o estratificación de pacientes son bien conocidos por un experto en la materia y se describen por ejemplo en *Ceppi F, et al., (2015) Risk factors for relapse in childhood acute lymphoblastic leukemia: prediction and prevention. Expert review of hematology 8: 57-70; Hunger SP, Mullighan CG (2015) Acute Lymphoblastic Leukemia in Children. The New England journal of medicine*

373: 1541-52; Teachey DT, Hunger SP (2013) Predicting relapse risk in childhood acute lymphoblastic leukaemia. British journal of haematology 162: 606-20):

En general, se distinguen 3 grupos de riesgo en función de una evaluación de los datos relativos a la edad del paciente y a los resultados obtenidos en una o más de las siguientes pruebas genéticas y/o bioquímicas:

Técnicas de laboratorio generalmente usadas en el diagnóstico y/o recaída:

- Hematimetría estándar con hemocitómetro.
- Citología convencional en extensión de muestra de médula ósea y en citospin de muestra de líquido cefalorraquídeo.
- Citometría de flujo multiparamétrica, preferiblemente en muestra de médula ósea.
- Técnicas citogenéticas clásicas y de bandeo de cromosomas, preferiblemente en muestra de médula ósea.
- RT-PCR con cebadores específicos para las translocaciones t(12;21), t(1;19), t(9;22) y t(4;11), preferiblemente en muestras de médula ósea.
- FISH con sondas para las translocaciones t(12;21), t(1;19), t(9;22) y break apart en MLL, preferiblemente en muestras de médula ósea.

En base a los resultados obtenidos en dichas pruebas, los pacientes de LLA se suelen clasificar en uno de los siguientes grupos de riesgo:

RIESGO ESTÁNDAR o RIESGO BAJO: El paciente debe reunir todos y cada uno de los siguientes criterios:

- Edad >1 y <10 años.
- Leucocitos <20 $x10^9$/l al diagnóstico.
- Inmunofenotipo no T.
- Ausencia de infiltración del SNC y/o testes.

- Citogenética (uno de los dos criterios siguientes es suficiente):
    - Alta Hiperdiploidía (51-67 cromosomas), índice de DNA 1,10-1,44 (siempre confirmado por otras técnicas citogenéticas).
    - t(12;21) positiva
- Ausencia de t(1;19)
- No reordenamiento MLL
- Presencia de < 1.000 blastos/mm3 en día +8 de la Inducción, en sangre periférica
- Presencia de < 5% de blastos y < 0,1% de enfermedad residual mínima (ERM) en médula ósea en día +15 de la Inducción y al final de la inducción IA

RIESGO ALTO: La existencia de cualquiera de los siguientes criterios determina la inclusión del paciente en este grupo:

- t(4;11) (MLL/AF4).
- t(9;22) p190.
- Hipodiploidía <44 cromosomas o índice DNA <0,81 (se requiere confirmación por otras técnicas).
- > 1.000 blastos en día +8 de la Inducción, en sangre periférica.
- > 25% de blastos y >10% de ERM en el día +15 de la Inducción, en médula ósea.
- ERM > 1% en el día +33 de la Inducción, en médula ósea.
- ERM > 0,1% antes de la Consolidación, en médula ósea.

RIESGO MEDIO: Aquellos pacientes que no reúnan los criterios de Riesgo Estándar ni de Riesgo Alto.

Dichos criterios son descritos en mayor detalle en la Guía de Recomendación Terapéutica SEHOP/PETHEMA 2013.

En un modo de realización particular, opcionalmente en combinación con una o más de las características definidas descritas, dicho método comprende además la detección o cuantificación de uno o más marcadores genéticos descritos con valor pronóstico para LLA, preferiblemente comprende la determinación de la presencia de una o más de las translocaciones seleccionadas del grupo que consiste en t(12;21), t(1;19), t(9;22) y t(4;11), Hipodiploidía <44 cromosomas o índice DNA <0,81.

Dicho método puede comprender además la determinación de uno o más de los parámetros bioquímicos y/o clínicos descritos más arriba habitualmente utilizados para la estratificación de los pacientes de LLA en función del riesgo.

El método de la invención puede comprender además el almacenaje de los resultados obtenidos en dispositivo de almacenamiento de datos. En un modo de realización, dicho dispositivo de almacenamiento de datos es una lámina de papel. En un modo de realización preferente, dicho dispositivo de almacenamiento de datos es un medio legible por ordenador. Como se usa en el presente documento, “un medio legible por ordenador” puede ser cualquier aparato que pueda incluir, almacenar, comunicar, propagar o transportar los resultados de la determinación del procedimiento de la invención. El medio puede ser un sistema (o aparato o dispositivo) electrónico, magnético, óptico, electromagnético, infrarrojo o semiconductor o un medio de propagación.

En un aspecto asociado, la presente invención se refiere al uso *in vitro* de los niveles de expresión de las isoformas TCFL5 y CHA del gen *TCFL5* determinados, preferiblemente de manera simultánea, en una muestra biológica aislada de un paciente con leucemia linfoblástica aguda (LLA) para el pronóstico y/o diagnóstico diferencial de pacientes con LLA, donde una reducción en los valores de expresión de TCFL5/CHA en la muestra del paciente respecto a valores de referencia es indicativa de LLA de alto riesgo.

Método para la personalización del tratamiento en función del riesgo

La invención hace también referencia a un método para la personalización del tratamiento o para la determinación del tratamiento más adecuado en función del perfil de riesgo del paciente, donde dicho método comprende el pronóstico y/o diagnóstico diferencial en función del riesgo según el método del primer aspecto de la invención.

En un aspecto relacionado, la invención hace referencia a un método para el tratamiento de pacientes de LLA que comprende la administración de una cantidad terapéuticamente efectiva de un fármaco o combinación de fármacos donde dicho tratamiento se determina en función de la clasificación o estratificación de dicho paciente en función del riesgo según el método del primer aspecto de la invención. Dicha terapia está típicamente formado por la asociación de varios fármacos, pudiendo incluir entre otros vincristina, prednisona y antraciclinas (daunorrubicina o idarubicinao doxorrubicina), L-asparaginasa, y metotrexato , etc((*Inaba, H., M. Greaves, and C.G. Mullighan, Acute lymphoblastic leukaemia. Lancet, 2013. 381(9881): p. 1943-55*).

En la práctica clínica actual aquellos pacientes de LLA clasificados como de alto riesgo recibirán un tratamiento quimioterapéutico de mayor intensidad. Una mayor intensidad del tratamiento se caracteriza por una mayor dosificación y/o frecuencia de las administraciones respecto a la administración del mismo fármaco o combinación de fármacos en un paciente clasificado como de riesgo medio o bajo. Asimismo, un tratamiento de mayor intensidad puede consistir también en la administración de otros fármacos cuya utilización se considera más agresiva, por ejemplo al estar asociados a una mayor toxicidad.

Método para la monitorización de la enfermedad y/o eficacia del tratamiento

En otro aspecto, la presente invención se refiere también a un método para la monitorización de la progresión de la enfermedad y/o eficacia del tratamiento en pacientes con LLA que comprende las siguientes etapas:

a. determinar, preferiblemente de manera simultánea, los niveles de expresión de las isoformas TCFL5 y CHA del gen *TCFL5* en una muestra biológica aislada de dicho paciente; y

b. comparar los niveles de expresión de TCFL5/CHA en la muestra del paciente con valores de referencia,

donde una reducción en los valores de la muestra del paciente respecto a los valores de referencia es indicativo de recaída. Típicamente, la toma de muestras se realiza al inicio del tratamiento, a lo largo del tratamiento y/o una vez finalizado el tratamiento, de manera cíclica o puntual.

En un aspecto relacionado, la invención hace referencia al uso *in vitro* de los niveles de expresión de las isoformas TCFL5 y CHA del gen *TCFL5* determinados, preferiblemente de manera simultánea, en una muestra biológica aislada de un paciente con leucemia linfoblástica aguda (LLA) para la monitorización de pacientes con LLA, donde una reducción en los valores de expresión de TCFL5/CHA en la muestra del paciente respecto a valores de referencia es indicativa de recaída.

En un aspecto adicional, la invención se refiere a un método de obtención de datos útiles para el pronóstico, diagnóstico diferencial y/o monitorización de pacientes con leucemia linfoblástica aguda (LLA) que comprende:

a. determinar, preferiblemente de manera simultánea, los niveles de expresión de las isoformas TCFL5 y CHA del gen *TCFL5* en una muestra biológica aislada de dicho paciente.

Realizaciones particulares y características preferidas de estos aspectos de la invención se han definido en aspectos anteriores, en particular en el primer aspecto de la invención.

Kit para el pronóstico, diagnóstico diferencial y/o monitorización de pacientes con LLA

En otro aspecto, la invención hace referencia a un kit para el pronóstico, diagnóstico diferencial y/o monitorización de pacientes con leucemia linfoblástica aguda (LLA) que comprende:

a. un cebador y/o sonda que comprende o consiste en una secuencia seleccionada del grupo que consiste en SEQ ID NO: 5 a SEQ ID NO: 70, y secuencias idénticas a cualquiera de las mismas en al menos 75%; y

b. opcionalmente, instrucciones para el uso de dicho cebador y/o sonda para la determinación de los niveles de expresión de dichas isoformas en una muestra biológica aislada de dicho paciente.

Preferiblemente, dicho cebador y/o sonda comprende o consiste en una secuencia seleccionada del grupo que consiste en SEQ ID NO: 45, SEQ ID NO: 46 y secuencias idénticas a cualquiera de las mismas en al menos 75%. Otras secuencias preferidas han sido descritas bajo el primer aspecto de la invención.

Realizaciones particulares y características preferidas de este aspecto de la invención se han definido en aspectos anteriores, en particular en el primer aspecto de la invención.

La invención se refiere también al uso de un kit para el pronóstico, diagnóstico diferencial y/o monitorización de pacientes con LLA en un método de acuerdo con los aspectos anteriores de la invención, donde dicho kit comprende:

a. un reactivo para determinar, preferiblemente de manera simultánea, los niveles de expresión de las isoformas TCFL5 y CHA del gen *TCFL5*; y
b. opcionalmente, instrucciones para el uso de dicho reactivo para la determinación de los niveles de expresión de dichas isoformas en una muestra biológica aislada de dicho paciente.

Realizaciones particulares y características preferidas de este aspecto de la invención se han definido en aspectos anteriores, en particular en el primer aspecto de la invención.

Secuencia nucleotídicas

En un aspecto adicional, la invención está relacionada con una secuencia nucleotídica, preferiblemente una sonda y/o cebador, que hibrida específicamente a SEQ ID NO: 1 y/o SEQ ID NO: 2. Dichas sondas y/o cebadores son parcialmente complementarias, preferiblemente sustancial o completamente complementarias a SEQ ID NO: 1 y/o SEQ ID NO: 2 o a fragmentos de las mismas tal y como se ha descrito en otros aspectos de la invención.

Preferiblemente, dicha secuencia es seleccionada del grupo que consiste en SEQ ID NO: 5 a SEQ ID NO: 70, y secuencias idénticas a cualquiera de las mismas en al menos 75%; más preferiblemente dicha secuencia es seleccionada del grupo que consiste en SEQ ID NO: 45, SEQ ID NO: 46, y secuencias idénticas a cualquiera de las mismas en al menos 75%. Otras secuencias preferidas han sido descritas bajo el primer aspecto de la invención.

En un aspecto relacionado, la invención se refiere al uso de una secuencia nucleotídica que hibrida específicamente a SEQ ID NO: 1 y/o SEQ ID NO: 2 como cebador y/o sonda en un método *in vitro* para el pronóstico, diagnóstico y/o monitorización de pacientes con LLA de acuerdo con los aspectos anteriores de la invención. Preferiblemente, dicha secuencia es seleccionada del grupo que consiste en SEQ ID NO: 5 a SEQ ID NO: 70, y secuencias idénticas a cualquiera de las mismas en al menos 75%. Más preferiblemente, dicha secuencia es seleccionada del grupo que consiste en SEQ ID NO: 45, SEQ ID NO: 46 y secuencias idénticas a cualquiera de las mismas en al menos 75%. Otras secuencias preferidas han sido descritas bajo el primer aspecto de la invención.

Realizaciones particulares y características preferidas de estos aspectos de la invención se han definido en aspectos anteriores, en particular en el primer aspecto de la invención.

En otro aspecto, la invención hace referencia a un método para la determinación *in vitro*, preferiblemente de manera simultánea, de los niveles de expresión de las isoformas TCFL5 y CHA del gen *TCFL5* en una muestra biológica aislada de un paciente con leucemia linfoblástica aguda (LLA); donde los niveles de expresión de las isoformas TCFL5 y CHA se determinan a nivel de RNAm mediante el uso de un cebador y/o sonda seleccionado del grupo que consiste en las secuencias SEQ ID NO: 5 a SEQ ID NO: 70, y secuencias idénticas a cualquiera de las mismas en al menos 75%. Preferiblemente, dicho cebador y/o sonda consiste en una secuencia nucleotídica seleccionada del grupo que consiste en SEQ ID NO: 45, SEQ ID NO: 46 y una secuencia idéntica a cualquiera de las mismas en al menos 75%. Otras secuencias preferidas han sido descritas bajo el primer aspecto de la invención.

Realizaciones particulares y características preferidas de estos aspectos de la invención se han definido en aspectos anteriores, en particular en el primer aspecto de la invención.

Se contempla que cualquier modo de realización analizado en esta memoria descriptiva se puede implementar con respecto a cualquier método, kit, reactivo o uso de la invención, y viceversa. En particular, aquellas características detalladas y realizaciones particulares relativas al primer aspecto de la invención podrán también implementarse respecto a otros aspectos de la invención. Se entenderá que los modos de realización particulares descritos en el presente documento se muestran a modo de ilustración y no como limitaciones de la invención. Los rasgos característicos principales de la presente invención se pueden emplear en varios modos de realización sin apartarse del alcance de la invención. Los expertos en la técnica reconocerán, o serán capaces de determinar usando no más que experimentación rutinaria, numerosos equivalentes a los procedimientos específicos descritos en el presente documento. Se considera que estos equivalentes están dentro del alcance de la presente invención y están contemplados por las reivindicaciones.

Todas las publicaciones y solicitudes de patentes mencionadas en la memoria descriptiva son indicativas del nivel de experiencia de los expertos en la técnica a la que pertenece la presente invención. Todas las publicaciones y solicitudes de patentes se incorporan en el presente documento por referencia en el mismo grado que si cada publicación o solicitud de patente individual se indicase específica e individualmente para incorporarse por referencia.

El uso de la palabra "un" o "una" cuando se usa junto con el término "que comprende" en las reivindicaciones y/o la memoria descriptiva puede querer decir "uno" pero también es coherente con el significado de "uno o más", "al menos uno" y "uno o más de uno". El uso del término "o" en las reivindicaciones se usa para querer decir "y/o" a menos que se indique explícitamente para referirse a alternativas solo o las alternativas son mutuamente exclusivas, aunque la divulgación respalda una definición que se refiere solo a alternativas y "y/o". En toda esta solicitud, el término "aproximadamente" se usa para indicar que un valor incluye la variación de error inherente para el dispositivo, empleándose el procedimiento para determinar el valor, o la variación que existe entre los sujetos de estudio.

Como se usa en esta memoria descriptiva y reivindicaciones, las palabras "comprender" (y cualquier forma de comprender, tal como "comprenden" y "comprende"), "tener" (y cualquier forma de tener, tal como "tienen" y "tiene"), "incluir" (y cualquier forma de incluir, tal como

"incluye" y "incluyen") o "contener" (y cualquier forma de contener, tal como "contiene" y "contienen") son inclusivas o abiertas y no excluyen elementos o etapas del procedimiento no citados, adicionales. El término “comprende” engloba y específicamente describe "consiste esencialmente en" y “consiste en”. Como se usa en el presente documento, la expresión "que consiste esencialmente en" limita el alcance de una reivindicación a los materiales o etapas especificados y a aquellos que no afectan materialmente a la(s) característica(s) básica(s) y novedosa(s) de la invención reivindicada. Como se usa en el presente documento, la expresión "que consiste en" excluye cualquier elemento, etapa o ingrediente no especificado en la reivindicación excepto, por ejemplo, impurezas habitualmente asociadas con el elemento o la limitación.

El término "o combinaciones de los mismos" como se usa en el presente documento se refiere a todas las permutaciones y combinaciones de los puntos enumerados que preceden al término. Por ejemplo, "A, B, C o combinaciones de los mismos" se pretende que incluya al menos uno de: A, B, C, AB, AC, BC o ABC, y si el orden es importante en un contexto particular, también BA, CA, CB, CBA, BCA, ACB, BAC o CAB. Continuando con este ejemplo, se incluyen expresamente combinaciones que contienen repeticiones de uno o más puntos o términos, tales como BBB, AAA, AB, BBC, AAABCCCC, CBBAAA, CABABB, y así sucesivamente. El experto en la técnica entenderá que típicamente no existe un límite sobre el número de puntos o términos en cualquier combinación, a menos que sea evidente de otro modo a partir del contexto.

Como se usa en el presente documento, palabras de aproximación tales como, sin limitación, "sobre”, "alrededor de”, “aproximadamente” se refieren a una condición que, cuando se modifica así, se entiende que no es necesariamente absoluta o perfecta sino que se consideraría lo suficientemente próxima para los expertos en la técnica para garantizar la designación de la condición como presente. La medida en que puede variar la descripción dependerá de lo grande que se pueda instituir un cambio y todavía reconozca un experto en la técnica que el rasgo característico modificado todavía tenga las características y capacidades requeridas del rasgo característico no modificado. En general, pero sujeto al análisis precedente, un valor numérico en el presente documento que se modifica por un palabra de aproximación tal como "aproximadamente" puede variar desde el valor establecido en ± 1, 2, 3, 4, 5, 6, 7, 10, 12 o 15 %, o menos, preferiblemente representa el valor establecido (± 0%)..

EJEMPLOS

Sitios WEB para el análisis de las bases de datos utilizados

- Secuencias:
  1. http://www.ncbi.nlm.nih.gov/
  2. http://www.ensembl.org/index.html
  3. http://vega.sanger.ac.uk/index.html
  4. http://genome.ewha.ac.kr/ECgene/

- Alineamientos:
  1. http://www.ncbi.nlm.nih.gov/
  2. http://www.ensembl.org/index.html
  3. http://multalin.toulouse.inra.fr/multalin/
  4. http://www.ncbi.nlm.nih.gov/spidey/

- Arrays:
  1. http://www.ncbi.nlm.nih.gov/
  2. https://www.oncomine.org/resource/login.html
  3. http://www-test.ebi.ac.uk/gxa/
  4. http://www.cbioportal.org/
  5. http://www.tumorportal.org

Materiales y Métodos

Obtención de muestras de pacientes

Se obtuvieron muestras de aspirado de médula ósea en el momento del diagnóstico o de la recaída, de 60 niños con LLA (edades 0-15 años) del Hospital Universitario Niño Jesús de Madrid. 18 LLA-T y 42 LLA-B

Clasificación de los pacientes

Se distinguieron 3 grupos de riesgo en función de una evaluación de los datos relativos a la edad del paciente y a los resultados obtenidos en una o más de las siguientes pruebas genéticas y/o bioquímicas:

- Hematimetría estándar con hemocitómetro.
- Citología convencional en extensión de muestra de médula ósea y en citospin de muestra de líquido cefalorraquídeo.
- Citometría de flujo multiparamétrica, preferiblemente en muestra de médula ósea.
- Técnicas citogenéticas clásicas y de bandeo de cromosomas, preferiblemente en muestra de médula ósea.
- RT-PCR con cebadores específicos para las translocaciones t(12;21), t(1;19), t(9;22) y t(4;11), preferiblemente en muestras de médula ósea.
- FISH con sondas para las translocaciones t(12;21), t(1;19), t(9;22) y break apart en MLL, preferiblemente en muestras de médula ósea.

En base a los resultados obtenidos en dichas pruebas, los pacientes de LLA se clasificaron en uno de los siguientes grupos de riesgo:

RIESGO ESTÁNDAR o RIESGO BAJO: El paciente debe reunir todos y cada uno de los siguientes criterios:

- Edad >1 y <10 años.
- Leucocitos <20 $x10^9$/l al diagnóstico.
- Inmunofenotipo no T.
- Ausencia de infiltración del SNC y/o testes.
- Citogenética (uno de los dos criterios siguientes es suficiente):
    - Alta Hiperdiploidía (51-67 cromosomas), índice de DNA 1,10-1,44 (siempre confirmado por otras técnicas citogenéticas).

- t(12;21) positiva

- Ausencia de t(1;19)
- No reordenamiento MLL
- Presencia de < 1.000 blastos/mm3 en día +8 de la Inducción, en sangre periférica
- Presencia de < 5% de blastos y < 0,1% de enfermedad residual mínima (ERM) en médula ósea en día +15 de la Inducción y al final de la inducción IA

RIESGO ALTO: La existencia de cualquiera de los siguientes criterios determina la inclusión del paciente en este grupo:

- t(4;11) (MLL/AF4).
- t(9;22) p190.
- Hipodiploidía <44 cromosomas o índice DNA <0,81 (se requiere confirmación por otras técnicas).
- > 1.000 blastos en día +8 de la Inducción, en sangre periférica.
- > 25% de blastos y >10% de ERM en el día +15 de la Inducción, en médula ósea.
- ERM > 1% en el día +33 de la Inducción, en médula ósea.
- ERM > 0,1% antes de la Consolidación, en médula ósea.

RIESGO MEDIO: Aquellos pacientes que no reúnan los criterios de Riesgo Estándar ni de Riesgo Alto.

Western blot

Se utilizaron 20 ng de proteína correspondientes a las muestras de los pacientes por pocillo en geles al 10% de SDS-poliacrilamida y se aplicaron 90Voltios hasta que el frente de azul de bromocresol migró a la posición deseada. Las proteínas fueron transferidas (100Voltios a 4ºC) a una membrana de nitrocelulosa durante 1h 30min con el *buffer* de transferencia.

Una vez transferidas las proteínas a la membrana de nitrocelulosa, se bloquearon uniones inespecíficas mediante su incubación durante una hora en una solución de BSA (Suero de Albúmina Bovina) al 5% en TBS-T. Antes de realizar el bloqueo, se realizó una tinción reversible con Rojo Ponceau que permite visualizar las proteínas transferidas.

Los tiempos de incubación al igual que las concentraciones utilizadas de los anticuerpos se realizaron siguiendo las recomendaciones del fabricante. Se utilizó el anticuerpo policlonal anti TCFL5/CHA (SAB4500152, Sigma-Aldrich) y el anticuerpo secundario policlonal de cabra anti conejo Goat (A-21431, Invitrogen). Entre dichas incubaciones se realizan 3 lavados de 15min con TBS-T y el último lavado con TBS.

El revelado de la membrana se realizó con el kit *SuperSignal* (Thermo Scientific) o el kit Immun-Star HRP (BIO_RAD) siguiendo las recomendaciones del fabricante.

Para la reutilización de las membranas se realizó un *stripping* con el *Restore Western Blot Stripping Buffer* (Thermo Scientific) siguiendo las recomendaciones del fabricante.

Aislamiento de células mononucleadas a partir de muestras de médula ósea

La extracción de células mononucleadas de pacientes se realizó mediante centrifugación por gradiente de densidad. Las muestras de medula osea de pacientes se diluyen a la mitad en PBS. En un tubo a parte se preparan 5 ml de Ficoll (Ficoll® Paque Plus, Sigma-Aldrich). La muestra de médula osea diluida se pone sobre el ficoll lentamente apoyando la punta sobre la pared del tubo para generar dos fases. Se centrifuga a 1500 rpm durante 25 minutos con deceleración baja a temperatura ambiente. Nos quedamos con la fase blanca que es donde se encuentran las células mononucleadas. Se lavan con PBS 1400 rpm durante 7 minutos para eliminar restos de ficoll. Tras el lavado, el pellet se resuspende en suero fetal bovino (FBS) y se criopreserva en viales de como máximo $12{,}5x10^6$ células.

Extracción de RNA y RT-qPCR (Protocolo utilizado en el ejemplo 5)

Una vez aisladas las células la extracción del RNA se llevó a cabo con RNeasy Plus Mini Kit (Qiagen) siguiendo las instrucciones del fabricante. Para la retrotranscripción se utilizó el Kit TaqMan Reverse Transcription Reagents (Applied Biosystem, Life Technologies) siguiendo el programa: un ciclo a 42°C durante 45 minutos y un último ciclo a 99°C durante 3 minutos. Por último para la PCR a tiempo real se utilizó SYBR GREEN PCR Master Mix (Applied Biosystem, Life Technologies) siguiendo el programa: un ciclo 50°C durante 2 min, 95°C duante 10 minutos, 40 ciclos a 95°C durante 15 segundos y 58°C durante 1 minutos, curva de melting continua. La secuencia de los oligonucleótidos utilizados son TCFL5/CHA_F (SEQ ID NO: 45): 5' GAGACTGACAAGGCCACAACT 3', TCFL5/CHA_R (SEQ ID NO:46): 5'

CCGCAAAATACGCTCTCAA 3', Beta-actina_F (SEQ ID NO: 71): CATGTACGTTGCTATCCAGGC; Beta-actina_R (SEQ ID NO: 72): CTCCTTAATGTCACGCACGAT.

Extracción de RNA y RT-qPCR (Protocolo utilizado en el ejemplo 6)

Las muestras se conservaron en TRIzol (Life Technologies) a 70ºC hasta su uso. La extracción del RNA total se realizó con RNeasy Plus Mini Kit (Qiagen) siguiendo las recomendaciones del fabricante.

RT-qPCR cuantitativa. El RNAm se retrotranscribió a DNA complementario (DNAc) y se cuantificaron los genes con oligonucleótidos específicos (SEQ ID NO: 45 (TCFL5/CHA 01 Forward) y SEQ ID NO: 46 (TCFL5/CHA 01 reverse) de la Tabla III) con el kit GoTaq 2-Step RT-PCR System (Promega) siguiendo las recomendaciones del fabricante.

Para la amplificación de los genes por PCR se sometieron las muestras a 95ºC durante 5min, a continuación se realizaron 45 ciclos de 15s a 95ºC, 30s a 59ºC y 15s a 72ºC; finalmente, se realiza un ciclo a 95ºC en el termociclador ABIPrism 7900 HT SDS (Applied Biosystems).

Los cálculos de cuantificación de la expresión de genes se realizaron por el método de comparación del ciclo umbral (*Comparative Threshold Method*, C_T). Los valores obtenidos para cada gen se normalizaron con el gen *House keeping* HPRT (Hypoxantina-guanina fosforibosiltransferasa) para obtener el ΔC_T, tras lo cual se normalizaron con el valor de la muestra control para obtener el $\Delta\Delta C_T$. La cantidad relativa, RQ, se calculó a partir de la siguiente formula $RQ = 2^{-\Delta\Delta CT}$. Los cebadores utilizados para cuantificar la expresión de HPRT fueron HPRT- Forward (SEQ ID NO: 73): CTGGAAAGAATGTCTTGATTGTGG; y HPRT-Reverse (SEQ ID NO: 74): CATCTTGGATTATACTGCCTGAC.

Análisis estadísticos

Los análisis de significancia y correlación estadísticos de expresión génica, se realizaron con las herramientas de análisis de GraphPad (5.0 para Windows) mediante los tests de ANOVA y T de Student para los análisis individuales y, de Spearman para los estudios de correlación tomando los datos como no paramétricos.

Resultados

Ejemplo 1.- Determinación de mensajeros consenso e identificación de nuevas isoformas de TCFL5

Se efectuó un análisis bioinformático partiendo de datos de secuencias de RNAm de TCFL5 publicadas y analizando lecturas de uniones de exones, se establecieron los mensajeros consenso expresados en humano aumentando la cifra de las 2 isoformas previamente descritas (TCFL5 y CHA) a 6 isoformas.

La Tabla I muestra los RNAm consenso o isoformas identificadas a partir de las secuencias de RNAm publicadas en las distintas bases de datos.

		NCBI	**Ensemble**	**Vega**	**EC gene**
Gene_ID		10732	ENSG00000101190	OTTHUMG00000032939	-
Variantes (mRNA)	**Exones**				
TCFL5R (R)	E1 E2a E3 E4 E5 E8	NM_006602, AB012124.1, AF070992.1	ENST00000335351	OTTHUMT00000080079	H20C8515.6, H20C8515.3
TCFL5R4b (R4b)	E1 E2a E3 E4b E5 E8	XM_005260185.1			
TCFL5R7 (R7)	E1 E2a E3 E4 E5 E7	XM_005260184.1, BC046933.1, H20C8515.9	ENST00000217162		
TCFL5R4b6 (R4b6)	E1 E2a E3 E4b E5 E6	XM_005260186.1, BC065520.1,			

TCFL5R6 (R6)	E1 E2a E3 E4 E5 E6	H20C8515.9			
CHA (R2b)	E2b E3 E4 E5 E8	AJ271337.1			H20C8515.5, H20C8515.2

Tabla I. Se indican los códigos de acceso para cada uno de los RNAm de TCFL5 descritos en bases de datos y la asociación de acuerdo con el análisis bioinformático efectuado por los inventores.

La Tabla II proporciona datos adicionales referentes a las variantes o isoformas y a las secuencias de RNAm correspondientes.

Variantes	Comentarios
R	Variante de referencia (RefSeq, NM_006602). H20C8515.3 presenta un 3'UTR más largo. AB012124.1 comienza 39 bases después saltándose el 1er ATG, por lo que podría ser un mensajero parcial. Además, presenta un 3'UTR más corto. Q9UL49-1 es la traducción de este mensajero. AF070992.1 comienza 170 bases después saltándose el 1er ATG, por lo que podría ser un mensajero
R4b	Esta variante es idéntica a R salvo porque su exón 4 tiene 3 bases menos en su extremo 5' (exón 4b). Existen EST que soportan la existencia de esta variante.
R7	Presenta un exón final alternativo (denominado exón 7). XM_005260184.1, ENST00000217162, BC046933.1 y H20C8515.8 muestran diferencias en la longitud de sus UTRs.
R4b6	Presenta un exón final alternativo (denominado exón 6) y utiliza el exón 4 alternativo (3 bases menos en su extremo 5', exón 4b). BC065520.1 presenta algunos errores de secuenciación.
R6	Idéntica a R4b6 pero utiliza el exón 4 completo (exón 4a). Soportado por una EST.
R2b	CHA. Utiliza un exón inicial alternativo (exón 2b) que contiene un 5'UTR y una región codificante coincidente con E2a. H20C8515.2 y H20C8515.5 presentan 3'UTRs más largos.

En base a dicho análisis bioinformático, se estableció que el gen *TCFL5* se compone de 8 exones (ver Figura 1) siendo los exones 3 (E3) y 5 (E5) comunes a todas las isoformas. El exón 1 (E1) está presente en 5 de las 6 isoformas, faltando su expresión únicamente en la isoforma denominada CHA (R2b). El último exón puede estar representado por el exón 8, en el caso de la isoforma consenso TCFL5 (TCFL5R) y la isoforma CHA, por el exón 7 o bien por el exón 6. A su vez, la isoforma consenso TCFL5 (TCFL5R) y la isoforma que acaba con

el exón 6, pueden presentar un exón 4 al cual le faltan los primeros codones (R4b6). Por su parte, la isoforma CHA presenta un exón 2 distinto al resto (E2b).

Ejemplo 2.- Diseño de oligonucleotidos para la detección de TCFL5/CHA

En base a dicha determinación de los exones y variantes de los mismos que componen cada una de las isoformas, se procedió al diseño de cebadores que cuantifican TCFL5, CHA de manera independiente y TCFL5/CHA conjuntamente.

Para la cuantificación de TCFL5/CHA se diseñaron cebadores que amplificaran una región común a ambas isoformas, es decir la región comprendida entre los exones 3 y 8 (ver Figs. 1 y 3).

Asimismo, para la cuantificación de las isoformas TCFL5 y CHA de manera específica, es decir cuantificando únicamente las isoformas TCFL5 y CHA (TCFL5R2b), se diseñaron cebadores específicos para la amplificación de una región comprendida entre los exones 5 y 8 de TCFL5 (ver Figs. 1 y 3).

Tabla III.-

	Primers para amplificar todas las isoformas				
Nombre Pareja cebadores	Nombre secuencia	Forward (5' - 3')	Nombre secuencia	Reverse (5' - 3')	Exon
TodoTCFL5_01	SEQ ID NO:5	TGCCTGAGCAAGTTTGGATT	SEQ ID NO:6	CGCATTCTACTCCGATTCCT	3 y 4
TodoTCFL5_02	SEQ ID NO:7	GAGTCCTCACAGGCAAACCT	SEQ ID NO:8	GATTCAACTCATCACAGCAAATG	4 y 5
TodoTCFL5_03	SEQ ID NO:9	CAACGTAGGGAGAGGCATAAC	SEQ ID NO:10	CAGAACGGCACTAAGAGATTCA	4 y 5
TodoTCFL5_04	SEQ ID NO:11	GCCTGAGCAAGTTTGGATTAAAG	SEQ ID NO:12	GTGTCCAACTGACGCATTCTA	3 y 4
TodoTCFL5_05	SEQ ID NO:13	CACTAAACAGACGTTAGGTAGTAGAA	SEQ ID NO:14	GATTCCTCTTCAGTGCTTGTTTG	3 y 4
TodoTCFL5_06	SEQ ID NO:15	AGGCATAACCGAATGGAAAG	SEQ ID NO:16	TTGCAGAACGGCACTAAGAG	4 y 5
TodoTCFL5_07	SEQ ID NO:17	TGCCTGAGCAAGTTTGGAT	SEQ ID NO:18	ACGCATTCTACTCCGATTCC	3 y 4
TodoTCFL5_08	SEQ ID NO:19	AGAAATTGAATCCACTAAACAGACG	SEQ ID NO:20	CCAACTGACGCATTCTACTCC	3 y 4

TodoTCFL5_09	SEQ ID NO:21	TCAACGTAGGGAGAGGCATA	SEQ ID NO:22	ACGGCACTAAGAGATTCAACTC	4 y 5
TodoTCFL5_10	SEQ ID NO:23	TGCAATTTACATACCCACTGTTTAC	SEQ ID NO:24	CTCTAGGCAATCCAATATCCTGG	2 y 3
TodoTCFL5_11	SEQ ID NO:25	TGAATCTCTTAGTGCCGTTCTG	SEQ ID NO:26	CTCCATGTCTTTCCTGGATGTAT	5
TodoTCFL5_12	SEQ ID NO:27	GAGCCCTTGGAGAGATTCAG	SEQ ID NO:28	AGGTGCCACCGCCACACAAG	5
TodoTCFL5_13	SEQ ID NO:29	CTCTTAGTGCCGTTCTGCAA	SEQ ID NO:30	CTGTGGTCCACTGCAGAGTT	5
TodoTCFL5_14	SEQ ID NO:31	GGGAGGAAGGTCTCAACGTA	SEQ ID NO:32	GGATTCTGCGCCTTCTATCT	4
TodoTCFL5_15	SEQ ID NO:33	GAGCTCTGCTCGGCTAGTCT	SEQ ID NO:34	GAGCTCTGCTCGGCTAGTCT	4
TodoTCFL5_16	SEQ ID NO:35	GCAAACAAGCACTGAAGAGG	SEQ ID NO:36	GGGCTCTTCGCTCTACATTT	4
TodoTCFL5_17	SEQ ID NO:37	TGCCTGAGCAAGTTTGGAT	SEQ ID NO:38	ACGCATTCTACTCCGATTCC	3 y 4
TodoTCFL5_18	SEQ ID NO:39	AGAGCCCTTGGAGAGATTCA	SEQ ID NO:40	CCTACGCTGAGACCTTCCTC	3 y 4
TodoTCFL5_19	SEQ ID NO:41	AAGGTCTCAGCGTAGGGAGA	SEQ ID NO:42	TTGCAGAACGGCACTAAGAG	3 y 4
	Primers para amplificar todas las isoformas menos CHA				
Nombre pareja cebadores	Nombre secuencia	Forward (5' - 3')	Nombre secuencia	Reverse (5' - 3')	Exon
*TCFL5*noCHA_01	SEQ ID NO:43	TTCAACAGCATCCCCGC	SEQ ID NO:44	GTTGCTGAAGAGGAACATTCAT	1 y 2
	Primers para amplificar TCFL5/CHA				
Nombre pareja cebadores	Nombre secuencia	Forward (5' - 3')	Nombre secuencia	Reverse (5' - 3')	Exon
TCFL5/CHA_01	SEQ ID NO:45	GAGACTGACAAGGCCACAACT	SEQ ID NO:46	CCGCAAAATACGCTCTCAA	5 y 8
TCFL5/CHA_02	SEQ ID NO:47	CTGACAAGGCCACAACTCTG	SEQ ID NO:48	CTTTAGCCTTCGGCCAGTT	5 y 8
TCFL5/CHA_03	SEQ ID NO:49	ATCTCTTAGTGCCGTTCTGC	SEQ ID NO:50	CCGCAAAATACGCTCTCAAATTC	5 y 8
TCFL5/CHA_04	SEQ ID NO:51	GTGATCGGACTGAACAGGAA	SEQ ID NO:52	CGAGTTTCAGGGATGACCTC	5 y 8
TCFL5/CHA_05	SEQ ID NO:53	CTGACAAGGCCACAACTCTG	SEQ ID NO:54	CTTTAGCCTTCGGCCAGTT	5 y 8

TCFL5/CHA_06	SEQ ID NO:55	GTCCCGTCAGCTTGAGTCACG	SEQ ID NO:56	CTGCAGTGGACCACAGCATTCC	3 y 8
TCFL5/CHA_07	SEQ ID NO:57	AAACTGGCCGAAGGCTAAA	SEQ ID NO:58	GTCCGATCACTTGATCTCCATC	8
TCFL5/CHA_08	SEQ ID NO:59	GTGATCGGACTGAACAGGAA	SEQ ID NO:60	TACTGGAGTCCCGTCAGCTT	8
TCFL5/CHA_9	SEQ ID NO:61	AATTTGAGAGCGTATTTTGCGG	SEQ ID NO:62	GGATTCCTGTTCAGTCCGATC	8
	Primers para amplificar CHA				
Nombre pareja cebadores	Nombre secuencia	Forward (5' - 3')	Nombre secuencia	Reverse (5' - 3')	Exon
SoloCHA_01	SEQ ID NO:63	GGAAGAGGAGGACTTCCACA	SEQ ID NO:64	AGCAATACGTGGCAGACAAG	2b
SoloCHA_02	SEQ ID NO:65	GCTCATCCCAGCCTGTAGTT	SEQ ID NO:66	TGTGGAAGTCCTCCTCTTCC	2b
SoloCHA_03	SEQ ID NO:67	CATTCTTGGATGTCAGGGAA	SEQ ID NO:68	TAGAGCTCAAAGCCTGCAAA	2b
SoloCHA_04	SEQ ID NO:69	GAGAGTTGGGTTGCTGTCTGG	SEQ ID NO:70	GTTCAGATGGATGTCGAATGAGAG	2b

Ejemplo 3.- Diferente expresión de TCFL5 y CHA en tejidos y líneas celulares humanos

Se analizó la expresión de E1 y E2 codificante en las bases de datos citadas anteriormente y se observó una expresión preferente de E1 en cerebro, corazón, hígado, pulmón, riñón y testículo, mientras que el E2 se expresó preferencialmente en la línea celular T Jurkat, células mononucleadas de sangre periférica (PBMNCs), monocitos y células T y B. Los resultados se muestran en la Fig. 2.

Ejemplo 4.- Determinación de los niveles de expresión de TCFL5 y CHA en líneas celulares tratadas con DAPT

Se determinaron mediante Western Blot los niveles de expresión de las isoformas TCFL5, CHA y HSP90 (como control), en células Jurkat (Clone E6-1, ATCC® TIB-152™). Los resultados en la Fig.5 muestran que se observa una diferencia de expresión entre las isoformas CHA y TCFL5 para distintas concentraciones de DAPT.

Ejemplo 5.- Comparación de los niveles de expresión de TCFL5/CHA RNAm en muestras de pacientes de LLA (LLA-B vs LLA-T, diagnóstico vs recaídas, riesgo alto vs medio vs bajo, y riesgo no alto vs riesgo alto)

Los pacientes fueron clasificados en función del riesgo tal y como se describe en materiales y métodos. Los niveles de expresión de RNAm de TCFL5/CHA fueron determinados mediante Real Time qPCR.

Los datos se muestran en la Tablas IV-IX y en las gráficas en la figura 6 A, B y C, correspondientes a los datos en las Tablas IV, V y VI respectivamente. Se observan diferencias significativas entre las muestras en el momento del diagnóstico vs recaídas, LLA-B vs LLA-T y entre muestras de pacientes con bajo, medio y alto riesgo de recaída.

Tabla IV.- LLA-B (B) versus LLA-T (T), solo muestras al diagnóstico:

	media B	eem B	media T	eem T	p
TCFL5 CHA	0,0108	0,0042	0,0009	0,0004	0,1150
TCFL5	0,0025	0,0008	0,0004	0,0002	0,3994
CHA	0,1299	0,0441	0,0404	0,0209	0,9484

Tabla V.- Diagnósticos (DX) versus recaídas (REC):

	media DX	eem DX	media REC	eem REC	p
TCFL5 CHA	0,0095	0,0037	0,0002	0,00005	0,0099
TCFL5	0,0023	0,0008	0,0003	0,00005	0,6847
CHA	0,1182	0,0386	0,0306	0,0275	0,0396

Tabla VI.- Riesgo bajo (B), medio (M), alto (A), sólo muestras al diagnóstico:

	media B	eem B	media M	eem M	media A	eem A
TCFL5 CHA	0,0171	0,0075	0,0035	0,0017	0,0011	0,0004
TCFL5	0,0033	0,0014	0,0011	0,0005	0,0013	0,0013
CHA	0,1381	0,0689	0,0667	0,0295	0,1504	0,0926

Tabla VII.- Riesgo no alto (NA) versus riesgo alto (A), solo muestras al diagnóstico:

	media NA	eem NA	media A	eem A	p
TCFL5 CHA	0,0117	0,0046	0,0011	0,0004	0,0734
TCFL5	0,0025	0,0009	0,0013	0,0013	0,4161
CHA	0,1095	0,0429	0,1504	0,0926	0,9714

Las tablas siguientes presentan dichas comparaciones efectuadas sólo con las muestras de LLA-B (el grupo más numeroso).

Tabla VIII.- Diagnósticos (DX) versus recaídas (REC):

	media DX	eem DX	media REC	eem REC	p
TCFL5 CHA	0,0108	0,0043	0,0003	0,00005	0,0217
TCFL5	0,0025	0,0008	0,0003	0,00005	0,8637
CHA	0,1199	0,0441	0,0349	0,0314	0,1215

Tabla IX.- Riesgo bajo (B), medio (M), alto (A), sólo muestras al diagnóstico:

	media B	eem B	media M	eem M	media A	eem A
TCFL5 CHA	0,0171	0,0075	0,0054	0,0028	0,0011	0,0004
TCFL5	0,0033	0,0014	0,0016	0,0007	0,0013	0,0013
CHA	0,1381	0,0689	0,0856	0,0488	0,1504	0,0926

Tabla X.- Riesgo no alto (NA) versus riesgo alto (A), solo muestras al diagnóstico:

	media NA	eem NA	media A	eem A	p
TCFL5 CHA	0,0138	0,0055	0,0011	0,0004	0,026
TCFL5	0,0028	0,001	0,0013	0,0013	0,3711
CHA	0,1234	0,0511	0,1504	0,0926	1

eem = error estandar de la media; p = significacion estadistica

A los resultados de la tabla VI se le aplicó un análisis de discriminantes mediante el programa IBM SPSS Statistics Base con el objeto de averiguar el valor pronóstico (predictivo de riesgo de recaída) de los valores de expresión de TCFL5/CHA. Los resultados indican que los valores de expresión de TCFL5/CHA permiten predecir un 87% de los pacientes de alto riesgo (Tabla XI)

Tabla XI.- Pronóstico según los grupos de riesgo de recaída bajo, medio y alto según los niveles de expresión de TCFL5/CHA.

			Pronóstico según los niveles de TCFL5/CHA			
		Riesgo	Bajo	Medio	Alto	Total
Clasificación clínica	Recuento	Bajo	6	6	6	18
		Medio	1	4	7	12
		Alto	0	1	7	8
	%	Bajo	33,3	33,3	33,3	100,0
		Medio	8,30	33,3	58,3	100,0
		Alto	0,0	12,5	87,5	100,0

Ejemplo 6.- Determinación de TCFL5/CHA RNAm en pacientes de LLA (diagnóstico vs recaídas)

Determinación de los niveles de expresión de las isoformas TCFL5 y CHA a nivel de RNAm y proteína en muestras de pacientes de LLA clasificadas de alto o bajo riesgo de acuerdo con los criterios indicados en materiales y métodos. Los resultados se muestran en las Figuras 7A y 7B donde se observan mayores niveles de expresión de TCFL5/CHA en el momento del diagnóstico.

REIVINDICACIONES

1. Método pronóstico y/o para el diagnóstico diferencial de pacientes con leucemia linfoblástica aguda (LLA) que comprende las siguientes etapas:
   a. determinar de manera específica y simultánea los niveles de expresión de las isoformas TCFL5 y CHA del gen *TCFL5* en una muestra biológica aislada de dicho paciente; y
   b. comparar los niveles de expresión de TCFL5/CHA en la muestra del paciente con valores de referencia,

   donde una reducción en los valores de la muestra del paciente respecto a los valores de referencia es indicativa de LLA de alto riesgo, y

   donde el término isoforma TCFL5 se refiere a las isoformas TCFL5R y TCFL5R4b.

2. El método según la reivindicación 1, donde dichas muestras de referencia corresponden a una o más muestras de pacientes con LLA o a líneas establecidas de LLA.

3. El método según cualquiera de las reivindicaciones 1 o 2, donde los niveles de expresión de las isoformas TCFL5 y CHA se determinan a nivel de RNA mensajero (RNAm).

4. El método según la reivindicación 3, donde los niveles de expresión de las isoformas

5. El método según cualquiera de las reivindicaciones 1 a 4, donde la determinación de los niveles de expresión de dichas isoformas se efectúa mediante el uso de una pareja de cebadores seleccionada del grupo que consiste en:
   a) un cebador que hibrida específicamente con una secuencia comprendida en el exón E3 y un cebador que hibrida con una secuencia comprendida en el exón E8;
   b) un cebador que hibrida específicamente con una secuencia comprendida en el exón E5 y un cebador que hibrida con una secuencia comprendida en el exón E8;
   c) una pareja de cebadores donde ambos hibridan específicamente con una secuencia comprendida en el exón E8.

6. El método según la reivindicación 5, donde dichos cebadores son seleccionados del grupo que consiste en las secuencias SEQ ID NO: 45 a SEQ ID NO: 62.

7. El método según la reivindicación 6, donde dichos cebadores son seleccionados del grupo que consiste en las secuencias SEQ ID NO: 45 y SEQ ID NO: 46.

8. El método según cualquiera de las reivindicaciones 1 o 2, donde los niveles de expresión de las isoformas TCFL5 y CHA se determinan a nivel proteico.

9. El método según la reivindicación 8, donde los niveles de expresión de las isoformas TCFL5 y CHA se determinan mediante un inmunoensayo.

10. El método según cualquiera de las reivindicaciones 1 a 9, donde dicha muestra biológica es una muestra aislada de médula ósea, de sangre periférica o de líquido cefalorraquídeo.

11. El método según la reivindicación 10, donde dicha muestra consiste en células mononucleares aisladas.

12. El método según cualquiera de las reivindicaciones 10 o 11, donde dicha muestra ha sido obtenida en el momento del diagnóstico.

13. El método según cualquiera de las reivindicaciones 1 a 12, donde dicha leucemia linfoblástica aguda es una leucemia linfoblástica aguda de células B (LLA-B) o una leucemia linfoblástica aguda de células T (LLA-T).

14. El método según la reivindicación 13, donde dicha leucemia linfoblástica aguda es LLA-B.

15. El método según cualquiera de las reivindicaciones 1 a 14, donde dichos pacientes tienen menos de 20 años.

16. El método según la reivindicación 15, donde dichos pacientes tienen entre 0 y 15 años.

17. El método según cualquiera de las reivindicaciones 1 a 16, donde dicho método comprende además la detección o cuantificación de uno o más marcadores genéticos asociados a LLA de alto riesgo.

18. El método según cualquiera de las reivindicaciones 1 a 17, donde dicho método comprende además la determinación de la presencia de una o más de las translocaciones seleccionadas del grupo que consiste en t(12;21), t(1;19), t(9;22) y t(4;11), Hipodiploidía <44 cromosomas o índice DNA <0,81.

19. El método según cualquiera de las reivindicaciones 1 a 18, donde dicho método comprende además almacenar los resultados de dicho método en un dispositivo de almacenamiento de datos.

20. Un método para la determinación del tratamiento más adecuado para un paciente de LLA que comprende el pronóstico y/o diagnóstico diferencial según el método de cualquiera de las reivindicaciones 1 a 19.

21. Método para la monitorización de la progresión de la enfermedad y/o eficacia del tratamiento en pacientes con LLA que comprende las siguientes etapas:
    a. determinar de manera específica y simultánea los niveles de expresión de las isoformas TCFL5 y CHA del gen *TCFL5* en una muestra biológica aislada de dicho paciente; y
    b. comparar los niveles de expresión de TCFL5/CHA en la muestra del paciente con valores de referencia,

    donde una reducción en los valores de la muestra del paciente respecto a los valores de referencia es indicativo de recaída; y

    donde el término isoforma TCFL5 se refiere a las isoformas TCFL5R y TCFL5R4b.

22. El método según la reivindicación 21, donde dichas muestras de referencia corresponden a una o más muestras de pacientes con LLA o a líneas establecidas de LLA.

23. Un kit para el pronóstico, diagnóstico diferencial y/o monitorización de pacientes con LLA que comprende:
    a. una pareja de cebadores donde dichos cebadores son seleccionados del grupo que consiste en las secuencias SEQ ID NO: 45 a SEQ ID NO: 62,; y
    b. opcionalmente, instrucciones para el uso de dicha pareja de cebadores para la determinación de los niveles de expresión de dichas isoformas en una muestra biológica aislada de dicho paciente.

24. Uso de un kit para el pronóstico, diagnóstico diferencial y/o monitorización de pacientes con LLA en un método según cualquiera de las reivindicaciones 1 a 22, donde dicho kit comprende:
    a. un reactivo para determinar de manera específica y simultánea los niveles de expresión de las isoformas TCFL5 y CHA del gen *TCFL5*; y
    b. opcionalmente, instrucciones para el uso de dicho reactivo para la determinación de los niveles de expresión de dichas isoformas en una muestra biológica aislada de dicho paciente;

    donde el término isoforma TCFL5 se refiere a las isoformas TCFL5R y TCFL5R4b.

25. Una secuencia nucleotídica seleccionada del grupo que consiste en SEQ ID NO: 45 a SEQ ID NO: 62.

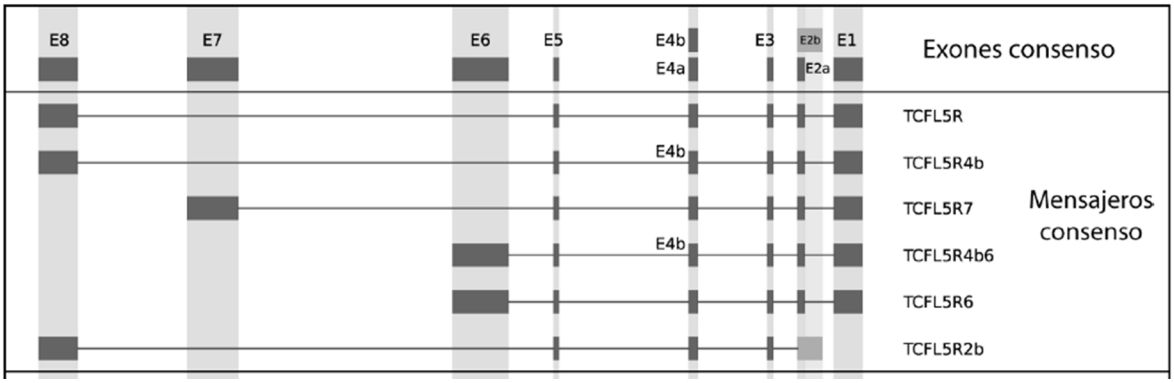

E8
E7
E6
E5
E4b
E4a
E3
E2b
E2a
E1
Exones consenso
TCFL5R
TCFL5R4b
TCFL5R7
TCFL5R4b6
TCFL5R6
TCFL5R2b
Mensajeros consenso

Fig. 1

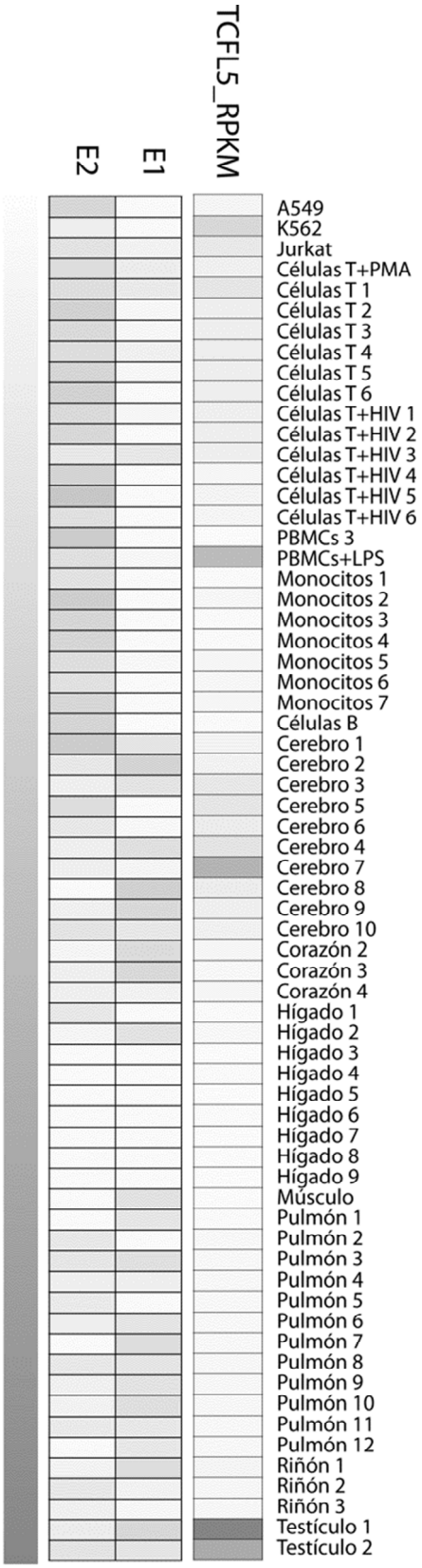

E2
E1
TCFL5_RPKM
A549
K562
Jurkat
Células T+PMA
Células T 1
Células T 2
Células T 3
Células T 4
Células T 5
Células T 6
Células T+HIV 1
Células T+HIV 2
Células T+HIV 3
Células T+HIV 4
Células T+HIV 5
Células T+HIV 6
PBMCs 3
PBMCs+LPS
Monocitos 1
Monocitos 2
Monocitos 3
Monocitos 4
Monocitos 5
Monocitos 6
Monocitos 7
Células B
Cerebro 1
Cerebro 2
Cerebro 3
Cerebro 5
Cerebro 6
Cerebro 4
Cerebro 7
Cerebro 8
Cerebro 9
Cerebro 10
Corazón 2
Corazón 3
Corazón 4
Hígado 1
Hígado 2
Hígado 3
Hígado 4
Hígado 5
Hígado 6
Hígado 7
Hígado 8
Hígado 9
Músculo
Pulmón 1
Pulmón 2
Pulmón 3
Pulmón 4
Pulmón 5
Pulmón 6
Pulmón 7
Pulmón 8
Pulmón 9
Pulmón 10
Pulmón 11
Pulmón 12
Riñón 1
Riñón 2
Riñón 3
Testículo 1
Testículo 2

Fig. 2

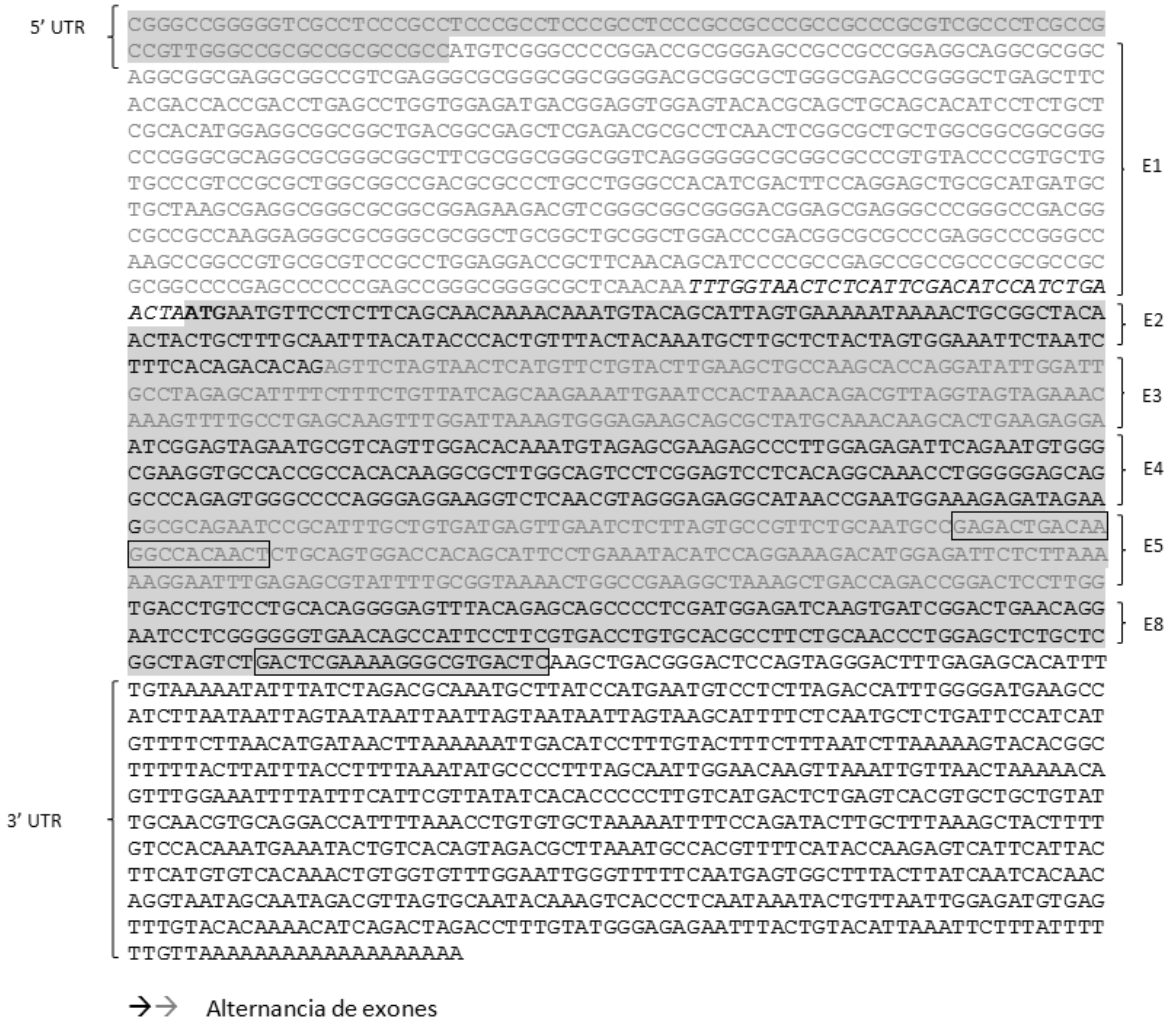

5' UTR
CGGGCCGGGGGTCGCCTCCCGCCTCCCGCCTCCCGCCTCCCGCCGCCCGCCGCCCGCGTCGCCCTCGCCG
CCGTTGGGCCGCGCCGCGCCGCCATGTCGGGCCCCGGACCGCGGGAGCCGCCGCCGGAGGCAGGCGCGGC
AGGCGGCGAGGCGGCCGTCGAGGGCGCGGGCGGCGGGGACGCGGCGCTGGGCGAGCCGGGGCTGAGCTTC
ACGACCACCGACCTGAGCCTGGTGGAGATGACGGAGGTGGAGTACACGCAGCTGCAGCACATCCTCTGCT
CGCACATGGAGGCGGCGGCTGACGGCGAGCTCGAGACGCGCCTCAACTCGGCGCTGCTGGCGGCGGCGGG
CCCGGGCGCAGGCGCGGGCGGCTTCGCGGCGGGCGGTCAGGGGGGCGCGGCGCCCGTGTACCCCGTGCTG
TGCCCGTCCGCGCTGGCGGCCGACGCGCCCTGCCTGGGCCACATCGACTTCCAGGAGCTGCGCATGATGC
TGCTAAGCGAGGCGGGCGCGGCGGAGAAGACGTCGGGCGGCGGGGACGGAGCGAGGGCCCGGGCCGACGG
CGCCGCCAAGGAGGGCGCGGGCGCGGCTGCGGCTGCGGCTGGACCCGACGGCGCGCCCGAGGCCCGGGCC
AAGCCGGCCGTGCGCGTCCGCCTGGAGGACCGCTTCAACAGCATCCCCGCCGAGCCGCCGCCCGCGCCGC
GCGGCCCCGAGCCCCCCGAGCCGGGCGGGGCGCTCAACAATTTGGTAACTCTCATTCGACATCCATCTGA
ACTAATGAATGTTCCTCTTCAGCAACAAAACAAATGTACAGCATTAGTGAAAAATAAAACTGCGGCTACA
ACTACTGCTTTGCAATTTACATACCCACTGTTTACTACAAATGCTTGCTCTACTAGTGGAAATTCTAATC
TTTCACAGACACAGAGTTCTAGTAACTCATGTTCTGTACTTGAAGCTGCCAAGCACCAGGATATTGGATT
GCCTAGAGCATTTTCTTTCTGTTATCAGCAAGAAATTGAATCCACTAAACAGACGTTAGGTAGTAGAAAC
AAAGTTTTGCCTGAGCAAGTTTGGATTAAAGTGGGAGAAGCAGCGCTATGCAAACAAGCACTGAAGAGGA
ATCGGAGTAGAATGCGTCAGTTGGACACAAATGTAGAGCGAAGAGCCCTTGGAGAGATTCAGAATGTGGG
CGAAGGTGCCACCGCCACACAAGGCGCTTGGCAGTCCTCGGAGTCCTCACAGGCAAACCTGGGGGAGCAG
GCCCAGAGTGGGCCCCAGGGAGGAAGGTCTCAACGTAGGGAGAGGCATAACCGAATGGAAAGAGATAGAA
GGCGCAGAATCCGCATTTGCTGTGATGAGTTGAATCTCTTAGTGCCGTTCTGCAATGCCGAGACTGACAA
GGCCACAACTCTGCAGTGGACCACAGCATTCCTGAAATACATCCAGGAAAGACATGGAGATTCTCTTAAA
AAGGAATTTGAGAGCGTATTTTGCGGTAAAACTGGCCGAAGGCTAAAGCTGACCAGACCGGACTCCTTGG
TGACCTGTCCTGCACAGGGGAGTTTACAGAGCAGCCCCTCGATGGAGATCAAGTGATCGGACTGAACAGG
AATCCTCGGGGGGTGAACAGCCATTCCTTCGTGACCTGTGCACGCCTTCTGCAACCCTGGAGCTCTGCTC
GGCTAGTCTGACTCGAAAAGGGCGTGACTCAAGCTGACGGGACTCCAGTAGGGACTTTGAGAGCACATTT
3' UTR
TGTAAAAATATTTATCTAGACGCAAATGCTTATCCATGAATGTCCTCTTAGACCATTTGGGGATGAAGCC
ATCTTAATAATTAGTAATAATTAATTAGTAATAATTAGTAAGCATTTTCTCAATGCTCTGATTCCATCAT
GTTTTCTTAACATGATAACTTAAAAAATTGACATCCTTTGTACTTTCTTTAATCTTAAAAAGTACACGGC
TTTTTACTTATTTACCTTTTAAATATGCCCCTTTAGCAATTGGAACAAGTTAAATTGTTAACTAAAAACA
GTTTGGAAATTTTATTTCATTCGTTATATCACACCCCCTTGTCATGACTCTGAGTCACGTGCTGCTGTAT
TGCAACGTGCAGGACCATTTTAAACCTGTGTGCTAAAAATTTTCCAGATACTTGCTTTAAAGCTACTTTT
GTCCACAAATGAAATACTGTCACAGTAGACGCTTAAATGCCACGTTTTCATACCAAGAGTCATTCATTAC
TTCATGTGTCACAAACTGTGGTGTTTGGAATTGGGTTTTTCAATGAGTGGCTTTACTTATCAATCACAAC
AGGTAATAGCAATAGACGTTAGTGCAATACAAAGTCACCCTCAATAAATACTGTTAATTGGAGATGTGAG
TTTGTACACAAAACATCAGACTAGACCTTTGTATGGGAGAGAATTTACTGTACATTAAATTCTTTATTTT
TTGTTAAAAAAAAAAAAAAAAA
Alternancia de exones
E1
E2
E3
E4
E5
E8

Fig. 3

```
TCFL5       ATGTCGGGCC CCGGACCGCG GGAGCCGCCG CCGGAGGCAG GCGCGGCAGG
CHA         ---------- ---------- ---------- ---------- ----------

            CGGCGAGGCG GCCGTCGAGG GCGCGGGCGG CGGGGACGCG GCGCTGGGCG
            ---------- ---------- ---------- ---------- ----------

            AGCCGGGGCT GAGCTTCACG ACCACCGACC TGAGCCTGGT GGAGATGACG
            ---------- ---------- ---------- ---------- ----------

            GAGGTGGAGT ACACGCAGCT GCAGCACATC CTCTGCTCGC ACATGGAGGC
            ---------- ---------- ---------- ---------- ----------

            GGCGGCTGAC GGCGAGCTCG AGACGCGCCT CAACTCGGCG CTGCTGGCGG
            ---------- ---------- ---------- ---------- ----------

            CGGCGGGCCC GGGCGCAGGC GCGGGCGGCT TCGCGGCGGG CGGTCAGGGG
            ---------- ---------- ---------- ---------- ----------

            GGCGCGGCGC CCGTGTACCC CGTGCTGTGC CCGTCCGCGC TGGCGGCCGA
            ---------- ---------- ---------- ---------- ----------

            CGCGCCCTGC CTGGGCCACA TCGACTTCCA GGAGCTGCGC ATGATGCTGC
            ---------- ---------- ---------- ---------- ----------

            TAAGCGAGGC GGGCGCGGCG GAGAAGACGT CGGGCGGCGG GGACGGAGCG
            ---------- ---------- ---------- ---------- ----------

            AGGGCCCGGG CCGACGGCGC CGCCAAGGAG GGCGCGGGCG CGGCTGCGGC
            ---------- ---------- ---------- ---------- ----------

            TGCGGCTGGA CCCGACGGCG CGCCCGAGGC CCGGGCCAAG CCGGCCGTGC
            ---------- ---------- ---------- ---------- ----------

            GCGTCCGCCT GGAGGACCGC TTCAACAGCA TCCCCGCCGA GCCGCCGCCC
            ---------- ---------- ---------- ---------- ----------

            GCGCCGCGCG GCCCCGAGCC CCCCGAGCCG GGCGGGGCGC TCAACAATTT
            ---------- ---------- ---------- ---------- ----------

            GGTAACTCTC ATTCGACATC CATCTGAACT AATGAATGTT CCTCTTCAGC
            ---------- ---------- ---------- -ATGAATGTT CCTCTTCAGC

            AACAAAACAA ATGTACAGCA TTAGTGAAAA ATAAAACTGC GGCTACAACT
            AACAAAACAA ATGTACAGCA TTAGTGAAAA ATAAAACTGC GGCTACAACT

            ACTGCTTTGC AATTTACATA CCCACTGTTT ACTACAAATG CTTGCTCTAC
            ACTGCTTTGC AATTTACATA CCCACTGTTT ACTACAAATG CTTGCTCTAC

            TAGTGGAAAT TCTAATCTTT CACAGACACA GAGTTCTAGT AACTCATGTT
            TAGTGGAAAT TCTAATCTTT CACAGACACA GAGTTCTAGT AACTCATGTT

            CTGTACTTGA AGCTGCCAAG CACCAGGATA TTGGATTGCC TAGAGCATTT
            CTGTACTTGA AGCTGCCAAG CACCAGGATA TTGGATTGCC TAGAGCATTT

            TCTTTCTGTT ATCAGCAAGA AATTGAATCC ACTAAACAGA CGTTAGGTAG
            TCTTTCTGTT ATCAGCAAGA AATTGAATCC ACTAAACAGA CGTTAGGTAG

            TAGAAACAAA GTTTTGCCTG AGCAAGTTTG GATTAAAGTG GGAGAAGCAG
            TAGAAACAAA GTTTTGCCTG AGCAAGTTTG GATTAAAGTG GGAGAAGCAG

            CGCTATGCAA ACAAGCACTG AAGAGGAATC GGAGTAGAAT GCGTCAGTTG
```

Fig. 4

```
CGCTATGCAA ACAAGCACTG AAGAGGAATC GGAGTAGAAT GCGTCAGTTG

GACACAAATG TAGAGCGAAG AGCCCTTGGA GAGATTCAGA ATGTGGGCGA
GACACAAATG TAGAGCGAAG AGCCCTTGGA GAGATTCAGA ATGTGGGCGA

AGGTGCCACC GCCACACAAG GCGCTTGGCA GTCCTCGGAG TCCTCACAGG
AGGTGCCACC GCCACACAAG GCGCTTGGCA GTCCTCGGAG TCCTCACAGG

CAAACCTGGG GGAGCAGGCC CAGAGTGGGC CCCAGGGAGG AAGGTCTCAA
CAAACCTGGG GGAGCAGGCC CAGAGTGGGC CCCAGGGAGG AAGGTCTCAG

CGTAGGGAGA GGCATAACCG AATGGAAAGA GATAGAAGGC GCAGAATCCG
CGTAGGGAGA GGCATAACCG AATGGAAAGA GATAGAAGGC GCAGAATCCG

CATTTGCTGT GATGAGTTGA ATCTCTTAGT GCCGTTCTGC AATGCCGAGA
CATTTGCTGT GATGAGTTGA ATCTCTTAGT GCCGTTCTGC AATGCCGAGA

CTGACAAGGC CACAACTCTG CAGTGGACCA CAGCATTCCT GAAATACATC
CTGACAAGGC CACAACTCTG CAGTGGACCA CAGCATTCCT GAAATACATC

CAGGAAAGAC ATGGAGATTC TCTTAAAAAG GAATTTGAGA GCGTATTTTG
CAGGAAAGAC ATGGAGATTC TCTTAAAAAG GAATTTGAGA GCGTATTTTG

CGGTAAAACT GGCCGAAGGC TAAAGCTGAC CAGACCGGAC TCCTTGGTGA
CGGTAAAACT GGCCGAAGGC TAAAGCTGAC CAGACCGGAC TCCTTGGTGA

CCTGTCCTGC ACAGGGGAGT TTACAGAGCA GCCCCTCGAT GGAGATCAAG
CCTGTCCTGC ACAGGGGAGT TTACAGAGCA GCCCCTCGAT GGAGATCAAG

TGA
---
```

Fig. 4 (cont.)

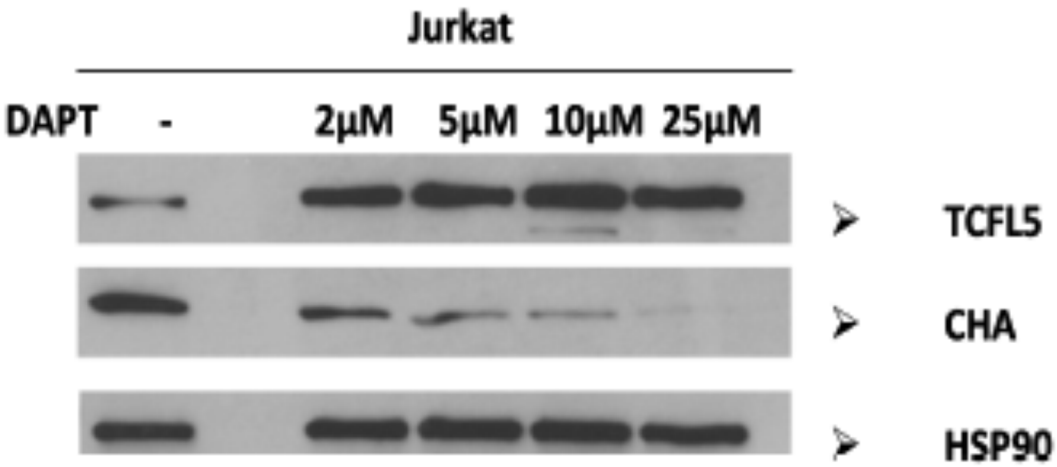


Fig. 5

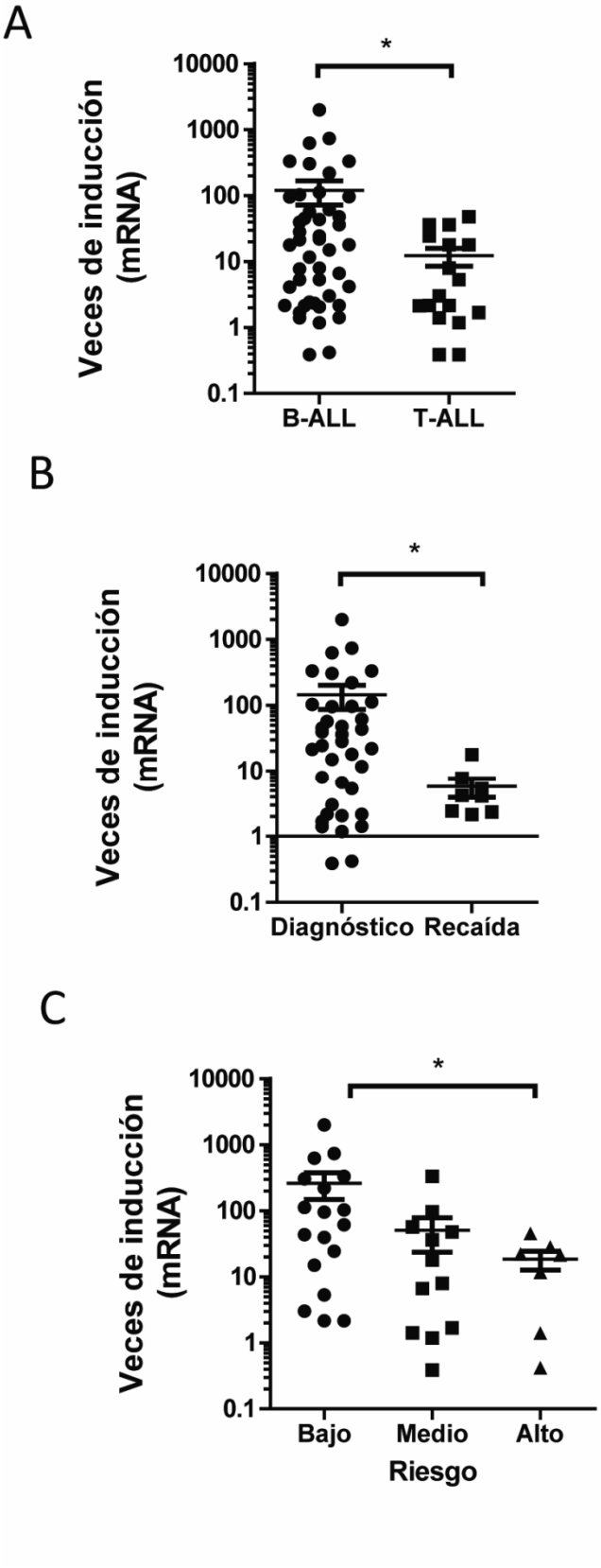
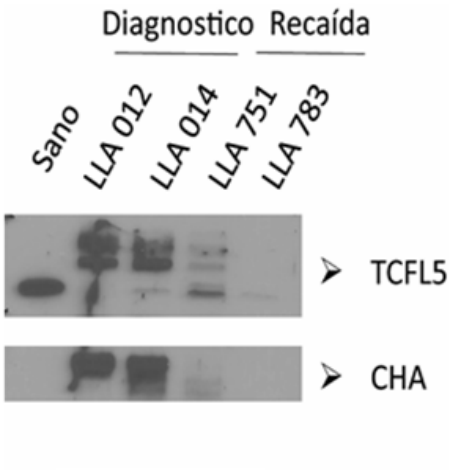

A
*
10000
1000
100
10
1
0.1
Veces de inducción (mRNA)
B-ALL
T-ALL
B
*
10000
1000
100
10
1
0.1
Veces de inducción (mRNA)
Diagnóstico
Recaída
C
*
10000
1000
100
10
1
0.1
Veces de inducción (mRNA)
Bajo
Medio
Alto
Riesgo

Fig. 6

A


Diagnostico
Recaída
Sano
LLA 012
LLA 014
LLA 751
LLA 783
TCFL5
CHA

B

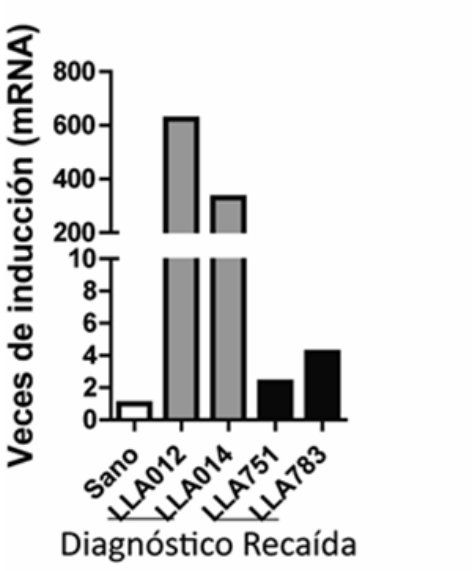

Veces de inducción (mRNA)
800
600
400
200
10
8
6
4
2
0
Sano
LLA012
LLA014
LLA751
LLA783
Diagnóstico Recaída

Fig. 7

LISTA DE SECUENCIAS

<110> Universidad Autónoma de Madrid
Fundación Investigación Biomédica Hospital Infantil Universitario Niño Jesús

<120> USO DE TCFL5/CHA COMO NUEVO MARCADOR PARA EL PRONÓSTICO Y/O DIAGNÓSTICO DIFERENCIAL DE LEUCEMIAS LINFOBLÁSTICAS AGUDAS

<130> 900 949

<160> 74

<170> PatentIn version 3.5

<210> 1
<211> 2575
<212> DNA
<213> Homo sapiens

<220>
<221> mRNA
<222> (1)..(2575)

<400> 1
cgggccgggg gtcgcctccc gcctcccgcc tcccgcctcc cgccgcccgc cgcccgcgtc 60

gccctcgccg ccgttgggcc gcgccgcgcc gccatgtcgg gccccggacc gcgggagccg 120

ccgccggagg caggcgcggc aggcggcgag gcggccgtcg agggcgcggg cggcggggac 180

gcggcgctgg gcgagccggg gctgagcttc acgaccaccg acctgagcct ggtggagatg 240

acggaggtgg agtacacgca gctgcagcac atcctctgct cgcacatgga ggcggcggct 300

gacggcgagc tcgagacgcg cctcaactcg gcgctgctgg cggcggcggg cccgggcgca 360

ggcgcgggcg gcttcgcggc gggcggtcag gggggcgcgg cgcccgtgta ccccgtgctg 420

tgcccgtccg cgctggcggc cgacgcgccc tgcctgggcc acatcgactt ccaggagctg 480

cgcatgatgc tgctaagcga ggcgggcgcg gcggagaaga cgtcgggcgg cggggacgga 540

gcgagggccc gggccgacgg cgccgccaag gagggcgcgg gcgcggctgc ggctgcggct 600

ggacccgacg gcgcgcccga ggcccgggcc aagccggccg tgcgcgtccg cctggaggac 660

cgcttcaaca gcatccccgc cgagccgccg cccgcgccgc gcggccccga gccccccgag 720

ccgggcgggg cgctcaacaa tttggtaact ctcattcgac atccatctga actaatgaat 780

gttcctcttc agcaacaaaa caaatgtaca gcattagtga aaaataaaac tgcggctaca 840

actactgctt tgcaatttac atacccactg tttactacaa atgcttgctc tactagtgga 900

aattctaatc tttcacagac acagagttct agtaactcat gttctgtact tgaagctgcc 960

aagcaccagg atattggatt gcctagagca ttttctttct gttatcagca agaaattgaa 1020

tccactaaac agacgttagg tagtagaaac aaagttttgc ctgagcaagt ttggattaaa 1080

gtgggagaag cagcgctatg caaacaagca ctgaagagga atcggagtag aatgcgtcag 1140

ttggacacaa atgtagagcg aagagcccтt ggagagattc agaatgtggg cgaaggtgcc 1200

accgccacac aaggcgcttg gcagtcctcg gagtcctcac aggcaaacct gggggagcag 1260

gcccagagtg ggccccaggg aggaaggtct caacgtaggg agaggcataa ccgaatggaa 1320

agagatagaa ggcgcagaat ccgcatttgc tgtgatgagt tgaatctctt agtgccgttc 1380

tgcaatgccg agactgacaa ggccacaact ctgcagtgga ccacagcatt cctgaaatac 1440

atccaggaaa gacatggaga ttctcttaaa aaggaatttg agagcgtatt ttgcggtaaa 1500

actggccgaa ggctaaagct gaccagaccg gactccttgg tgacctgtcc tgcacagggg 1560

agtttacaga gcagcccctc gatggagatc aagtgatcgg actgaacagg aatcctcggg 1620

gggtgaacag ccattccttc gtgacctgtg cacgccttct gcaaccctgg agctctgctc 1680

ggctagtctg actcgaaaag ggcgtgactc aagctgacgg gactccagta gggactttga 1740

gagcacattt tgtaaaaata tttatctaga cgcaaatgct tatccatgaa tgtcctctta 1800

gaccatttgg ggatgaagcc atcttaataa ttagtaataa ttaattagta ataattagta 1860

agcattttct caatgctctg attccatcat gttttcttaa catgataact taaaaaattg 1920

acatcctttg tactttcttt aatcttaaaa agtacacggc tttttactta tttacctttt 1980

aaatatgccc ctttagcaat tggaacaagt taaattgtta actaaaaaca gtttggaaat 2040

tttatttcat tcgttatatc acacccccctt gtcatgactc tgagtcacgt gctgctgtat 2100

tgcaacgtgc aggaccattt taaacctgtg tgctaaaaat tttccagata cttgctttaa 2160

agctactttt gtccacaaat gaaatactgt cacagtagac gcttaaatgc cacgttttca 2220

taccaagagt cattcattac ttcatgtgtc acaaactgtg gtgtttggaa ttgggttttt 2280

caatgagtgg ctttacttat caatcacaac aggtaatagc aatagacgtt agtgcaatac 2340

aaagtcaccc tcaataaata ctgttaattg gagatgtgag tttgtacaca aaacatcaga 2400

ctagaccttt gtatgggaga gaatttactg tacattaaat tctttatttt ttgttactct 2460

gtagccagtc attaatctga aggtttaata tatcatttta ttgggatgag atcatagtct 2520

ttacacaaat gctatgtaaa caagttactg aatatttaaa aaaaaaaaaa aaaaa 2575

<210> 2
<211> 1458
<212> DNA
<213> Homo sapiens

<220>
<221> mRNA
<222> (1)..(1458)

<400> 2
gggaaccatg tgctccgggt ggcattcttg gatgtcaggg aagcctagga aggcccgacg 60

ggggtttgca ggctttgagc tctaaaatga ggctccagcc tgcagccgcc cccttggcga 120

ccctcaggga tggctgtgga aggggctgcg tccgcagttt tctgctcatc ccagcctgta 180

gttgcccatc gtggtcagac tggaggatgg caaaagtgag attggtgaca tgatgtgatg 240

agagcaggga ggtggcgcag ttttgctttt cctccatagg aagaggagga cttccacagc 300

agcgatagtt gctaattgtt gtgtagagat ttccaggatg agagttgggt tgctgtctgg 360

aaaaactagc cagagccatg ggagtgggat ggtggatgct tgtctgccac gtattgctga 420

cacttctggg ctttcttttg tcttggcagt ttggtaactc tcattcgact tccatctgaa 480

ctaatgaatg ttcctcttca gcaacaaaac aaatgtacag cattagtgaa aaataaaact 540

gcggctacaa ctactgcttt gcaatttaca tacccactgt ttactacaaa tgcttgctct 600

actagtggaa attctaatct ttcacagaca cagagttcta gtaactcatg ttctgtactt 660

gaagctgcca agcaccagga tattggattg cctagagcat tttctttctg ttatcagcaa 720

gaaattgaat ccactaaaca gacgttaggt agtagaaaca aagttttgcc tgagcaagtt 780

tggattaaag tgggagaagc agcgctatgc aaacaagcac tgaagaggaa tcggagtaga 840

atgcgtcagt tggacacaaa tgtagagcga agagcccttg gagagattca gaatgtgggc 900

gaaggtgcca ccgccacaca aggcgcttgg cagtcctcgg agtcctcaca ggcaaacctg 960

ggggagcagg cccagagtgg gccccaggga ggaaggtctc agcgtaggga gaggcataac 1020

cgaatggaaa gagatagaag gcgcagaatc cgcatttgct gtgatgagtt gaatctctta 1080

gtgccgttct gcaatgccga gactgacaag gccacaactc tgcagtggac cacagcattc 1140

ctgaaataca tccaggaaag acatggagat tctcttaaaa aggaatttga gagcgtattt 1200

tgcggtaaaa ctggccgaag gctaaagctg accagaccgg actccttggt gacctgtcct 1260

gcacagggga gtttacagag cagcccctcg atggagatca agtgatcgga ctgaacagga 1320

atcctcgggg ggtgaacagc cattccttcg tgacctgtgc acgccttctg caaccctgga 1380

gctctgctcg gctagtctga ctcgaaaagg gcgtgactct tcgactgccc tgaggtcatc 1440

cctgaaactc gagcacgg 1458

<210> 3
<211> 500
<212> PRT
<213> Homo sapiens

<400> 3

Met Ser Gly Pro Gly Pro Arg Glu Pro Pro Pro Glu Ala Gly Ala Ala
1 5 10 15

Gly Gly Glu Ala Ala Val Glu Gly Ala Gly Gly Gly Asp Ala Ala Leu
20 25 30

Gly Glu Pro Gly Leu Ser Phe Thr Thr Thr Asp Leu Ser Leu Val Glu
35 40 45

Met Thr Glu Val Glu Tyr Thr Gln Leu Gln His Ile Leu Cys Ser His
50 55 60

Met Glu Ala Ala Ala Asp Gly Glu Leu Glu Thr Arg Leu Asn Ser Ala
65 70 75 80

Leu Leu Ala Ala Ala Gly Pro Gly Ala Gly Ala Gly Gly Phe Ala Ala
85 90 95

Gly Gly Gln Gly Gly Ala Ala Pro Val Tyr Pro Val Leu Cys Pro Ser

```
            100                 105                 110

Ala Leu Ala Ala Asp Ala Pro Cys Leu Gly His Ile Asp Phe Gln Glu
        115                 120                 125

Leu Arg Met Met Leu Leu Ser Glu Ala Gly Ala Ala Glu Lys Thr Ser
    130                 135                 140

Gly Gly Gly Asp Gly Ala Arg Ala Arg Ala Asp Gly Ala Ala Lys Glu
145                 150                 155                 160

Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Pro Asp Gly Ala Pro Glu
                165                 170                 175

Ala Arg Ala Lys Pro Ala Val Arg Val Arg Leu Glu Asp Arg Phe Asn
            180                 185                 190

Ser Ile Pro Ala Glu Pro Pro Pro Ala Pro Arg Gly Pro Glu Pro Pro
        195                 200                 205

Glu Pro Gly Gly Ala Leu Asn Asn Leu Val Thr Leu Ile Arg His Pro
    210                 215                 220

Ser Glu Leu Met Asn Val Pro Leu Gln Gln Gln Asn Lys Cys Thr Ala
225                 230                 235                 240

Leu Val Lys Asn Lys Thr Ala Ala Thr Thr Thr Ala Leu Gln Phe Thr
                245                 250                 255

Tyr Pro Leu Phe Thr Thr Asn Ala Cys Ser Thr Ser Gly Asn Ser Asn
            260                 265                 270

Leu Ser Gln Thr Gln Ser Ser Ser Asn Ser Cys Ser Val Leu Glu Ala
        275                 280                 285

Ala Lys His Gln Asp Ile Gly Leu Pro Arg Ala Phe Ser Phe Cys Tyr
    290                 295                 300

Gln Gln Glu Ile Glu Ser Thr Lys Gln Thr Leu Gly Ser Arg Asn Lys
305                 310                 315                 320

Val Leu Pro Glu Gln Val Trp Ile Lys Val Gly Glu Ala Ala Leu Cys
                325                 330                 335

Lys Gln Ala Leu Lys Arg Asn Arg Ser Arg Met Arg Gln Leu Asp Thr
            340                 345                 350

Asn Val Glu Arg Arg Ala Leu Gly Glu Ile Gln Asn Val Gly Glu Gly
        355                 360                 365

Ala Thr Ala Thr Gln Gly Ala Trp Gln Ser Ser Glu Ser Ser Gln Ala
    370                 375                 380
```

```
Asn Leu Gly Glu Gln Ala Gln Ser Gly Pro Gln Gly Gly Arg Ser Gln
385                 390                 395                 400

Arg Arg Glu Arg His Asn Arg Met Glu Arg Asp Arg Arg Arg Arg Ile
                405                 410                 415

Arg Ile Cys Cys Asp Glu Leu Asn Leu Leu Val Pro Phe Cys Asn Ala
            420                 425                 430

Glu Thr Asp Lys Ala Thr Thr Leu Gln Trp Thr Thr Ala Phe Leu Lys
        435                 440                 445

Tyr Ile Gln Glu Arg His Gly Asp Ser Leu Lys Lys Glu Phe Glu Ser
    450                 455                 460

Val Phe Cys Gly Lys Thr Gly Arg Arg Leu Lys Leu Thr Arg Pro Asp
465                 470                 475                 480

Ser Leu Val Thr Cys Pro Ala Gln Gly Ser Leu Gln Ser Ser Pro Ser
                485                 490                 495

Met Glu Ile Lys
            500
```

<210> 4
<211> 273
<212> PRT
<213> Homo sapiens

<400> 4

```
Met Asn Val Pro Leu Gln Gln Gln Asn Lys Cys Thr Ala Leu Val Lys
1               5                   10                  15

Asn Lys Thr Ala Ala Thr Thr Thr Ala Leu Gln Phe Thr Tyr Pro Leu
            20                  25                  30

Phe Thr Thr Asn Ala Cys Ser Thr Ser Gly Asn Ser Asn Leu Ser Gln
        35                  40                  45

Thr Gln Ser Ser Ser Asn Ser Cys Ser Val Leu Glu Ala Ala Lys His
    50                  55                  60

Gln Asp Ile Gly Leu Pro Arg Ala Phe Ser Phe Cys Tyr Gln Gln Glu
65                  70                  75                  80

Ile Glu Ser Thr Lys Gln Thr Leu Gly Ser Arg Asn Lys Val Leu Pro
                85                  90                  95

Glu Gln Val Trp Ile Lys Val Gly Glu Ala Ala Leu Cys Lys Gln Ala
            100                 105                 110
```

Leu Lys Arg Asn Arg Ser Arg Met Arg Gln Leu Asp Thr Asn Val Glu
115 120 125

Arg Arg Ala Leu Gly Glu Ile Gln Asn Val Gly Glu Gly Ala Thr Ala
130 135 140

Thr Gln Gly Ala Trp Gln Ser Ser Glu Ser Ser Gln Ala Asn Leu Gly
145 150 155 160

Glu Gln Ala Gln Ser Gly Pro Gln Gly Gly Arg Ser Gln Arg Arg Glu
165 170 175

Arg His Asn Arg Met Glu Arg Asp Arg Arg Arg Arg Ile Arg Ile Cys
180 185 190

Cys Asp Glu Leu Asn Leu Leu Val Pro Phe Cys Asn Ala Glu Thr Asp
195 200 205

Lys Ala Thr Thr Leu Gln Trp Thr Thr Ala Phe Leu Lys Tyr Ile Gln
210 215 220

Glu Arg His Gly Asp Ser Leu Lys Lys Glu Phe Glu Ser Val Phe Cys
225 230 235 240

Gly Lys Thr Gly Arg Arg Leu Lys Leu Thr Arg Pro Asp Ser Leu Val
245 250 255

Thr Cys Pro Ala Gln Gly Ser Leu Gln Ser Ser Pro Ser Met Glu Ile
260 265 270

Lys

<210> 5
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_01 Forward

<220>
<221> misc_feature
<222> (1)..(20)

<400> 5
tgcctgagca agtttggatt 20

<210> 6
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_01 Reverse

<220>
<221> misc_feature
<222> (1)..(20)

<400> 6
cgcattctac tccgattcct 20

<210> 7
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_02 Forward

<220>
<221> misc_feature
<222> (1)..(20)

<400> 7
gagtcctcac aggcaaacct 20

<210> 8
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_02 Reverse

<220>
<221> misc_feature
<222> (1)..(23)

<400> 8
gattcaactc atcacagcaa atg 23

<210> 9
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_03 Forward

<220>
<221> misc_feature
<222> (1)..(21)

<400> 9
caacgtaggg agaggcataa c 21

<210> 10
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_03 Reverse

<220>
<221> misc_feature
<222> (1)..(22)

<400> 10
cagaacggca ctaagagatt ca 22

<210> 11
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_04 Forward

<220>
<221> misc_feature
<222> (1)..(23)

<400> 11
gcctgagcaa gtttggatta aag 23

<210> 12
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_04 Reverse

<220>
<221> misc_feature
<222> (1)..(21)

<400> 12
gtgtccaact gacgcattct a 21

<210> 13
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_05 Forward

<220>
<221> misc_feature
<222> (1)..(26)

<400> 13
cactaaacag acgttaggta gtagaa 26

<210> 14
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_05 Reverse

<220>

<221> misc_feature
<222> (1)..(23)

<400> 14
gattcctctt cagtgcttgt ttg 23

<210> 15
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_06 Forward

<220>
<221> misc_feature
<222> (1)..(20)

<400> 15
aggcataacc gaatggaaag 20

<210> 16
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_06 Reverse

<220>
<221> misc_feature
<222> (1)..(20)

<400> 16
ttgcagaacg gcactaagag 20

<210> 17
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_07 Forward

<220>
<221> misc_feature
<222> (1)..(19)

<400> 17
tgcctgagca agtttggat 19

<210> 18
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_07 Reverse

<220>
<221> misc_feature

<222> (1)..(20)

<400> 18
acgcattcta ctccgattcc 20

<210> 19
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_08 Forward

<220>
<221> misc_feature
<222> (1)..(25)

<400> 19
agaaattgaa tccactaaac agacg 25

<210> 20
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_08 Reverse

<220>
<221> misc_feature
<222> (1)..(21)

<400> 20
ccaactgacg cattctactc c 21

<210> 21
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_09 Forward

<220>
<221> misc_feature
<222> (1)..(20)

<400> 21
tcaacgtagg gagaggcata 20

<210> 22
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_09 Reverse

<220>
<221> misc_feature
<222> (1)..(22)

<400> 22
acggcactaa gagattcaac tc 22

<210> 23
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_10 Forward

<220>
<221> misc_feature
<222> (1)..(25)

<400> 23
tgcaatttac atacccactg tttac 25

<210> 24
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_10 Reverse

<220>
<221> misc_feature
<222> (1)..(23)

<400> 24
ctctaggcaa tccaatatcc tgg 23

<210> 25
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_11 Forward

<220>
<221> misc_feature
<222> (1)..(22)

<400> 25
tgaatctctt agtgccgttc tg 22

<210> 26
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_11 Reverse

<220>
<221> misc_feature
<222> (1)..(23)

<400> 26
ctccatgtct ttcctggatg tat 23

<210> 27
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_12 Forward

<220>
<221> misc_feature
<222> (1)..(20)

<400> 27
gagcccttgg agagattcag 20

<210> 28
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_12 Reverse

<220>
<221> misc_feature
<222> (1)..(20)

<400> 28
aggtgccacc gccacacaag 20

<210> 29
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_13 Forward

<220>
<221> misc_feature
<222> (1)..(20)

<400> 29
ctcttagtgc cgttctgcaa 20

<210> 30
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_13 Reverse

<220>
<221> misc_feature
<222> (1)..(20)

<400> 30

ctgtggtcca ctgcagagtt 20

<210> 31
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_14 Forward

<220>
<221> misc_feature
<222> (1)..(20)

<400> 31
gggaggaagg tctcaacgta 20

<210> 32
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_14 Reverse

<220>
<221> misc_feature
<222> (1)..(20)

<400> 32
ggattctgcg ccttctatct 20

<210> 33
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_15 Forward

<220>
<221> misc_feature
<222> (1)..(20)

<400> 33
gagctctgct cggctagtct 20

<210> 34
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_15 Reverse

<220>
<221> misc_feature
<222> (1)..(20)

<400> 34
gagctctgct cggctagtct 20

<210> 35
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_16 Forward

<220>
<221> misc_feature
<222> (1)..(20)

<400> 35
gcaaacaagc actgaagagg 20

<210> 36
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_16 Reverse

<220>
<221> misc_feature
<222> (1)..(20)

<400> 36
gggctcttcg ctctacattt 20

<210> 37
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_17 Forward

<220>
<221> misc_feature
<222> (1)..(19)

<400> 37
tgcctgagca agtttggat 19

<210> 38
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_17 Reverse

<220>
<221> misc_feature
<222> (1)..(20)

<400> 38
acgcattcta ctccgattcc 20

<210> 39
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_18 Forward

<220>
<221> misc_feature
<222> (1)..(20)

<400> 39
agagcccttg gagagattca 20

<210> 40
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_18 Reverse

<220>
<221> misc_feature
<222> (1)..(20)

<400> 40
cctacgctga gaccttcctc 20

<210> 41
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_19 Forward

<220>
<221> misc_feature
<222> (1)..(20)

<400> 41
aaggtctcag cgtagggaga 20

<210> 42
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> TodoTCFL5_19 Reverse

<220>
<221> misc_feature
<222> (1)..(20)

<400> 42
ttgcagaacg gcactaagag 20

<210> 43
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> TCFL5noCHA_01 Forward

<220>
<221> misc_feature
<222> (1)..(17)

<400> 43
ttcaacagca tccccgc 17

<210> 44
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> TCFL5noCHA_01 Reverse

<220>
<221> misc_feature
<222> (1)..(22)

<400> 44
gttgctgaag aggaacattc at 22

<210> 45
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> TCFL5/CHA_01 Forward

<220>
<221> misc_feature
<222> (1)..(21)

<400> 45
gagactgaca aggccacaac t 21

<210> 46
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> TCFL5/CHA_01 Reverse

<220>
<221> misc_feature
<222> (1)..(19)

<400> 46
ccgcaaaata cgctctcaa 19

<210> 47

<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> TCFL5/CHA_02 Forward

<220>
<221> misc_feature
<222> (1)..(20)

<400> 47
ctgacaaggc cacaactctg 20

<210> 48
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> TCFL5/CHA_02 Reverse

<220>
<221> misc_feature
<222> (1)..(19)

<400> 48
ctttagcctt cggccagtt 19

<210> 49
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> TCFL5/CHA_03 Forward

<220>
<221> misc_feature
<222> (1)..(20)

<400> 49
atctcttagt gccgttctgc 20

<210> 50
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> TCFL5/CHA_03 Reverse

<220>
<221> misc_feature
<222> (1)..(23)

<400> 50
ccgcaaaata cgctctcaaa ttc 23

<210> 51
<211> 20

<212> DNA
<213> Artificial Sequence

<220>
<223> TCFL5/CHA_04 Forward

<220>
<221> misc_feature
<222> (1)..(20)

<400> 51
gtgatcggac tgaacaggaa 20

<210> 52
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> TCFL5/CHA_04 Reverse

<220>
<221> misc_feature
<222> (1)..(20)

<400> 52
cgagtttcag ggatgacctc 20

<210> 53
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> TCFL5/CHA_05 Forward

<220>
<221> misc_feature
<222> (1)..(20)

<400> 53
ctgacaaggc cacaactctg 20

<210> 54
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> TCFL5/CHA_05 Reverse

<220>
<221> misc_feature
<222> (1)..(19)

<400> 54
ctttagcctt cggccagtt 19

<210> 55
<211> 21
<212> DNA

<213> Artificial Sequence

<220>
<223> TCFL5/CHA_06 Forward

<220>
<221> misc_feature
<222> (1)..(21)

<400> 55
gtcccgtcag cttgagtcac g 21

<210> 56
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> TCFL5/CHA_06 Reverse

<220>
<221> misc_feature
<222> (1)..(22)

<400> 56
ctgcagtgga ccacagcatt cc 22

<210> 57
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> TCFL5/CHA_07 Forward

<220>
<221> misc_feature
<222> (1)..(19)

<400> 57
aaactggccg aaggctaaa 19

<210> 58
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> TCFL5/CHA_07 Reverse

<220>
<221> misc_feature
<222> (1)..(22)

<400> 58
gtccgatcac ttgatctcca tc 22

<210> 59
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> TCFL5/CHA_08 Forward

<220>
<221> misc_feature
<222> (1)..(20)

<400> 59
gtgatcggac tgaacaggaa 20

<210> 60
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> TCFL5/CHA_08 Reverse

<220>
<221> misc_feature
<222> (1)..(20)

<400> 60
tactggagtc ccgtcagctt 20

<210> 61
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> TCFL5/CHA_9 Forward

<220>
<221> misc_feature
<222> (1)..(22)

<400> 61
aatttgagag cgtattttgc gg 22

<210> 62
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> TCFL5/CHA_9 Reverse

<220>
<221> misc_feature
<222> (1)..(21)

<400> 62
ggattcctgt tcagtccgat c 21

<210> 63
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> SoloCHA_01 Forward

<220>
<221> misc_feature
<222> (1)..(20)

<400> 63
ggaagaggag gacttccaca 20

<210> 64
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> SoloCHA_01 Reverse

<220>
<221> misc_feature
<222> (1)..(20)

<400> 64
agcaatacgt ggcagacaag 20

<210> 65
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> SoloCHA_02 Forward

<220>
<221> misc_feature
<222> (1)..(20)

<400> 65
gctcatccca gcctgtagtt 20

<210> 66
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> SoloCHA_02 Reverse

<220>
<221> misc_feature
<222> (1)..(20)

<400> 66
tgtggaagtc ctcctcttcc 20

<210> 67
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> SoloCHA_03 Forward

<220>
<221> misc_feature
<222> (1)..(20)

<400> 67
cattcttgga tgtcagggaa 20

<210> 68
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> SoloCHA_03 Reverse

<220>
<221> misc_feature
<222> (1)..(20)

<400> 68
tagagctcaa agcctgcaaa 20

<210> 69
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> SoloCHA_04 Forward

<220>
<221> misc_feature
<222> (1)..(21)

<400> 69
gagagttggg ttgctgtctg g 21

<210> 70
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> SoloCHA_04 Reverse

<220>
<221> misc_feature
<222> (1)..(24)

<400> 70
gttcagatgg atgtcgaatg agag 24

<210> 71
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Beta-actina_F

<220>
<221> misc_feature
<222> (1)..(21)

<400> 71
catgtacgtt gctatccagg c 21

<210> 72
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Beta-actina_R

<220>
<221> misc_feature
<222> (1)..(21)

<400> 72
ctccttaatg tcacgcacga t 21

<210> 73
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> HPRT- Forward

<220>
<221> misc_feature
<222> (1)..(24)

<400> 73
ctggaaagaa tgtcttgatt gtgg 24

<210> 74
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> HPRT-Reverse

<220>
<221> misc_feature
<222> (1)..(23)

<400> 74
catcttggat tatactgcct gac 23