

UNIVERSIDAD AUTÓNOMA DE MADRID

FACULTAD DE MEDICINA

Departamento de Medicina

**BASES MOLECULARES DE LA ASOCIACIÓN DEL
MICROSATÉLITE
DEL GEN CD40LG A AUTOINMUNIDAD**

Ignacio Losada Fernández

TESIS DOCTORAL

Madrid, 2015

ÍNDICE

INTRODUCCIÓN.....	1
· CD40 Ligando	1
· Regulación de la expresión de CD40LG.....	6
· Regulación de la transcripción de CD40LG.....	7
· Regulación posttranscripcional de CD40LG.....	8
· CD40LG y autoinmunidad.....	12
· Microsatélite de CD40LG. Asociación genética a LES y AR.....	13
OBJETIVOS.....	16
MATERIALES Y MÉTODOS.....	16
Pacientes y controles.....	17
· Series de pacientes	17
· Series de controles.....	18
Materiales	18
· Cepas bacterianas de <i>E. coli</i>	18
· Líneas celulares	19
· Vectores.....	19
Métodos	21
· Cultivos de bacterias.....	21
· Transformación de bacterias con DNA plasmídico.....	21
· Cultivo de líneas celulares.....	21
· Aislamiento de DNA genómico.....	22
· Aislamiento de DNA plasmídico.....	22
· Aislamiento de RNA.....	23
· Síntesis de cDNA mediante retrotranscripción (RT)	23
· Amplificación génica mediante la reacción en cadena de la polimerasa (PCR).....	23
· Análisis de polimorfismos.....	25
· Clonaje de la región 3'UTR de CD40LG	25
· Secuenciación de DNA plasmídico.....	26
· Transfección y obtención de transfectantes estables	26
· Cinéticas de caída del mRNA	27
· Medida de la cantidad de mRNA en las cinéticas de caída	27
Herramientas informáticas.....	27

Desarrollo de un cálculo para estimar la penetrancia y el riesgo relativo para un genotipo a partir de datos de un estudio de casos y controles.....	30
RESULTADOS.....	36
· Asociación del MSCD40LG a AR y LES.....	36
· Series iniciales de pacientes y controles	36
· Nuevas series de sanos y pacientes de LES.....	38
· MSCD40LG y estabilidad del mRNA.....	45
· Clonaje de variantes alélicas comunes del MSCD40LG en vectores de expresión	46
· Obtención de transfectantes estables de las construcciones GFP-3UTR.CAn	54
· Cinéticas de caída de los mRNA GFP-3UTR.CAn.....	57
· Comparación de la estabilidad del mRNA en función del alelo del MSCD40LG	60
· MSCD40LG, estructura secundaria del mRNA y estabilidad del mRNA	63
· MSCD40LG y procesamiento del pre-mRNA	68
DISCUSIÓN.....	70
CONCLUSIONES	79
BIBLIOGRAFÍA.....	80

ABREVIATURAS

(CA)n: dinucleótido citosina adenina
ΔG: descenso en energía interna
°C: grado centígrado
μg: microgramo
3'UTR: extremo 3' no traducido
ACR: *American College of Rheumatology*
Act-D: actinomicina D
AIRE: *AutoImmune REgulator*
AKNA: *AT-Hook Transcription Factor*
AMV: virus de la mieloblastosis aviar (*Avian Myeloblastosis Virus*)
AP-1: factor de transcripción proteína activadora 1 (*Activator Protein 1*)
AR: artritis reumatoide
ARE: región rica en AU
AUF-1: *AU-Rich Element RNA Binding Protein 1*
BCL2: proto-oncogen 2 del linfoma de células B (*B-cell lymphoma 2*)
BRASS: *Brigham and Women's Rheumatoid Arthritis Sequential Study*
CAMKIV: kinasa IV dependiente de calcio-calmodulina
cDNA: ácido desoxirribonucleico complementario
c-fos: *FBJ Murine Osteosarcoma Viral Oncogene Homolog*
Cfu: unidades formadoras de colonias
ClaI: enzima de restricción I de *Caryophanon latum L*
CMV: citomegalovirus
Ct: umbral de ciclo
CU-rich: región rica en citosina y uracilo
DAG: diacilglicerol
Dam: deoxyadenosin metilasa
Dcm: DNA citosina metilasa
DNA: ácido desoxirribonucleico
DSC: donantes sanas canarias
DSP: donantes sanas peninsulares
E. coli: *Escherichia coli*
Egr1: factor de transcripción de respuesta rápida al crecimiento 1 (*Early Growth Response 1*)
EIRA : *Epidemiological Investigation of Rheumatoid Arthritis*

ERK1/2 : kinasas 1/2 reguladas por señales extracelulares (*Extracellular-signal Regulated Kinases 1/2*)
F: prueba F de suma de cuadrados extra
g: gramo
GADPH: gliceraldehído-3-fosfato deshidrogenasa
GC: guanina-citosina
GFP: proteína verde fluorescente
GM-CSF: factor de crecimiento de colonias de granulocitos y macrófagos
GWAS: estudio de asociación de genoma completo (*Whole Genome Association Study*)
h: hora
HaeIII: enzima de restricción III de *Haemophilus aegypticus*
HIGM1: síndrome de hiperIgM ligado al cromosoma X
HLA: antígenos de histocompatibilidad
HLA-DR1: alelo 1 del gen de histocompatibilidad DRB1
HLA-DR4 : alelo 4 del gen de histocompatibilidad DRB1
hnRNPL: *heterogeneous nuclear ribonucleoprotein L*
h-num :
IC 95%: intervalo de confianza al 95%
IFN-γ: interferón gamma
IgG: inmunoglobulina G
IgM: inmunoglobulina M
IL: interleucina
Io: ionomicina
IRAK1: kinasa 1 asociada al receptor de IL-1 (*interleukin-1 Receptor-Associated Kinase 1*)
kcal: kilocaloría
KDa: kilodalton
l: litro
LB: Luria Broth
LES: lupus eritematoso sistémico
LFA-3: antígeno 3 de función leucocitaria
log molec: logaritmo del número de moléculas de mRNA
Mac-1: antígeno 1 de macrófagos (*Macrophage Antigen-1*)

MAPK: proteína quinasas activadas por mitógenos
(*Mitogen-Activated Protein Kinases*)

min: minuto

ml: mililitro

mol: mol

MPP: metaloproteasas

mRNA: ácido ribonucleico mensajero

MSCD40LG: microsatélite de CD40L

mTEC: células epiteliales de la médula del timo

N: número

NARAC: *North American Arthritis Consortium*

NFAT: factor nuclear de células T activadas (*Nuclear factor of activated T-cells*)

NFκB: factor nuclear potenciador de las cadenas ligeras kappa de las células B activadas

ng: nanogramo

NHGRI: *National Human Genome Research Institute*

NZB: *New Zealand Black*

NZW: *New Zealand White*

OR: *odds ratio*

pb: par de bases

PBS: tampón fosfato salino

PCR: reacción en cadena de la polimerasa

PHA: fitohemaglutinina

PIP: fosfolípidos de inositol trifosfato

PKC: proteína quinasa C

PMA: forbol-12-miristato acetato

p-num:

poli-A: cola de adeninas

poli-C: tira de citosinas

PTB: *polypyrimidine tract-binding protein*

qPCR: *quantitative PCR*

RE: retículo endoplásmico

RNA: ácido ribonucleico

RNP: ribonucleoproteína

RR: riesgo relativo

RT: retrotranscripción

SDS: dodecil sulfato sódico

SNP: polimorfismo de sustitución de base (*Single Nucleotide polymorphism*)

s-num:

SV40: virus 40 vacuolado del simio

t: prueba t de Student

t^{1/2}: tiempo medio

t0: 0 minutos de tratamiento (tiempo inicial)

t90: 90 minutos de tratamiento

t120: 120 minutos de tratamiento

T4: bacteriófago T4

TB: *Terrific Broth*

TCR: receptor de antígeno de linfocitos T

Th: linfocitos T cooperadores

Tm: temperatura de *melting*

TNF: factor de necrosis tumoral

TRAP: proteína de activación relacionada con el

TNF: (*TNF-Related Activation Protein*)

TTP: tristetraprolina

U: unidad

VEGF: factor de crecimiento endotelial vascular

XbaI: enzima de restricción I de *Xanthomonas badrii*

Y:

μl: microlitro

χ²: prueba Chi cuadrado

X: valor medio

INTRODUCCIÓN

· CD40 Ligando

El ligando de CD40 (CD40LG, CD40L, CD154, TNFSF5, TRAP, gp39, T-BAM) es una proteína transmembrana tipo II de la familia del factor de necrosis tumoral (TNF). Se expresa en la membrana plasmática de los linfocitos T y de otros tipos celulares de forma transitoria en respuesta a activación. Se produce también una forma soluble a partir de la proteína de membrana mediante proteólisis ¹⁻³. La forma activa tanto de membrana como soluble es un trímero, al igual que ocurre con otros miembros de la familia como TNF- α o TNF- β ^{4,5} (Figura 1).

CD40LG fue descubierta a principios de los años 90 en la búsqueda de mediadores de la respuesta humoral, para la que ya se conocía la importancia crucial que juega el contacto directo entre linfocitos T cooperadores (Th) y linfocitos B ⁶. Varios grupos, de forma casi paralela, purificaron una única proteína responsable de la activación policlonal de linfocitos B en reposo por linfocitos T activados. En la universidad de Columbia, un grupo había aislado en 1991 un subclon surgido de forma espontánea de la línea celular humana de linfocito T Jurkat, (Jurkat D1.1). A diferencia de la línea parental, Jurkat D1.1 inducía la activación de linfocitos B en reposo ⁷. En presencia de lectinas inducía la proliferación de Linfocitos B y su diferenciación a células productoras de IgG, actividad que era independiente de antígeno y de HLA pero que requería el contacto celular. Mediante un anticuerpo monoclonal (5c8) precipitaron una proteína de 30 KDa de la membrana de Jurkat D1.1 que llamaron T-BAM (*T cell-B cell Activation Molecule*) ⁸. Ese mismo año otros dos grupos ^{9,10}, trabajando con líneas celulares de linfocitos Th de ratón, aislaron y secuenciaron un DNA complementario (cDNA) que codificaba para una glicoproteína que tenía las mismas propiedades que T-BAM. Uno de ellos denominó a esta proteína gp39 ¹⁰ por su masa aparente de 39 kDa en geles de poliacrilamida. También el mismo año se publicaba el aislamiento de un cDNA humano que fue llamado TRAP (TNF-Related Activation Protein) a partir de una librería de cDNA elaborada con linfocitos T activados ¹¹. TRAP tenía gran homología tanto en el cDNA como en la proteína con la recientemente descrita gp39 en ratón. El gen que codifica para TRAP/gp39/T-BAM se localizó en el genoma humano en el cromosoma X, en Xq26.3-q27.1 ¹². En 1993 se identificaron mutaciones en el gen de gp39, ya denominado CD40 ligando, como responsables del síndrome de hiperIgM ligado al cromosoma X (HIGM1) ^{12,13}.

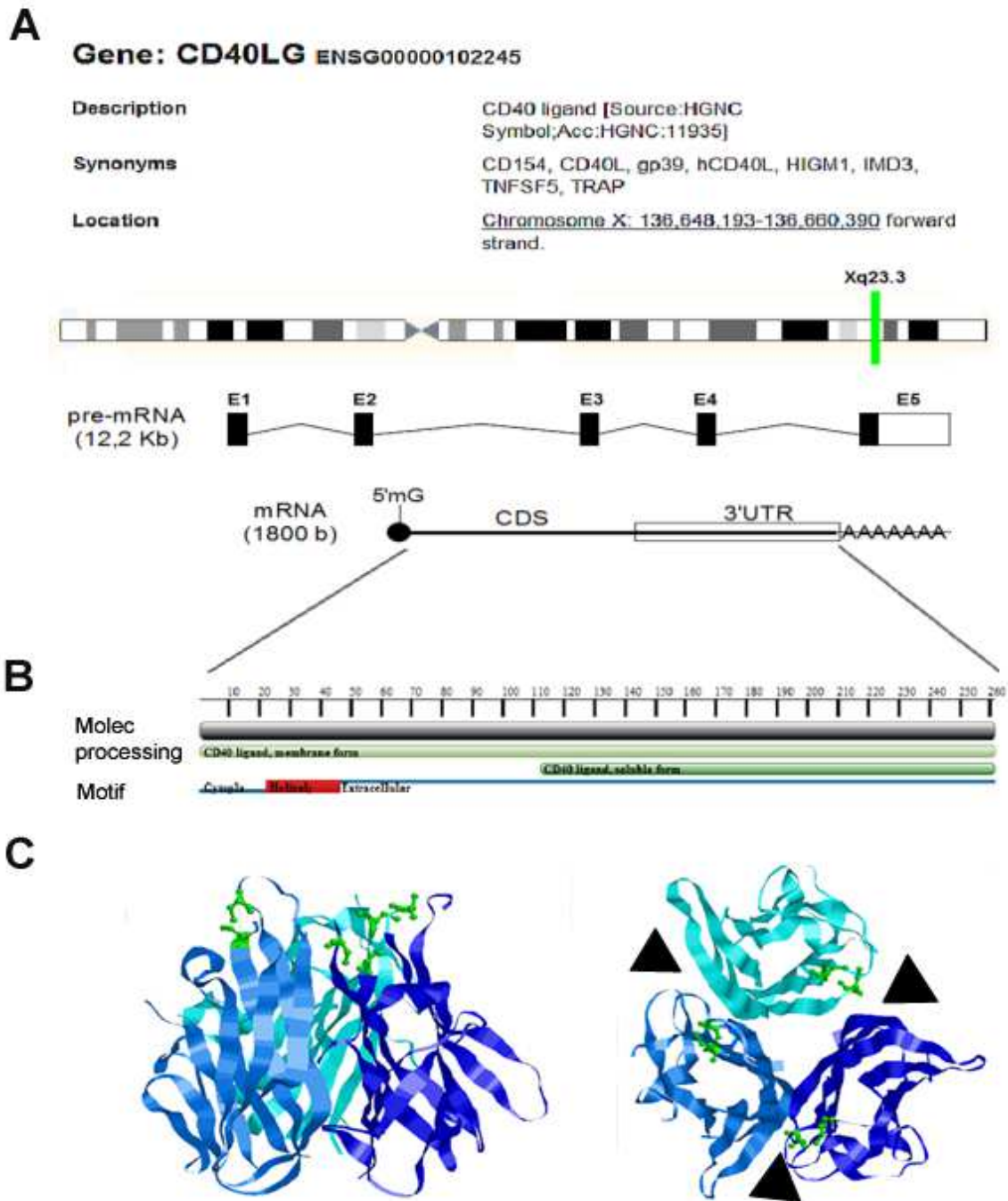


Figura 1. CD40 ligando. (A) Referencia del CD40LG humano extraída de la base de datos Ensembl Genome Browser (línea 'Description' en el gráfico). Se muestran además los sinónimos ('Synonymus') y la localización ('Location') en el genoma en pares de bases según la secuencia de referencia de la versión GRCh38 del genoma humano. (B) Características de CD40LG basadas en la secuencia de aminoácidos (escala en la parte superior) del Protein Data Bank. La proteína de membrana tiene 261 aminoácidos, con el extremo N-terminal en la región citoplásmica (aminoácidos 1-22, línea 'Motif' en el gráfico). A continuación hay un dominio transmembrana (aminoácidos 23-46, en rojo en la línea 'Motif'). Las regiones citoplásmica y transmembrana están codificadas en el primer exón. Los otros cuatro exones codifican la región extracelular, que contiene el dominio tipo TNF. La forma soluble contiene los aminoácidos 113 a 261 (línea 'Molec. Processing'). (C) Imágenes de una parte de la región extracelular tomadas de la estructura 3LKJ del Protein Data Bank con el programa RasMol¹⁴. La estructura se cristalizó como un trímero con una molécula inhibidora de la unión a CD40. El inhibidor se ha eliminado de las imágenes, donde aparecen dos vistas perpendiculares del CD40LG trimérico, con un color para cada monómero. En verde están resaltadas las cisteínas 178 y 218 que forman un puente disulfuro. En la vista cenital (derecha) los triángulos negros señalan los sitios de unión del receptor CD40, situados entre cada par de monómeros.

A raíz de los primeros estudios que llevaron al descubrimiento y caracterización de CD40LG se puso de manifiesto la función de esta proteína en la respuesta humoral T-dependiente, esencial para la formación de centros germinales ^{15,15,16}, el cambio de isotipo de inmunoglobulina por los linfocitos B y la generación de linfocitos B de memoria ¹⁷⁻²⁰. El papel clave de CD40LG en estos procesos se ponía de manifiesto con la ausencia de CD40LG funcional en el síndrome HIGM1. Esta patología se caracteriza por la desorganización del tejido linfoide con ausencia de centros germinales en los ganglios linfáticos, números normales o elevados de linfocitos B que expresan inmunoglobulinas de clase M y D en la superficie y una ausencia total o prácticamente total de linfocitos B que expresan inmunoglobulinas de las clases A, G y E ²¹. El bloqueo en el cambio de clase de las inmunoglobulinas reflejaba el papel clave de CD40LG en la respuesta humoral, pero la aparición de frecuentes infecciones por patógenos oportunistas diversos (virus, bacterias, hongos...) apuntaban además a un papel de CD40LG en la inmunidad celular, corroborado en modelos murinos en los que la ausencia de CD40LG impide el control de infecciones virales ²².

Estas funciones están mediadas por la unión de CD40LG en la membrana de linfocitos T activados al receptor CD40, que se expresa de forma constitutiva en todos los linfocitos B maduros y en otros tipos celulares como monocitos, células dendríticas, células epiteliales o endotelio ²³. La señal CD40-CD40LG participa por tanto en la comunicación entre diferentes tipos celulares, contribuyendo a la regulación del tipo y el grado de respuesta con que el organismo responde a diferentes antígenos. Además de la activación de linfocitos B, que desencadena la formación de centros germinales y el cambio de isotipo de inmunoglobulina, la señal CD40-CD40LG participa en la respuesta inmune a través de la interacción con el CD40 presente en células dendríticas y en monocitos/macrófagos, en los que induce su activación y maduración (Figura 2C)²⁴⁻²⁷.

La activación de células dendríticas por la señal CD40LG-CD40 es importante para la inducción de respuestas citotóxicas por linfocitos CD8 ^{28,29}. Ésta activación se ha descrito de forma secuencial, con la activación inicial de la célula dendrítica por linfocitos T CD4 cooperadores vía CD40LG y la posterior activación de linfocitos T CD8 por la célula dendrítica previamente activada ³⁰. La señal CD40LG proporcionada por linfocitos T CD4 a células dendríticas parece ser crucial para la generación de linfocitos T CD8 citotóxicos de memoria ^{31,32}. Esta interacción resulta

esencial en algunos casos de forma directa para la eliminación por parte de los macrófagos de algunos parásitos intracelulares en modelos de enfermedad en ratones ³³⁻³⁵.

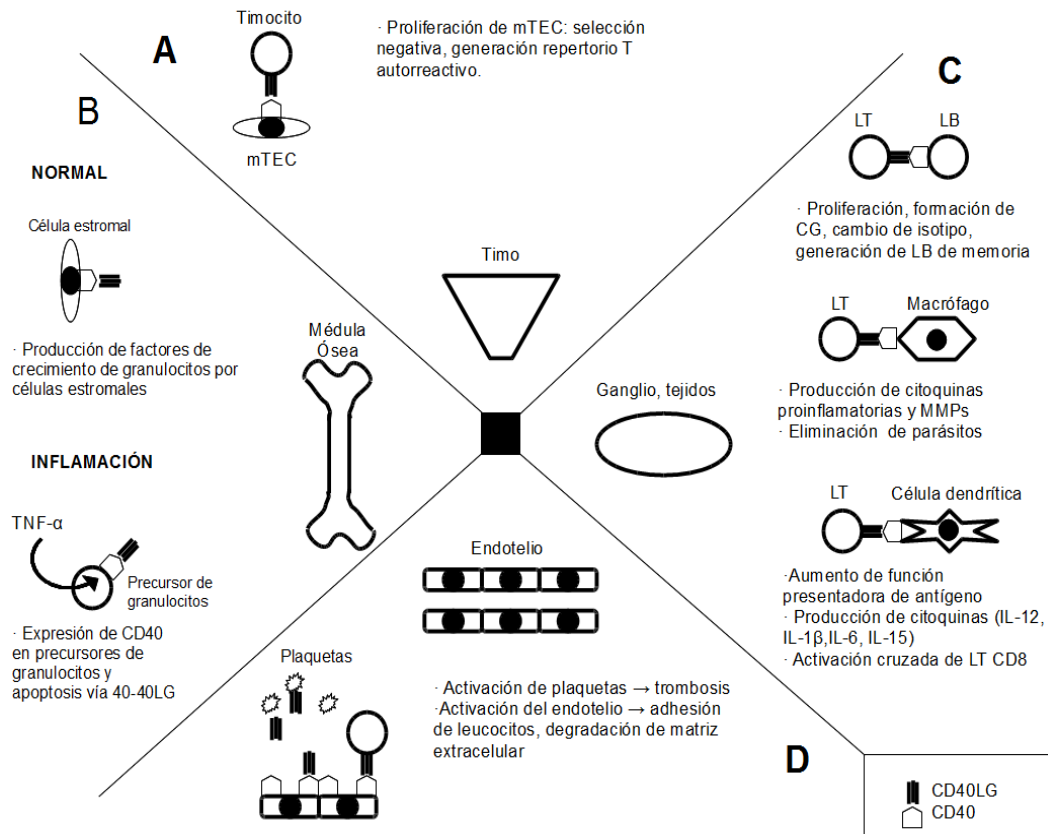


Figura 2. Funciones de la unión CD40LG-CD40 en diferentes tipos celulares. La señal CD40LG-CD40 participa en numerosas funciones en el sistema inmune. (A) La señal entre timocitos y células epiteliales de la médula del timo (mTECs) participa en el mantenimiento de la médula tímica, la selección negativa y la generación del repertorio T regulador. (B) Tiene también un papel en el desarrollo de progenitores de granulocitos en médula ósea. (C) En la periferia, CD40LG se expresa en linfocitos T activados, donde su expresión regula la respuesta inmune llevada a cabo por distintos tipos celulares (linfocitos B, macrófagos, células dendríticas). (D) Participa en la trombosis con las formas de membrana y soluble expresadas por las plaquetas. CG: Centros germinales. MMPs: metaloproteasas. LT: linfocitos T; LB: linfocitos B.

CD40LG también participa en el desarrollo y maduración de las células epiteliales de la médula del timo (mTEC), que expresan el receptor CD40 (Figura 2A). Entre ellas se encuentra la subpoblación que expresa el gen AIRE (*AutoImmune REgulator*), que se encarga de la expresión de antígenos tisulares que dan lugar a la selección negativa de timocitos

autorreactivos y a la generación del repertorio de linfocitos T reguladores ³⁶⁻³⁸. CD40LG participa también en el desarrollo de progenitores hematopoyéticos de granulocitos en médula ósea (Figura 2B)³⁹⁻⁴¹

Aunque originalmente CD40LG se había descrito como una molécula presente en linfocitos T activados, principalmente CD4⁺, posteriormente se describió su expresión en otros tipos celulares como linfocitos B ⁴²⁻⁴⁴, plaquetas ⁴⁵, eosinófilos ⁴⁶, macrófagos, endotelio, células de músculo liso ⁴⁷ o células dendríticas ⁴⁸. La inducción de la expresión de CD40LG por ciertos virus en estas últimas puede estimular respuestas citotóxicas por linfocitos T CD8 de forma independiente de linfocitos T CD4 ⁴⁸ aunque como se ha mencionado antes, la señal de CD40LG proporcionada por células dendríticas a linfocitos T CD8 no parece poder sustituir a la ayuda prestada por linfocitos T CD4 CD40LG⁺ para la generación de linfocitos CD8 de memoria y la producción de respuestas secundarias frente a un antígeno ³². Por otra parte, se han descrito hasta la fecha otros tres receptores para CD40LG distintos de CD40 (Tabla 1), las integrinas $\alpha 5\beta 1$, $\alpha M\beta 2$ (Mac-1) y $\alpha IIb\beta III$. La integrina $\alpha 5\beta 1$ se expresa en monocitos y plaquetas y su unión a CD40LG en monocitos produce una señal activadora a través de las quinasas ERK1/2 y la expresión del gen de la interleucina 8 (IL-8) ^{49,50}. La unión a Mac-1, presente en monocitos y neutrófilos, participa en la adhesión a endotelio y la migración transendotelial ^{51,52}. El tercer receptor, la integrina $\alpha IIb\beta III$, se expresa en plaquetas. Las plaquetas expresan tres de los cuatro receptores conocidos para CD40LG, además de expresar CD40LG. El descubrimiento de la expresión de CD40LG en plaquetas ⁴⁵ significó un cambio en la visión que se tenía de la función inmunológica de éstas, situándolas como intermediarias entre la inmunidad innata y la adaptativa, para la que CD40LG tiene un papel fundamental ^{53,54}. El CD40LG de plaquetas, al igual que el expresado por linfocitos T y en colaboración con éste, induce la maduración de células dendríticas aumentando tanto su capacidad presentadora de antígeno como su capacidad de activar la respuesta de linfocitos T CD8, y de activar a linfocitos B para la formación de centros germinales y el cambio de isotipo ⁵⁵. Por otra parte, tiene un papel relevante en el proceso de trombosis ⁵⁶, participando en la señalización entre los distintos tipos celulares implicados. Las plaquetas contienen CD40LG preformado almacenado en los gránulos α ⁵⁷, que se moviliza a la membrana de forma rápida tras su activación. Una vez en la membrana, se libera en forma soluble por un mecanismo que se dispara a través de la integrina $\alpha IIb\beta III$ ^{58,59}. El CD40LG soluble liberado

induce la activación y la liberación de metaloproteasas (MPP) por parte del endotelio circundante, que degradan la matriz extracelular ⁶⁰ (Figura 2D).

La unión de CD40LG a sus receptores no sólo produce una señal en la célula que expresa el receptor al que se une. En 1997 se describió cómo la porción citoplásmica de CD40LG activa vías de señalización en el linfocito T CD40LG⁺ en respuesta a la unión a CD40 ⁶¹. En la línea celular Jurkat D1.1, la señal a través de CD40LG produce un aumento de Ca²⁺ intracelular y activa distintas isoformas de la proteína kinasa C (PKC α y PKC δ), induciendo la fosforilación de las MAP kinasas ERK1/2 y p38 MAPK ⁶². Esta señal parece ser importante para ciertas funciones como la regulación de la producción de IL-4 ⁶³.

TABLA 1. Otros receptores para CD40LG y funciones de la unión CD40LG-receptor

Integrina	Expresión	Función	Referencias
α II β III	Plaquetas	· Activación, \uparrow CD40LG, liberación CD40LG soluble	58
			59
			60
α 5 β 1	Macrófagos	· Activación, producción de IL-8, MMP-2, MMP-9	50
	Plaquetas		49
Mac-1	Neutrófilos, Macrófagos	· Adhesión a endotelio y migración transendotelial	52
			51
			64

· Regulación de la expresión de CD40LG

Los primeros experimentos realizados en linfocitos T y en clones de linfocitos Th1 y Th2 ⁶⁵ para estudiar la expresión de CD40LG parecían indicar que el principal factor que determina el grado de expresión de CD40LG es la intensidad de la señal antigénica a través del receptor de antígeno (TCR). La cual a su vez se veía incrementada por la presencia de moléculas de adhesión, que facilitan un contacto más firme entre el linfocito T y la célula presentadora de antígeno ⁶⁶. Numerosos trabajos, sin embargo, han descrito posteriormente cómo diferentes citocinas presentes en el medio influyen también en la expresión de CD40LG, de forma diferente en distintos subtipos de linfocitos T ⁶⁷⁻⁷¹, habiendo diferentes requerimientos para una máxima expresión de CD40LG según los subtipos de linfocitos T ⁶⁵.

Cuando se activan linfocitos T *in vitro*, en las horas siguientes al inicio de la activación hay un aumento progresivo de CD40LG en la membrana, que requiere la transcripción y traducción del RNA mensajero (mRNA), alcanzando un máximo de expresión entre las 6 y las 24 h en función del estímulo y la composición celular de las muestras activadas, y declinando progresivamente en las siguientes 24-72 h ^{66,68}. La expresión de CD40LG en respuesta a la activación vía TCR varía entre los distintos subtipos de linfocitos T, y forma parte del programa de expresión génica que define sus funciones. Así, los linfocitos T reguladores no expresan CD40LG tras estimulación *in vitro*, lo que contribuye a generación de tolerancia frente a antígenos propios ^{72,73}, mientras que los linfocitos T efectoros o de memoria contienen CD40LG preformado que se moviliza en pocos minutos a la membrana tras activación ^{74,75}, lo que probablemente contribuye a la mayor rapidez de respuesta de los linfocitos T de memoria ⁷⁶.

·Regulación de la transcripción de CD40LG

La transcripción de CD40LG se activa *in vitro* en linfocitos T con la estimulación a través del TCR mediante anticuerpos anti-CD3. También puede activarse su transcripción disparando directamente las vías de señalización que se inducen tras la señal antígeno-TCR utilizando compuestos activadores de la proteína kinasa C (PKC) como el forbol-12-miristato acetato (PMA), o con ionomicina (Io), que produce un aumento de Ca²⁺ en el citoplasma. El aumento de la concentración de Ca²⁺ intracelular es suficiente para la transcripción de CD40LG, actuando a través de dos vías que llevan finalmente a la activación de los factores de transcripción NFAT y AP-1 ⁷⁷⁻⁸⁰ (Figura 3), aunque la máxima expresión se consigue con la combinación de PMA/Io ⁸¹. Se ha descrito la participación de otros tres factores de transcripción en la expresión de CD40G como Egr1 ⁸² y AKNA ⁸³, que tienen sitios de unión en el promotor, y NFκB, que se une al extremo 3' del gen actuando como un *enhancer* ^{84,85}.

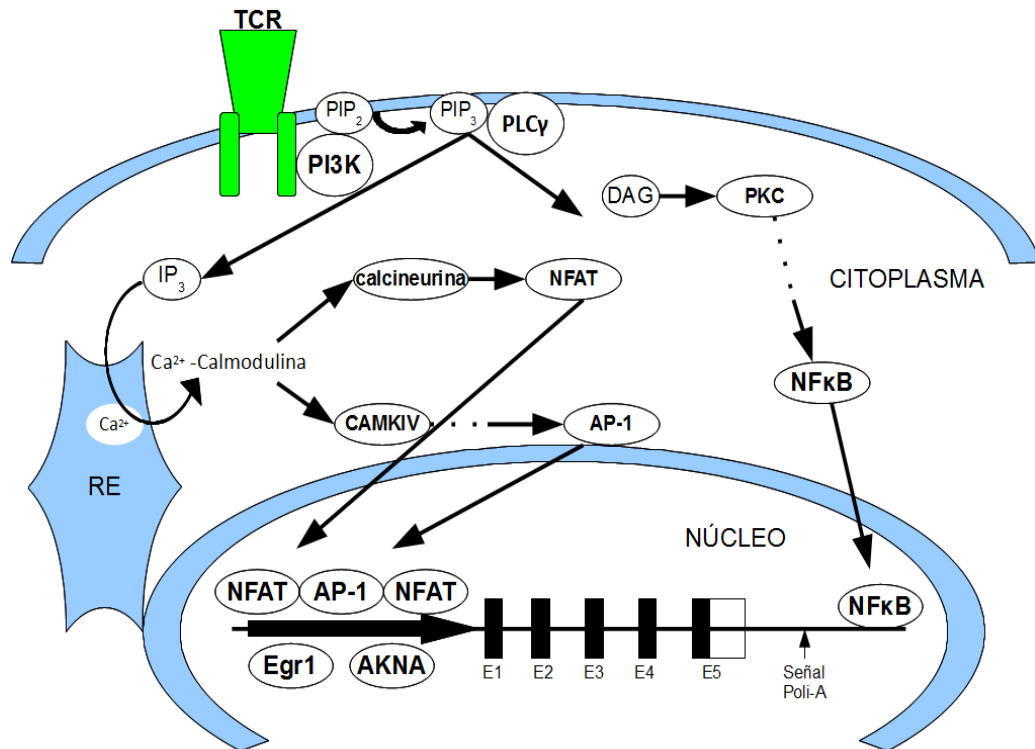


Figura 3. Esquema simplificado de las vías de señalización para la transcripción del gen CD40LG. La activación del TCR conduce a la producción de fosfolípidos de inositol trifosfato (PIP₃) en la membrana por la enzima fosfatidilinositol 3 quinasa (PI3K). La fosfolipasa C gamma (PLC γ) rompe el PIP₃ en diacilglicerol (DAG), que activa a la PKC, y fosfatidil inositol trifosfato (PIP₃), que abre canales de calcio en la membrana del retículo endoplásmico (RE) produciendo un aumento de Ca²⁺ en el citoplasma. El Ca²⁺ unido a calmodulina activa a las proteínas calcineurina y quinasa IV dependiente de calcio-calmodulina (CAMKIV). Éstas activan los factores de transcripción proteína activadora 1 (AP-1) y factor nuclear de células T activadas (NFAT), que mediante su unión al promotor activan la transcripción del gen. En el promotor de CD40LG hay además sitios de unión para los factores de transcripción Egr1 y AKNA, y en el extremo 3' del gen hay un sitio de unión para NF κ B que actúa como enhancer. Flecha: indica activación. E1-E5: Exones de CD40LG.

·Regulación posttranscripcional de CD40LG

La regulación posttranscripcional implica el control de cambios en la estabilidad de los mRNA y de la eficiencia con que se traducen, y es característica de genes que codifican para factores de transcripción, proteínas reguladoras del ciclo celular o citoquinas y factores de crecimiento ⁸⁶, de forma que su expresión puede adaptarse de forma rápida a cambios en el entorno ⁸⁷. En el sistema inmune, la regulación posttranscripcional tiene un papel fundamental en el control de la respuesta frente a patógenos, como lo demuestran varios modelos de ratón con inactivación de genes que codifican para proteínas como TTP o AUF-1, que regulan la estabilidad de mensajeros que codifican para mediadores importantes como el factor de necrosis tumoral (TNF- α) o la IL1- β . ^{88,89}, modelos que presentan una respuesta inflamatoria descontrolada tras la estimulación con un antígeno.

El mRNA que se transcribe del gen CD40LG tras la activación de linfocitos T tiene inicialmente una vida media corta, de aproximadamente 20-30 min ⁹⁰. Sin embargo, una activación prolongada vía TCR (>24h) ⁹⁰ estabiliza el mRNA de CD40LG aumentando su vida media de unos 30 a 120 min (Tabla2). Determinados estímulos como PMA o la señal proporcionada por la molécula de adhesión LFA-3 ^{81,91} también son capaces de estabilizar el mRNA de CD40LG de forma rápida. La estabilización del mRNA de CD40LG, a diferencia de lo que ocurre con los mRNA que codifican para otras citoquinas como IL-2, TNF- α , IFN- γ o GM-CSF, es independiente de la coestimulación a través de CD28 ⁶⁶.

TABLA 2. Vida media del mRNA de CD40LG en trabajos publicados

mRNA	Tipo celular	Estímulo	Parada de transcripción	t $\frac{1}{2}$ (min) ¹	Referencias
CD40LG	LT CD4 (ratón)	Anti-CD3/28 ² , 6h	DRB ³	30	92
		Anti-CD3/28, 24h	DRB	55	
CD40LG	PBMCs ⁴	Io	Act-D ⁵	50	81
CD40LG	PBMCs	PMA/Io	Act-D	120	
luc-3'UTR ⁶	HeLa	-	Doxiciclina	50-90	93
luc-3'UTR ⁸	HeLa	-	Doxiciclina	50	94
CD40LG	LT CD4	En reposo	DRB	30	95
CD40LG	LT CD4	PHA ⁷ /PMA, 2h	DRB	120	
CD40LG	Jurkat D1.1	-	DRB	120	95,96
CD40LG	Jurkat D1.1	-	Act-D	120	97
CD40LG	LT CD4	Anti-CD3, 2h	Act-D	40	90
		Anti-CD3, 8h	Act-D	43	
		Anti-CD3, 24h	Act-D	84	
		Anti-CD3, 48h	Act-D	132	
CD40LG	LT CD4	Anti-CD3/28, 6h	Act-D	40	98
		Anti-CD3/28, 24h	Act-D	160	
CD40LG	LT	En Reposo	Act-D	51- >360	99
CD40LG	LT	Anti-CD3/28, 3h	Act-D	45 - 79	

¹Tiempo de vida media. ²Anticuerpos frente a CD3 y CD28. ³5,6-Dicloro-1-beta-D-ribofuranosilbenzimidazol. ⁴Células mononucleares de sangre periférica. ⁵ Actinomicina D. ⁶Región 3'UTR de CD40LG clonada en un vector de expresión a continuación de la enzima luciferasa. ⁷Fitohemaglutinina. ⁸3'UTR del gen CD40LG de ratón.

El cambio en la estabilidad del mRNA de CD40LG está regulado mediante la unión de proteínas a secuencias presentes en la región 3' no traducida (3'UTR). Entre las secuencias diana para la unión de proteínas reguladoras la clase más común y estudiada se conoce como *AU-rich elements* (secuencias ricas en adenina y uracilo, ARE). Las secuencias ARE se localizan en la región 3'UTR de los mRNA de expresión rápida y transitoria que codifican para citoquinas, proto-oncogenes y factores de transcripción ^{100,101}. El mRNA de CD40LG contiene secuencias ARE de dos tipos (Figura 4). Por un lado hay cinco repeticiones del pentámero AUUUA, tres en la primera mitad de la región 3'UTR y dos hacia el final (varios pentámeros AUUUA agrupados en una zona limitada forman los ARE tipo I, cuyo ejemplo más representativo es el del mRNA de c-fos ¹⁰²). El último pentámero forma el cuerpo de un nonámero UUAUUUAUU (al menos dos nonámeros seguidos son característicos de los ARE tipo II, que se encuentran en los mRNA de varias citoquinas como TNF- α o GM-CSF ¹⁰²). El mRNA de CD40LG tiene, por tanto, varias secuencias tipo ARE, sin embargo éstas no tienen las características típicas de los ARE tipo I o II, y según los estudios publicados hasta ahora, la regulación de la estabilidad del mRNA de CD40LG parece ser independiente ellas ¹⁰³.

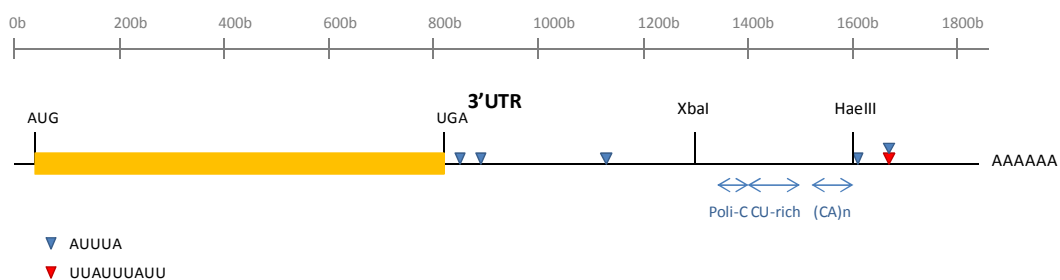


Figura 4. Secuencias ARE en el mRNA de CD40LG. Esquema del mRNA de CD40LG con la secuencia codificante señalada en naranja. En la región 3'UTR (a partir del codón stop UGA) están señaladas las secuencias tipo ARE (▼; ▼). La escala superior marca de forma aproximada las bases en la secuencia de referencia del mRNA de CD40LG (NM_000074.2). También se señalan en azul, entre las dianas de restricción XbaI y HaeIII, otras tres regiones con secuencias repetidas para las que se ha descrito un papel regulador: poli-C, CU-rich y un microsatélite formado por repeticiones en tándem (CA)_n.

Varios trabajos han descrito la unión de complejos de proteínas a la región 3'UTR del mRNA de CD40LG que regulan la estabilidad del mRNA. La situación de estos complejos en la región 3'UTR de CD40LG se muestran en la Figura 5, donde están señaladas tres regiones consecutivas de secuencias repetitivas que consisten en una tira de citosinas (poli-C), una

región rica en citosina y uracilo (*CU-rich*) y un microsatélite (MSCD40LG) formado por repeticiones en tándem del dinucleótido (CA)_n.

En 1999 se describió por primera vez cómo varias proteínas se unían a la región 3'UTR del mRNA de CD40LG ¹⁰³, con polipéptidos de 25 y 50 KDa como componentes principales. En linfocitos T activados a tiempos cortos con fitohemaglutinina (PHA), el mRNA unía los polipéptidos de 25 y 50 KDa. Al añadir PMA, en paralelo a la estabilización del mRNA se perdía la unión del péptido de 25 KDa y aparecían además, unidos a la misma región, polipéptidos de tamaño intermedio entre 25 y 50 KDa. Más adelante ese mismo grupo caracterizó los polipéptidos de 25 y 50 KDa, y los identificó como la ribonucleoproteína PTB (*polypyrimidine tract binding protein*, hnRNP I) y una isoforma truncada de 25 kDa que llamaron PTB-T ⁹³. PTB-T y PTB se unían a los mismos sitios con efectos antagónicos en la estabilidad del mRNA de CD40LG, PTB-T promovía inestabilidad mientras que PTB aumentaba la estabilidad.

Otro grupo describió dos complejos de proteínas, que llamaron *Complex I* y *Complex II*, que se unían a la misma parte de la región 3'UTR de CD40LG, situada entre las dianas XbaI y HaeIII ⁹⁷ (Figura 5). La unión correlacionaba con la estabilización del mRNA de CD40LG en el transcurso de la activación de linfocitos T CD4⁺, y se daba también en el mRNA estabilizado de forma constitutiva de la línea Jurkat D1.1. Aislaron una proteína de 55 kDa que se unía al mRNA y la identificaron como PTB. En trabajos posteriores se caracterizaron otras dos proteínas que formaban parte de estos complejos: *Nucleolin* (en el *Complex I*) y hnRNPL (*heterogeneous nuclear ribonucleoprotein L*, en el *Complex II*) ^{94,104,105}.

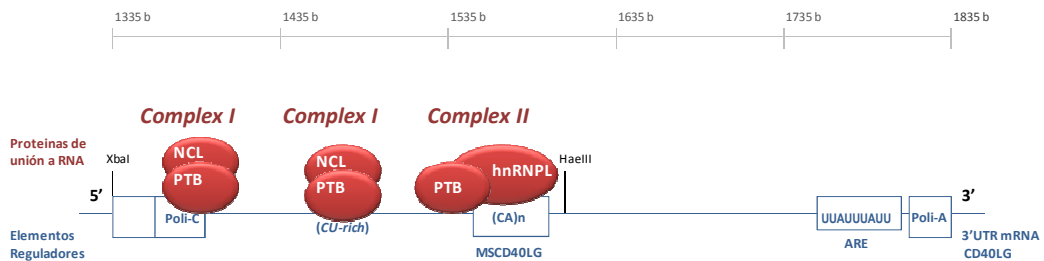


Figura 5. Complejos proteicos de unión al mRNA de CD40LG. Esquema de la región 3'UTR del mRNA de CD40LG entre la diana de restricción XbaI y la cola de poli-A, con los complejos de polipéptidos que regulan la estabilidad del mRNA. La ribonucleoproteína PTB tiene una isoforma truncada (PTB-T) que se une a los mismos sitios que PTB. NCL: *Nucleolin*. hnRNPL: (*heterogeneous nuclear ribonucleoprotein*),

Así pues, la región que regula la estabilidad del mRNA de CD40LG comprende una secuencia de unos 360 nucleótidos que comienza en la diana de restricción XbaI (Figura 5). Esta región contiene una repetición del mononucleótido de citosina (poli-C), una zona rica en citosinas y uracilos (*CU-rich*), y un microsatélite (MSCD40LG).

· CD40LG y autoinmunidad

Poco después de su descubrimiento, por su importante papel en la regulación de la respuesta humoral y en la producción de anticuerpos, se empezó a relacionar a CD40LG con distintas patologías, inicialmente enfermedades autoinmunes en las que se producen autoanticuerpos con un papel patológico en la enfermedad ^{106,107}. Se ha descrito un aumento del número de linfocitos T que expresan CD40LG o de CD40LG soluble en plasma en muchas de estas enfermedades, como en el lupus eritematoso sistémico (LES) ¹⁰⁸⁻¹¹¹, esclerosis múltiple ¹¹², enfermedad inflamatoria intestinal ¹¹³ o artritis reumatoide (AR) ^{114,115}. La elevada expresión de CD40LG en estas enfermedades es un reflejo de la activación crónica del sistema inmune, y en algunas patologías como en la AR se ha descrito cómo los linfocitos T CD4⁺ extraídos del sinovio y activados *in vitro* expresan CD40LG en pocos minutos y de forma independiente de transcripción y traducción, lo que indica la presencia de linfocitos T CD4⁺ de memoria con CD40LG preformado implicados en la reacción autoinmune frente a componentes de la articulación. En este sentido, la cantidad de linfocitos T que expresan CD40LG en superficie o de CD40LG soluble generalmente correlacionan con la actividad de la enfermedad ^{116,117}. El aumento de la expresión de CD40LG no implica que éste sea causante de una enfermedad, aunque el síndrome HIGM1 ¹¹⁸ y modelos transgénicos de ratón que expresan CD40LG de forma deslocalizada ^{119,120} muestran cómo la desregulación drástica tanto por defecto como por exceso pueden llevar por sí mismas a la aparición de autoinmunidad.

La elevada expresión de CD40LG en varias enfermedades autoinmunes, unida al papel fundamental de CD40LG en la formación de centros germinales donde se producen el cambio de isotipo, la hipermutación somática y la maduración de la afinidad de los anticuerpos, señalaron a la interacción CD40-CD40LG como una atractiva diana terapéutica. Una de las enfermedades donde más se ha estudiado esta posibilidad es el LES, para la que hay modelos murinos de la enfermedad desde hace décadas, siendo el más estudiado la

primera generación del cruce entre las cepas congénicas New Zealand Black (NZB) y New Zealand White (NZW), que desarrolla de forma espontánea una enfermedad de curso agresivo con producción de autoanticuerpos anti-DNA e histonas y glomerulonefritis por inmunocomplejos. En este modelo, el bloqueo de la señal CD40-CD40LG en etapas preclínicas de la enfermedad inhibe la producción de anticuerpos anti DNA e histonas y previene la aparición de glomerulonefritis durante el tratamiento. Este efecto se mantiene durante un tiempo tras la interrupción del tratamiento ^{121,122} y prolonga la vida de los animales. En animales con enfermedad establecida, el bloqueo de la señal CD40-CD40L induce remisión parcial de la enfermedad en algunos casos, pero la eficacia es mucho menor comparada con el bloqueo en etapas tempranas ¹²¹. De forma similar, en un modelo de IBD inducido en la cepa murina SJL por inyección intrarrectal de ácido 2,4,6-trinitro-bencenosulfónico, la administración de un anticuerpo bloqueante anti-CD40LG de forma simultánea a la inducción de la enfermedad previene la acumulación de linfocitos T CD4⁺ productores de interferón gamma (IFN γ) y evita la aparición de signos histológicos de enfermedad, mientras que el anticuerpo anti-CD40LG administrado cuando la enfermedad está establecida no tiene efecto en la actividad de la misma ¹²³. En estos dos modelos, CD40LG tiene un papel clave en el desarrollo inicial de la enfermedad, pero resulta prescindible una vez establecida para la perpetuación de la inflamación crónica. El bloqueo de CD40LG con anticuerpos monoclonales se ha tratado de llevar a la práctica clínica y se han llegado a realizar ensayos en fase II con anticuerpos monoclonales frente a CD40LG para el tratamiento del LES ¹²⁴, pero éstos tuvieron que ser interrumpidos por la aparición de problemas de tromboembolismo ¹²⁵. Recientemente se ha publicado la síntesis de un anticuerpo modificado que podría evitar las complicaciones de tromboembolismo ¹²⁶.

· **Microsatélite de CD40LG. Asociación genética a LES y AR**

Numerosas patologías crónicas, entre las que se encuentran las enfermedades autoinmunes mediadas por autoanticuerpos como LES, AR, MS o T1DM tienen un componente genético complejo que aún no está competamente definido ¹²⁷. Como se ha mencionado anteriormente, CD40LG tiene un papel importante en el funcionamiento del sistema inmune y la desregulación en la expresión de esta proteína puede llevar a la aparición de autoinmunidad, tanto por exceso como por defecto. Parece posible, por tanto, que variantes genéticas que afecten a la expresión de CD40LG puedan formar parte del componente genético que modifica la predisposición a desarrollar autoinmunidad. En la

región 3' no traducida (3'UTR) de CD40LG se sitúa un microsatélite (MSCD40LG), situado entre las dianas de restricción XbaI y HaeIII, en la zona donde se encuentran las dianas conocidas para la unión de las proteínas que regulan la expresión del mRNA (Figura 5).

Trabajos anteriores realizados por nuestro grupo describieron la distribución alélica del MSCD40LG en población española de las Islas Canarias y de la península ¹²⁸. El alto grado de polimorfismo del MSCD40LG, cuyas variantes alélicas consisten en el diferente número de repeticiones (CA)_n que forman el microsatélite, junto con la importante función de CD40LG en la regulación de la respuesta inmune y la alteración de su expresión descrita en numerosas enfermedades autoinmunes, sugerían que podía tratarse de un buen marcador para estudiar la posible asociación genética de CD40LG a estas enfermedades, asociación que estudiamos en dos enfermedades autoinmunes, LES y AR.

En el estudio de asociación a LES ¹²⁸ encontramos que el alelo mayoritario en la población (24CA) aparecía con menor frecuencia en pacientes de LES que en controles sanos. Uniendo los alelos mayores y menores de 24CA (mayor o menor número de repeticiones (CA)_n en dos grupos (>24CA y <24CA), los alelos >24CA tenían mayor frecuencia en pacientes de LES que en controles. Esta asociación se daba sólo en mujeres. En otro estudio posterior se analizó la asociación a AR ⁹⁸ y encontramos también que el alelo 24CA tenía menor frecuencia en pacientes de AR que en controles sanos cuando se encontraba en homocigosis. Esta asociación también se daba sólo en mujeres y era independiente de HLA. En este estudio los alelos >24CA no tenían mayor frecuencia en pacientes, y los alelos distintos del 24CA se unieron en un único grupo (X) resultando tres genotipos (2424, 24X y XX). Con esta agrupación había una menor frecuencia de mujeres homocigotas 2424 en AR que en controles.

Aparte de los dos trabajos mencionados, se han publicado un escaso número de estudios de asociación del MSCD40LG a enfermedades autoinmunes entre las que se encuentran la AR ^{129,130} y la esclerosis múltiple ¹³¹. En cuanto a la AR, en los trabajos mencionados se estudiaba la posible asociación de marcadores cercanos a distintos genes candidatos, entre ellos CD40LG. El trabajo de John *et al.* (John *et al.*, 1998) estudiaba la transmisión de alelos y su correlación con artritis en familias con dos miembros afectados. Como los autores del trabajo indican, para poder detectar efectos pequeños en un diseño de este tipo se necesitaría una muestra mucho mayor de la que disponían. El estudio de Gomolka ¹²⁹, con un diseño de casos y controles en población alemana, tampoco encontró

asociación del alelo mayoritario, que los autores también designaron como 24CA, y que por la distribución alélica que muestran los autores parece coincidir con el alelo 24CA en población canaria. Sí describían una mayor frecuencia de portadores del alelo 21CA en individuos que no son portadores de alelos de HLA-DR4 ni HLA-DR1 (Controles: 0,02; AR: 0,21). Resulta interesante que esta asociación se replicaba en nuestras muestras de población canaria (Controles: 0,03; AR: 0,18).⁹⁸.

Además de la asociación genética, en nuestros trabajos previos habíamos encontrado una menor expresión de CD40LG en la membrana de linfocitos T CD4⁺ de muestras homocigotas para el alelo 24CA, en ensayos *in vitro* en los que PBMCs de controles se activaban con diferentes estímulos^{98,132}. Además, en el trabajo de Martín-Donaire et al. se activaron PBMCs de controles sanos y pacientes de AR homocigotos para el alelo 24CA o portadores de alelos distintos (NO24CAs) con anti-CD3/28 durante 6 h y 24 h y se midió la estabilidad del mRNA. Tras 6 h de activación, las muestras con alelos 24CA tenían una vida media ($t_{1/2}$) que en promedio aproximadamente doblaba la de las muestras NO24CA. Tras 24 h de activación, el mRNA se estabilizaba de forma notable y nuevamente las muestras con 24CA tenían en promedio el doble de vida media, lo que parecía indicar que el MSCD40LG tenía un efecto en la estabilidad del mRNA. El diseño de este trabajo, con todos los alelos distintos del alelo 24CA unidos en un mismo grupo, no permitía establecer las diferencias en la estabilidad del mRNA en función del número de repeticiones del dinucleótido (CA)_n del MSCD40LG. Además, los ensayos se realizaron en muestras de PBMCs expandidos y reestimulados, por lo que diferencias en la composición celular entre muestras probablemente contribuían de forma significativa a la diferencia individual de la vida media del mRNA CD40LG medida¹³³.

En el presente trabajo hemos querido profundizar en este hallazgo y tratar de determinar si el MSCD40LG funciona por sí sólo como un elemento en "cis" que afecta a la estabilidad del mRNA, y medir la magnitud del cambio en estabilidad en función del número de repeticiones del dinucleótido (CA)_n. Para ello, hemos clonado el fragmento de la región 3'UTR de CD40LG que contiene las secuencias reguladoras de la estabilidad (donde se encuentra el MSCD40LG) en vectores de expresión cuyos fragmentos 3'UTR se diferencian únicamente en el número de repeticiones (CA)_n del MSCD40LG, y hemos insertado estas construcciones en una línea celular para estudiar la estabilidad de los mRNA quiméricos que llevan la región 3'UTR de CD40LG a continuación de la proteína verde fluorescente (GFP).

OBJETIVOS

Los objetivos generales del presente trabajo han sido:

- 1- Revisar la asociación genética del microsatélite del gen CD40LG a artritis reumatoide y a lupus eritematoso sistémico, descrita previamente por nuestro grupo, mediante el genotipado de nuevas muestras de DNA genómico de individuos sanos y pacientes de LES.
- 2- Estudiar el efecto de variaciones en el número de repeticiones del dinucleótido (CA)_n del microsatélite de CD40LG en la estabilidad del RNA mensajero mediante el clonaje de distintas variantes alélicas comunes en vectores de expresión, transfección en una línea celular y ensayos de caída del mRNA.

MATERIALES Y MÉTODOS

Pacientes y controles

Se han analizado muestras de DNA de varios grupos de pacientes y donantes sanos voluntarios disponibles en el laboratorio para el análisis de genotipos. Todos los individuos dieron su consentimiento para ser incluidos en estudios de genética que habían sido aprobados por los Comités de ética de las instituciones participantes.

· Series de pacientes

LESC1

Está constituida por 80 mujeres no relacionadas diagnosticadas de lupus eritematoso sistémico que cumplían 4 o más criterios del *American College of Rheumatology* (ACR) ¹³⁴, procedentes del Servicio de Reumatología del Hospital Universitario de Gran Canaria Dr. Negrín (Las Palmas de Gran Canaria), donde habían sido clínicamente evaluadas y diagnosticadas.

LESC2

Reúne a 123 mujeres no relacionadas diagnosticadas de lupus eritematoso sistémico. La totalidad de los casos cumplían 4 o más criterios del *American College of Rheumatology* (ACR) ¹³⁴, procedentes del Servicio de Reumatología del Hospital Universitario de Gran Canaria Dr. Negrín (Las Palmas de Gran Canaria), donde habían sido clínicamente evaluadas y diagnosticadas.

LESP1

Reúne a 26 mujeres no relacionadas diagnosticadas de lupus eritematoso sistémico. La totalidad de los casos cumplían 4 o más criterios del *American College of Rheumatology* (ACR) ¹³⁴, procedentes del Servicio de Medicina Interna del Hospital Universitario Puerta de Hierro, Majadahonda (Madrid), donde habían sido clínicamente evaluadas y diagnosticadas.

ARC1

Está constituida por 140 mujeres no relacionadas diagnosticadas de artritis reumatoide según los criterios descritos por el *American College of Rheumatology* (ACR) ¹³⁵, procedentes del Servicio de Reumatología del Hospital Universitario de Gran Canaria Dr.

Negrín (Las Palmas de Gran Canaria), donde habían sido clínicamente evaluadas y diagnosticadas.

· **Series de controles**

DSC1

Está constituida por 80 mujeres donantes sanas voluntarias sin antecedentes de enfermedades autoinmunes y procedentes del Hospital Universitario de Gran Canaria Dr. Negrín (Las Palmas de Gran Canaria).

DSC2

Está constituida por 74 mujeres donantes sanas voluntarias sin antecedentes de enfermedades autoinmunes y procedentes del Hospital Universitario de Gran Canaria Dr. Negrín (Las Palmas de Gran Canaria).

DSP1

Está constituida por 241 mujeres donantes sanas voluntarias, trabajadoras de Ediciones y Promociones S.A. (SAPE), todas ellas procedentes de Madrid.

DSP2

Está constituida por 54 mujeres donantes sanas voluntarias y reúne a donantes de sangre del Hospital puerta de Hierro Majadahonda.

Materiales

· **Cepas bacterianas de *E. coli***

DH5- α

Cepa DH5- α de bacterias *E. coli* competentes (Life Technologies, Madrid, España), tratadas químicamente para favorecer el clonaje. Proporcionan una eficiencia de transformación $> 1 \times 10^6$ unidades formadoras de colonias por μg de DNA plasmídico.

JM110

JM110 (Life Technologies) es una cepa de *E. coli* deficiente para dos metilasas (*dam* y *dcm*). Especialmente útil para el crecimiento y purificación de DNA plasmídico que será digerido con enzimas de restricción sensibles a metilación *dam* o *dcm*. En nuestro caso *ClaI*.

· **Líneas celulares**

HEK293

Línea celular derivada de riñón de embrión humano de morfología epitelial, que crece en cultivo formando monocapas de células adherentes. Adquirida de la *American Type Culture Collection* (ATCC) (CRL-1573).

· **Vectores**

pCIneo

Vector plasmídico de expresión en células de mamífero con el promotor de citomegalovirus (CMV) y el gen de resistencia a neomicina (Promega, Figura 7).

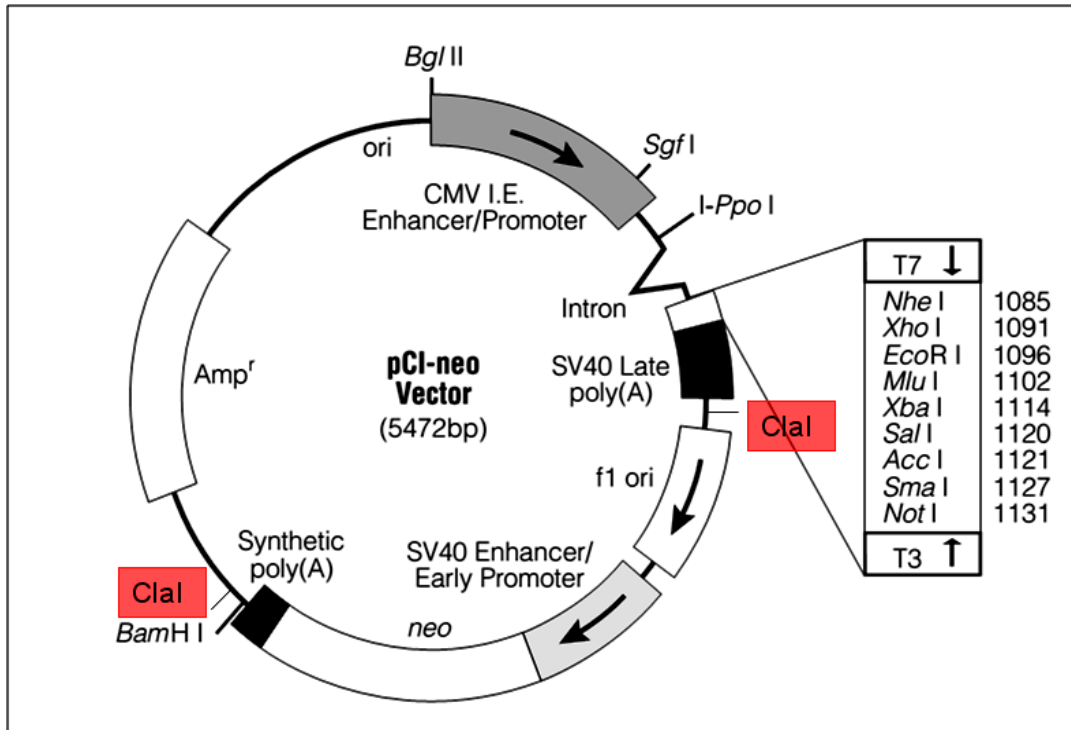


Figura 7. Esquema del vector pCIneo

phMGFP

Vector plasmídico de expresión en células de mamífero (Promega, Figura 8) que contiene el gen de la proteína verde fluorescente (*Monster Green Fluorescent Protein®*), codificada por una versión modificada del gen de la proteína verde fluorescente originalmente clonada a partir de *Montastrea cavernosa*. El vector hMGFP contiene el promotor de CMV y adicionalmente un intrón quimérico y una señal de poliadenilación tardía del virus SV40, los cuales favorecerán el correcto procesamiento del RNA de hMGFP en las células transfectadas.

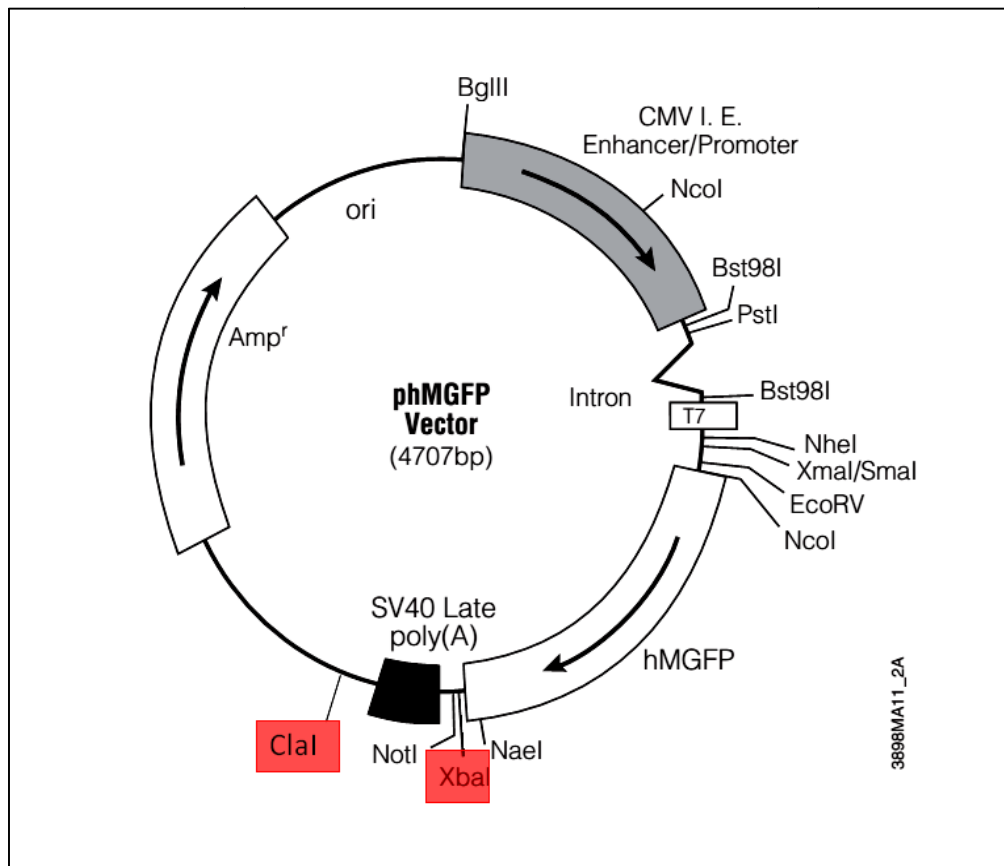


Figura 8. Esquema del vector phMGFP

Métodos

· Cultivos de bacterias

Para crecer las distintas cepas de *E. coli* se utilizaron los medios líquidos *Luria Broth* (LB) y *Terrific Broth* (TB) (ambos de Pronadisa, Madrid, España) preparados según las instrucciones del fabricante. El medio TB se utilizó para crecer bacterias que contenían la región 3'UTR de CD40LG con el microsatélite. En estos casos las bacterias se crecieron a una temperatura de 25-30 °C. En el resto de casos se crecieron a 37 °C con agitación.

Para crecer colonias en medio sólido se prepararon medios LB y TB con agar bacteriológico (Pronadisa) a una concentración final de 15 g/L en placas de Petri de 10 cm de diámetro (Soria Genlab, Madrid, España). En el caso de las placas que contenían ampicilina, ésta se añadió al medio a una concentración final de 100 µg/ml (Britapen, Laboratorio Reig-Jofré S.A., Madrid, España).

· Transformación de bacterias con DNA plasmídico

En las tres cepas de *E. coli* competentes empleadas, 50 µl de bacterias se transformaron con aproximadamente 0,25 ng de DNA plasmídico, siguiendo las instrucciones del fabricante para cada cepa. Para el almacenamiento de bacterias transformadas con diferentes plásmidos, una o varias colonias bacterianas se resuspendieron en 1 ml de medio de cultivo al 15% de glicerol (Sigma-Aldrich) y se congelaron a -80°C.

· Cultivo de líneas celulares

La línea celular HEK293 se cultivó en medio Advanced DMEM (Gibco-Life Technologies) suplementado con un 1% penicilina/estreptomicina (Gibco) y un 10% de suero fetal bovino (FBS, Life Technologies) (medio completo), en placas de poliestireno de 6 y 24 pocillos (Corning Inc., NY, EE. UU.) a 37 °C y con una atmósfera al 5% de CO₂. Se mantuvieron las placas hasta una confluencia del 80-90%. Se realizaron pases cada 3-4 días mediante tratamiento con tripsina-EDTA (Gibco) a 37 °C durante 5 min. Tras la neutralización con medio completo se sembraron nuevas placas utilizando una dilución 1/10 de la suspensión celular.

· Aislamiento de DNA genómico

El DNA genómico de pacientes y controles se obtuvo a partir de muestras de 5 ml de sangre venosa extraída en un tubo con heparina de litio (Terumo Europe NV Leuven, Bélgica). Las muestras de sangre se trataron primero con una solución de lisis de hematíes que contenía Tris-HCl 10 mM pH 7,6; MgCl₂ 5 mM y NaCl 50 mM (todos de Sigma) para obtener los leucocitos.

El DNA genómico de la línea HEK293 y de los diferentes clones obtenidos a partir de ella se aisló a partir de $2-4 \times 10^6$ células.

Para la extracción de DNA a partir de las muestras indicadas previamente se utilizó una modificación del método de Blin y Stafford¹³⁶. Las células se resuspendieron en un tampón para la lisis de leucocitos WCLB (Tris-HCl 10 mM, pH 7,6; EDTA 10 mM, pH 8; NaCl 50 mM), al que se añadieron SDS (Sigma-Aldrich, St. Louis, MO), proteinasa K y RNasa A (ambas de Roche Diagnostics GmbH, Mannheim, Alemania) a concentraciones finales del 0,16% (v/v), 160 µg/ml y 2 µg/ml respectivamente. La mezcla se incubó entre 8 y 16 h a 42 °C en agitación, tras lo cual se realizaron dos extracciones con fenol/cloroformo/alcohol isoamílico (PCI, 25:24:1, Sigma-Aldrich) seguidas de una nueva extracción con cloroformo/alcohol isoamílico (CI, 24:1). Finalmente, el DNA se precipitó en la fase acuosa con 2 volúmenes de isopropanol (Sigma-Aldrich) en presencia de NaCl 0,1 M (Sigma-Aldrich) durante una noche a 4 °C. Las muestras de DNA obtenidas se resuspendieron en tampón TE (Tris 10 mM, pH 8; EDTA 1 mM) y la concentración del DNA se midió por absorbancia y en geles de agarosa comparando las muestras con un patrón de cantidad de DNA (Low DNA Mass Ladder, Invitrogen Life Technologies).

· Aislamiento de DNA plasmídico

Para el aislamiento de DNA plasmídico se usaron los kits Plasmid Mini kit o Plasmid Plus Maxi kit (ambos de Qiagen GmbH, Hilden, Alemania), siguiendo las instrucciones del fabricante.

· **Aislamiento de RNA**

El RNA total de las muestras de células HEK293 utilizadas en los ensayos cinéticos se aisló con el kit High Pure RNA Isolation kit (Roche Diagnostics GmbH). La concentración y pureza de los RNA se midió por absorbancia en un espectrofotómetro Nanodrop 1000 (Thermo Fisher Scientific Inc. Waltham, MA EEUU).

· **Síntesis de cDNA mediante retrotranscripción (RT)**

Para la síntesis de cDNA se retrotranscribieron 600 µg de RNA total en un volumen final de 10 µl con la transcriptasa en reverso de AMV (New England Biolabs), siguiendo las instrucciones del fabricante. La retrotranscripción se hizo usando oligos dTVN, similares a los oligos dT convencionales con la diferencia de que se anclan siempre al comienzo de la cola de poli-A al llevar combinaciones de bases diferentes de timina en las dos posiciones del extremo 5'. Las muestras se incubaron 5 min a 68 °C para eliminar estructuras secundarias en el RNA, y la reacción de RT se llevó a cabo en un ciclador AB 2720 Thermal Cycler en placas de 24 tubos de 0,2 ml (Thermo Scientific) a 46 °C durante 1 h, con un paso final de 5 min a 80 °C para inactivar la retrotranscriptasa. Las RT se diluyeron 1/20 en agua destilada tras finalizar la reacción, y se almacenaron a -20 °C hasta su uso.

· **Amplificación génica mediante la reacción en cadena de la polimerasa (PCR)**

Los *primers*, mezclas de reacción y programas de todas las PCR que aparecen en este trabajo se resumen en la Tabla 3. Las PCR convencionales se llevaron a cabo en un ciclador 2720 Thermal Cycler (Applied Biosystems, Life Technologies) y Las del apartado RT-PCR y de los ensayos Taqman® en un equipo de PCR a tiempo Real 7900 HT Fast Real-time PCR (Applied Biosystems). Se realizaron PCR utilizando muestras de DNA de distintas procedencias: DNA genómico de muestras de pacientes y controles, DNA plasmídico y DNA de la línea celular HEK293.

Tabla 3: Primers, mezclas de reacción y programas de las PCR

PCR	Primers	Mezcla reacción	Programa
Tipajes			
MSCD _{4o} LG	5' FAM- CTCTTCCTCCCCAGTCTC 5'- AAGAAGAGAACTGACTAGCAACG	Pwo ; 0,2 mM dNTPs; 1,5 mM MgSO ₄	94 °C, 60 s; Nx [94 °C, 10 s + 58 °C, 30 s + 72 °C, 30 s]; 72 °C, 5 min
rs1800896 [T/C]	C_1747360_101 ¹	Taqman® Genotyping Mastermix	95 °C, 10 min; Nx [95 °C, 15 s + 60 °C, 60 s]; 60 °C, 1 min
rs1800871 [A/G]	C_1747362_10	Taqman® Genotyping Mastermix	95 °C, 10 min; Nx [95 °C, 15 s + 60 °C, 60 s]; 60 °C, 1 min
rs1800871 [T/G]]	C_1747363_10	Taqman® Genotyping Mastermix	95 °C, 10 min; Nx [95 °C, 15 s + 60 °C, 60 s]; 60 °C, 1 min
Poli-CCD _{4o} LG	5' FAM- GCCAGGCTCTAGAACGTCTAAC 5'- GAATGAATAACTGTCCGAGAGG	Pwo; 0,2 mM dNTPs; 1,5 mM MgSO ₄	94 °C, 5 min; Nx [94 °C, 30 s + 62 °C, 40 s + 72 °C, 40 s]; 72 °C, 7 min
Amplificaciones			
3'UTRCD _{4o} LG	5'- AACCCCACTGATGCAGACAT 5'- TCAATCTCCTGCACTTTTCA	Pwo; 0,2 mM dNTPs; 2 mM MgSO ₄	94 °C, 2 min; 30x [94 °C, 30 s + 58 °C, 30 s + 72 °C, 55 s]; 72 °C, 7 min
GFP-MSCD _{4o} LG	5' AGCCACAGCGGACTACC 5'- AAGAAGAGAACTGACTAGCAACG	Pwo; 0,2 mM dNTPs; 1,5 mM MgSO ₄	95 °C, 5 min; 35x [94 °C, 30 s + 57 °C, 30 s + 72 °C, 40 s]; 72 °C, 5 min
RT-PCR			
GFP ₃ UTRCA _n	5' CTGCAGAAGTTGGTCGTGAG 5' CCATCCGCAGCTTGATCTTC	Gene expression Mastermix	95 °C, 10 min; 40x [95 °C, 15 s + 60 °C, 60 s]; 60 °C, 1 min
GADPH	5' GAAGGTGAAGGTCGGAGT 5' GAAGATGGTGATGGGATTTC	Gene expression Mastermix	95 °C, 10 min; 40x [95 °C, 15 s + 60 °C, 60 s]; 60 °C, 1 min
c-fos	5' CGGGCTTCAACGCAGACTA 5' CTGGTCGAGATGGCAGTGA	Gene expression Mastermix	95 °C, 10 min; 40x [95 °C, 15 s + 60 °C, 60 s]; 60 °C, 1 min

¹Ensayos Taqman® preformados (Applied biosystems) que contienen una pareja de primers específicos y sondas de discriminación alélica. s: segundos; min: minutos; N: número variable de ciclos (indicado en cada experimento). FAM: 5'-Fluoresceína. Pwo: DNA polimerasa de *Pyrococcus woesei*.

·Análisis de polimorfismos

Las PCR MSCD40LG y Poli-C CD40LG sirvieron para el tipaje alélico de estos marcadores a partir de muestras de DNA genómico, DNA plasmídico o cDNA en distintos casos. Alícuotas de estas PCR (aproximadamente 1 ng de producto de PCR por marcador y muestra) se sometieron a electroforesis en un secuenciador capilar ABI 3700 (Applied Biosystems) en la Unidad de Genómica del Parque Científico de Madrid. Los archivos con la lectura de las electroforesis se analizaron con el programa Peak Scanner (Applied Biosystems). En las PCRs MS40L y poli-C, los tamaños de los fragmentos de PCR se transformaron en repeticiones (CA)_n ó (C)_n por comparación con la secuencia de referencia del cDNA en Genbank.

Los tres polimorfismos de base (SNP) situados en el promotor del gen de la IL-10: rs1800896, rs1800871 y rs1800872 (situados en posiciones -1087, -824 y -597 respectivamente) fueron tipados mediante PCR utilizando ensayos Taqman® (TaqMan, Applied biosystem) de discriminación alélica siguiendo las instrucciones del fabricante. Este ensayo emplea una pareja de *primers* no marcados y sondas marcadas en 5' con los agentes fluorescentes 6-carboxifluoresceína (FAM™) o VIC® y un *quencher* no fluorescente en el extremo 3'. La reacción de genotipado se llevó a cabo en placas de 486 pocillos utilizando las condiciones indicadas en Tabla 3 y partiendo de 20 ng de DNA, en un volumen total de 25 µl,

·Clonaje de la región 3'UTR de CD40LG

La región 3'UTR de CD40LG se clonó en el vector pMGFP (Promega, Madrid, España). Para ello, los fragmentos de la región 3'UTR de CD40LG se amplificaron a partir de DNA genómico de individuos de tipaje conocido utilizando las condiciones indicadas en la Tabla 3 (PCR 3'UTR40LG) y el producto de PCR se purificó con el kit MinElute PCR Purification kit (Quiagen). Tanto el producto de la PCR como el vector pMGFP, se cortaron con las enzimas de restricción XbaI y ClaI (New England Biolabs, MA, EE. UU.), y ambos se ligaron con la enzima DNA ligasa del fago T4 (New England Biolabs, MA, EE. UU.), siguiendo las instrucciones del fabricante. Las uniones se amplificaron mediante transformación de bacterias *E. coli* competentes de las cepas DH5-α (Invitrogen) o JM110 (Promega) siguiendo las instrucciones del fabricante en cada caso. La cepa JM110 tiene inactivada la enzima *dam* metilasa, y se usó en los casos en que había que desmetilar dianas ClaI que estaban

boqueadas para el corte por la metilación *dam*. Las bacterias transformadas se cultivaron en medio TB con ampicilina a 100 µg/ml para el aislamiento de colonias. La presencia del inserto en las colonias se comprobó mediante el análisis de los fragmentos de corte con enzimas de restricción y con la PCR GFP-MSCD40LG (Tabla 3).

Para introducir la resistencia a neomicina en los vectores pHMGFP con los fragmentos 3'UTR insertados (vectores GFP-3UTR.CAn), el vector pCI-Neo transformado en la cepa JM110 de *E.coli* se cortó con ClaI y se aisló el fragmento del corte que contenía la resistencia a partir de geles de agarosa con el kit Gel Extraction Kit (Qiagen). El fragmento aislado se unió con la ligasa del fago T4 a las construcciones GFP-3UTR.CAn cortadas con ClaI y tratadas con fosfatasa alcalina (Alkaline phosphatase, shrimp; Roche Diagnostics GmbH) para evitar la recircularización de los vectores. De nuevo se crecieron colonias en medio con ampicilina, y se comprobó la presencia de la resistencia a neomicina mediante el análisis de los fragmentos de corte con enzimas de restricción.

· **Secuenciación de DNA plasmídico**

Los vectores GFP-3UTR.CAn se secuenciaron con el *primer* 5' de la PCR GFP-MSCD40LG, que se une al extremo 3' de la secuencia codificante de la GFP en la Unidad de Genómica del Parque Científico de Madrid en un secuenciador ABI 3700 (Applied Biosystems).

· **Transfección y obtención de transfectantes estables**

Para obtener las condiciones adecuadas de transfección, las células HEK293 se crecieron en medio completo en placas de 24 pocillos hasta alcanzar una confluencia del 80-90%. Se mezclaron 250 µg de los vectores GFP-3UTR.CAn con 5 µl del compuesto Saint-18 (Synvolux Therapeutics B.V., Groningen, Países Bajos) siguiendo las instrucciones del fabricante. Tras 24 h se cambió el medio y se reemplazó por 1 ml de medio completo suplementado con 400 µg/ml del antibiótico G418 (Sigma) para la selección de colonias de transfectantes con el DNA plasmídico integrado en el genoma de forma estable. Las colonias

estables se traspasaron a placas nuevas para establecer clones con distintas construcciones integradas. Los clones aislados se mantuvieron en cultivo con 250 µg/ml de G418.

· Cinéticas de caída del mRNA

Los clones de células HEK293 con las distintas construcciones GFP-3UTR.CAn se cultivaron en placas de 24 pocillos hasta llegar al 80-90% de confluencia. Antes de comenzar la cinética, se retiró el medio y se añadió 1 ml de medio completo a cada pocillo. Tras 1 h con el medio fresco a 37 °C se paró la transcripción añadiendo actinomicina D (Act-D, Sigma) a una concentración final de 10 µg/ml. Trascorridos 0, 30, 60 ó 90 min desde la adición de Act-D las células se lavaron dos veces con PBS (Iberlabo), se recogieron y se centrifugaron 1 min a 7000 g para almacenarlas en *pellet* a -80°C hasta su uso posterior.

· Medida de la cantidad de mRNA en las cinéticas de caída

La cantidad de mRNA de las construcciones GFP-3UTR.CAn se midió por RT-PCR. Se amplificaron mediante PCR 2,5 µl de las RT (previamente diluidas 1/20) en un volumen final de reacción de 12,5 µl utilizando parejas de *primers* para las PCR GFP3UTR, GADPH o c-fos (Tabla 3) en reacciones independientes. Se utilizó una mezcla de reacción (Taqman Gene Expression Mastermix, Applied Biosystems) que lleva la DNA polimerasa Taq Gold y todos los componentes necesarios para la reacción de PCR, a excepción del DNA molde y los *primers*. A la mezcla se le añadió el compuesto fluorescente Sybr Green I (Sigma-Aldrich) a una dilución final de 1/50000 para la cuantificación de los fragmentos amplificados en un equipo de PCR a tiempo real 7900 HT Fast Real-time PCR (Applied Biosystems). Las reacciones se llevaron a cabo por duplicado en placas de 384 pocillos (Thermo) en la Unidad de Genómica del Parque Científico de Madrid.

Herramientas informáticas

Primer 3 (<http://www-genome.wi.mit.edu/cgi-bin/primer/primer3>)

El programa Primer 3, creado por el Whitehead Institute for Biomedical Research (Cambridge, MA, EE. UU.), permite diseñar y contrastar los oligonucleótidos con varias

librerías génicas y ajustar parámetros tales como T_m , porcentaje de GC, longitud, autocomplementariedad y complementariedad de los extremos 3'.

Oligo Analyzer (<https://eu.idtdna.com/analyzer/Applications/OligoAnalyzer/>)

Plataforma on-line para el diseño y análisis de *primers* y sondas. Creado por Integrated DNA Technologies, Inc.

WebCutter (<http://rna.lundberg.gu.se/cutter2/>)

Programa para la creación de mapas de restricción de secuencias nucleotídicas.

NCBI (<http://www.ncbi.nlm.nih.gov>)

El NCBI (National Centre for Biotechnology Information), creado como la rama informática de la NLM (National Library of Medicine) del NIH (National Institutes of Health), aglutina bases de datos de secuencias de genes y cDNA del proyecto genoma humano, y del resto de organismos secuenciados hasta la fecha, así como bases de datos de proteínas y de expresión génica, con un gran número de herramientas informáticas para el análisis de la información conectada entre las diferentes bases de datos.

Ensembl (<http://www.ensembl.org/Multi/blastview>)

Ensembl es un proyecto conjunto que comprende EBI (European Bioinformatics Institute). Es el equivalente europeo del NCBI.

RBPmap (<http://rbpmap.technion.ac.il/>)

Servidor web desarrollado por el Instituto de Tecnología israelí Technion en colaboración con la Universidad de California para la predicción y localización de sitios de unión de proteínas que unen RNA¹³⁷. Se basa en un algoritmo que considera la tendencia a la agrupación de los sitios de unión y la tendencia de conservación de las regiones reguladoras.

RBPDB (<http://rbpdb.ccb.utoronto.ca/index.php>)

RNA Binding Protein Database, desarrollada por el departamento de Genética Molecular de la Universidad de Toronto (Ontario, Canadá)⁽¹³⁸⁾, es una colección de proteínas con motivos de unión a RNA conectada a bases de datos de ensayos de unión a RNA revisados y publicados.

Chromas Lite v2.33. (<http://technelysium.com.au>).

Permite observar el gráfico de la secuenciación del fragmento de DNA, y obtener la secuencia para analizarla cuando no es necesario el alineamiento de múltiples secuencias.

EMBOSS merger (<http://emboss.bioinformatics.nl/cgi-bin/emboss/merger>)

El *European Molecular Biology Open Software Suite* ¹³⁹ es un servidor para el análisis de secuencias *on-line*. Una de las herramientas integradas en esta plataforma, “*merger*”, desarrollada por Gary Williams del Centro de Investigación Genómica Rosalind Franklin (Wellcome Trust Sanger Institute, Cambridgeshire, Reino Unido), se utilizó para el alineamiento de secuencias.

Peak Scanner Software v1.0

Permite el análisis de datos generados utilizando los analizadores genéticos de *Applied Biosystems*®. Muestra el perfil de separación de fragmentos de DNA y calcula el tamaño de los fragmentos.

Mfold

Servidor web desarrollado por la escuela de medicina de la Universidad de Washington que reúne varias herramientas para la predicción de estructuras secundarias de ácidos nucleicos de cadena sencilla ¹⁴⁰ (<http://www.bioinfo.rpi.edu/applications/mfold>)

Ras Mol (<http://rasmol.org/>)

Es un programa para la visualización gráfica de estructuras moleculares originalmente desarrollado por Roger Sayle ¹⁴.

Genepop v4.2 (http://genepop.curtin.edu.au/genepop_op1.html)

Paquete de *software* de análisis genético de poblaciones *on-line*, desarrollado por Michel Raymond y Francois Rousset en el Laboratoire de Genetique et Environment (Montpellier, France). Se ha empleado para la estimación y comparación estadística de frecuencias alélicas y genotípicas en múltiples marcadores mediante la prueba exacta de Fisher, y para las estimaciones del equilibrio Hardy-Weinberg.

GraphPad Prism v5

GraphPad Software, Inc. (La Jolla, CA, EE. UU.). Permite la construcción de gráficos científicos, ajuste de curvas y análisis estadístico. Se empleó para el estudio de la asociación alélica o genotípica utilizando tablas de contingencia. Se utilizó la prueba de χ^2 o prueba exacta de Fisher, en caso necesario. La magnitud de asociación entre las diferentes variables se expresó por medio de la razón de ventajas *odds ratio* (OR) con intervalos de confianza al 95%. El ajuste de datos a una línea recta se realizó utilizando la regresión de mínimos cuadrados no lineal. Se empleó la prueba F de suma de cuadrados para la comparación de las pendientes de distintas rectas entre sí. El nivel de significación estadística utilizado ha sido de $p < 0.05$.

G*power 3 (<http://www.gpower.hhu.de>)

G*Power es una herramienta gratuita disponible *on line* para el cálculo del poder estadístico de diferentes tests (*t* tests, *F* tests, χ^2 tests...) ¹⁴¹. Los cálculos del poder estadístico ($1-\beta$) y de las OR mínimas detectables en función del tamaño de muestra están obtenidos con este programa.

Desarrollo de un cálculo para estimar la penetrancia y el riesgo relativo para un genotipo a partir de datos de un estudio de casos y controles

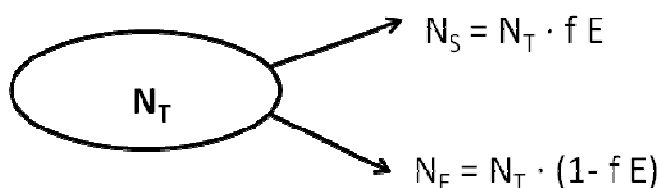
A continuación se expone el desarrollo de un cálculo sencillo que hemos elaborado y que permite estimar las penetrancias y los riesgos relativos (RR) para genotipos de un sistema bialélico (AA, Aa, aa), o para la presencia de un factor de riesgo dado, a partir de datos obtenidos en un estudio de casos y controles. Este cálculo hasta donde sabemos no ha sido publicado anteriormente.

El riesgo relativo (RR) se define como la razón de la probabilidad de enfermedad en un sujeto en función de la presencia o ausencia de un determinado factor de riesgo. En nuestro caso, el RR de un genotipo del MSCD40LG, por ejemplo 2424 en AR, sería la frecuencia de AR en mujeres homocigotas 2424 dividido por la frecuencia de AR en mujeres que no son 2424. El RR da una medida de la magnitud del efecto de una variante alélica de riesgo en la probabilidad de aparición de enfermedad. En un estudio de casos y controles, el número de casos y controles portadores de cada genotipo de interés es elegido a priori, normalmente de forma que haya un número similar de casos y controles, con lo que no están en la misma proporción en la que se encuentran en la población general. Esto hace que las

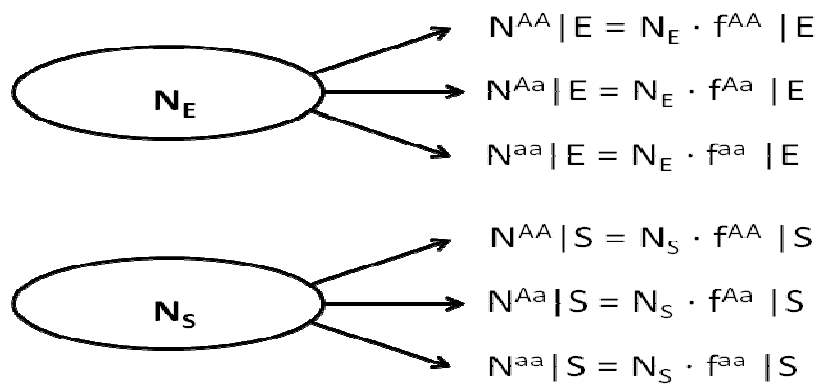
penetrancias (frecuencia de enfermedad) para cada genotipo no se puedan estimar de forma directa a partir de la muestra de estudio, por lo que no es posible calcular el RR de acuerdo con su definición matemática ^{142,143}. Esta es la razón de que, en estudios con un diseño de casos y controles, para estimar la fuerza de la asociación a enfermedad de los distintos genotipos o alelos se usen mayoritariamente las *odds ratio* (OR). Las OR o “razón de ventajas”, se definen como el cociente de las *odds* (“ventajas”) para la presencia y ausencia de un factor de riesgo. Las *odds* a su vez se definen como el cociente entre la probabilidad de éxito y la probabilidad de fracaso en un suceso binomial. Las *odds*, y por tanto las OR, se pueden calcular fácilmente a partir de los datos obtenidos en un estudio de casos y controles, y cuando la probabilidad de éxito es baja, como ocurre por ejemplo en AR o LES, donde la probabilidad de “éxito” sería la aparición del suceso “enfermedad”, las OR se aproximan a los RR, y tienden hacia el RR según disminuye la probabilidad del suceso. Esto hace que las OR sean el valor más utilizado en estudios con diseño de casos y controles para medir la fuerza de la asociación a enfermedad conferida por un factor de riesgo.

Si bien la frecuencia de enfermedad para un genotipo de un marcador de interés no es conocida a priori, la frecuencia general de la enfermedad en la población sí suele ser una cifra de estimación conocida, que puede encontrarse en publicaciones y boletines estadísticos de gobiernos y organizaciones nacionales e internacionales. A partir de la frecuencia de la enfermedad en la población, y de las frecuencias genotípicas obtenidas en un estudio de asociación de casos y controles, se pueden calcular las penetrancias y los riesgos relativos para los genotipos de la siguiente manera:

Supongamos una población con un total de N_T individuos. Una fracción de esa población desarrollará una enfermedad a la que denominamos de forma genérica E. Podemos dividir la población en individuos sanos (N_S) e individuos con la enfermedad E (N_E), donde el número de enfermos es igual al número total de individuos multiplicado por la frecuencia de la enfermedad (f^E):



Si tenemos un gen G , con las variantes alélicas A y a , que confieren distinta probabilidad de desarrollar la enfermedad E , podemos dividir a las dos subpoblaciones de sanos y enfermos en función de los tres genotipos para los alelos A y a del gen G , AA , Aa y aa :



Por definición, el riesgo relativo para un genotipo g_i es:

$$RR_{g_i} = \frac{f^E | g_i}{f^E | \text{Resto}} \quad ; f^E | g_i : \text{frecuencia de enfermedad en individuos con el genotipo } g_i$$

La frecuencia de enfermedad para un genotipo g_i ($f^E | g_i$) en la población sería:

$$1. \quad f^E | g_i = \frac{N^E | g_i}{N_{g_i}}$$

Es decir, el número de individuos con el genotipo g_i que tienen la enfermedad E, dividido por el número total de individuos con el genotipo g_i . El total de individuos con el genotipo g_i puede dividirse entre sanos y enfermos:

$$2. N^{g_i} = N^S|g_i + N^E|g_i, \text{ donde}$$

$$3. N^S|g_i = N_T \cdot f^{g_i}|S \cdot (1 - f^E), \text{ y}$$

$$4. N^E|g_i = N_T \cdot f^{g_i}|E \cdot f^E$$

Sustituyendo 2., 3. y 4. en 1.:

$$5. f^E|g_i = \frac{N_T \cdot f^{g_i}|E \cdot f^E}{N_T \cdot f^{g_i}|S \cdot (1 - f^E) + N_T \cdot f^{g_i}|E \cdot f^E}$$

Que resulta:

$$5. \text{ Penetrancia : } f^E|g_i = \frac{f^{g_i}|E \cdot f^E}{f^{g_i}|S \cdot (1 - f^E) + f^{g_i}|E \cdot f^E}$$

En 5. aparece la frecuencia de enfermedad para un genotipo g_i , es decir, la penetrancia de la enfermedad para ese genotipo, expresada en función de las frecuencias genotípicas en sanos y enfermos. El denominador en 5. es igual a la frecuencia del genotipo g_i en la población total, así que de forma abreviada podemos expresar la penetrancia del genotipo g_i como:

$$5.1. f^E|g_i = \frac{f^{g_i|E}}{f^{g_i|Pob}} \cdot f^E ; \quad f^{g_i|Pob} = f^{g_i|S} \cdot (1 - f^E) + f^{g_i|E} \cdot f^E$$

De igual manera se puede calcular la frecuencia de enfermedad en individuos que no tienen el genotipo g_i , cambiando las frecuencias de g_i en sanos y enfermos por $f \cdot (1 - g_i)$, que son las frecuencias del resto de genotipos:

$$6. f^E|Resto = \frac{(1 - f^{g_i|E}) \cdot f^E}{(1 - f^{g_i|S}) \cdot (1 - f^E) + (1 - f^{g_i|E}) \cdot f^E}$$

Las fórmulas 5. y 6. son el numerador y denominador de la fórmula para el RR.

$$7. \text{ Riesgo Relativo: } RR = \frac{f^{g_i|E} \cdot f^E / f^{g_i|S} \cdot (1 - f^E) + f^{g_i|E} \cdot f^E}{(1 - f^{g_i|E}) \cdot f^E / (1 - f^{g_i|S}) \cdot (1 - f^E) + (1 - f^{g_i|E}) \cdot f^E}$$

De esta forma, con datos que se obtienen en un estudio de casos y controles (frecuencias de los genotipos en sanos y enfermos) y el valor de la frecuencia de la enfermedad en la población, se pueden calcular los riesgos relativos asociados a los diferentes genotipos.

Este mismo cálculo se puede aplicar para estimar el RR conferido por cualquier factor de riesgo, sustituyendo la frecuencia de los genotipos en sanos y enfermos por la frecuencia de exposición en sanos y enfermos al factor de riesgo dado.

El desarrollo se puede aplicar también al cálculo de las OR. Las OR se definen como:

$$OR = \frac{(f^E|g_i / (1 - f^E|g_i))}{(f^E|Resto / (1 - f^E|Resto))}$$

Sustituyendo $f^E|g_i$ y $f^E|Resto$ por 5. y 6., podemos calcular las OR para el genotipo g_i . Este cálculo da un resultado idéntico al cálculo tradicional a partir de los datos representados en tablas de contingencia.

Las penetrancias y los RR que aparecen en los resultados del estudio de asociación genética de este trabajo se han obtenido utilizando estos cálculos. Para las dos enfermedades con las que tratamos en este trabajo, según el estudio EPISER 2000, la prevalencia de AR es de un 0,8 % en mujeres y un 0,2 % en hombres. EN LES, la prevalencia según el EPISER 2000 es de un 0,1 %, pero la muestra de este estudio no es suficiente para estimar con precisión la prevalencia de una enfermedad de tan baja frecuencia (el IC 95 % es de 30-390 casos por 100000 habitantes). No hay datos de prevalencia por sexos (sólo hay dos casos de LES en el EPISER 2000 y los dos son mujeres). Basándonos en un estudio de prevalencia de LES realizado en Asturias en 2003 ¹⁴⁴ con un total de 367 casos registrados, la prevalencia en mujeres es de aproximadamente 6 en 10000, y en hombres de 0,8 en 10000. Para los cálculos hemos utilizado los valores de 0,8% en AR y 0,06% en LES.

RESULTADOS

· Asociación del MSCD40LG a AR y LES

Nuestros trabajos previos ^{98,132} habían mostrado la existencia de una asociación del MSCD40LG a LES y AR, que en los dos casos se daba sólo en mujeres y con una menor frecuencia del alelo 24CA en ambas patologías comparada con controles sanos. En función de la presencia de alelos 24CA frente a alelos diferentes, los dos trabajos presentaban diferencias. En LES había una mayor frecuencia de enfermas portadoras de alelos mayores de 24CA comparado con mujeres sanas, mientras que en AR no se daba esta diferencia, pero sí una mayor frecuencia de enfermas homocigotas para alelos distintos del 24CA, y a su vez una menor frecuencia de homocigotas 2424.

· Series iniciales de pacientes y controles

Para comparar mejor entre sí los resultados previos mencionados, en el presente trabajo hemos analizado los datos genéticos de que disponíamos (dos series iniciales de pacientes: LESC1 y ARC1 y una serie de controles DSC1, Tabla 7) agrupando todos los alelos distintos del 24CA en un grupo (X), como se hizo en el estudio de asociación a AR, de manera que se forman tres genotipos (2424, 24X y XX) en función de la presencia de dos, una o ninguna copias del alelo 24CA. De esta manera los “alelos” 24CA y X se comportan como un sistema bialélico A/a, y se encuentran en equilibrio de *Hardy-Weinberg* en la serie de controles DSC1. Así comprobamos que las frecuencias genotípicas (Figura 8) y los riesgos relativos (RR) (Tabla 4) son muy similares en ambas enfermedades para el alelo 24CA y los tres genotipos 2424, 24X y XX.

Figura 8. Frecuencias genotípicas del MSCD40LG en las series iniciales de mujeres sanas (DSC1), pacientes de LES (LESC1) y AR (ARC1). Se muestran las frecuencias de los genotipos 2424, 24X y XX formados uniendo los alelos distintos del alelo 24CA en un grupo (X). DSC1: N=80; LESC1: N=80; ARC1: N=140; * $p < 0,05$ en la prueba de χ^2 .

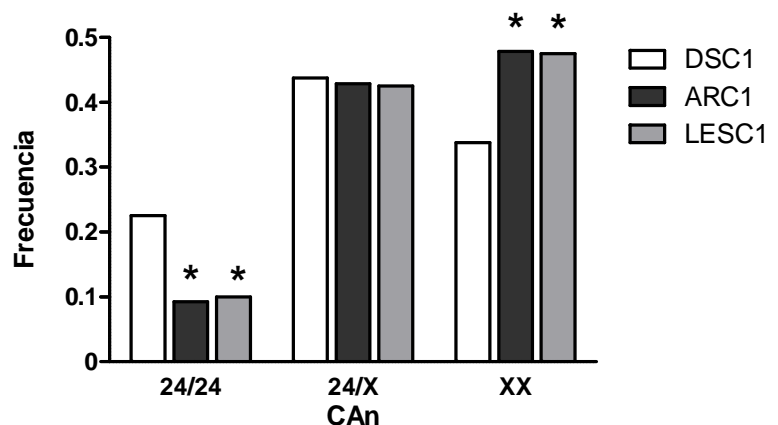


Tabla 4. Riesgos relativos en los estudios previos para el alelo 24CA y los genotipos 2424 24X y XX en LES y AR

	Genotipo			Alelo
	XX	24X	2424	24CA
LES	1,78	0,95	0,38	0,57
AR	1,79	0,96	0,35	0,56

Por la menor frecuencia de mujeres homocigotas 2424 en AR frente a controles, pero no de heterocigotas 24X (Figura 8), inicialmente interpretamos el alelo 24CA como alelo protector cuando se encontraba en homocigosis ⁹⁸, lo que indicaba un carácter recesivo del alelo 24CA. Sin embargo, la relación de RR entre los genotipos 2424, 24X y XX (Tabla 4) que observamos ahora ofrece una interpretación alternativa, ya que cada alelo X presente parece aumentar de forma proporcional el RR, de forma muy similar en LES y AR. Así, desde el punto de vista del alelo 24CA éste se comporta como alelo protector, ya que el RR de AR y LES disminuye con cada alelo 24CA presente, pero no sólo cuando se encuentra en homocigosis. En la Tabla 5 aparece un esquema de la relación matemática de RR entre genotipos en los modelos más comunes de penetrancia para un sistema bialélico (basado en ¹⁴⁵). A la vista de los datos anteriores los RR de la asociación del MSCD40LG a LES y AR en las series estudiadas inicialmente se ajustarían a un modelo multiplicativo, de forma que con cada alelo de riesgo (alelos X) se multiplica el RR en un factor (γ) que resulta muy parecido para LES y AR (tabla 6).

Tabla 5. Modelos básicos de penetrancia para variantes bialélicas

Modelo	Penetrancia			Riesgo Relativo		
	aa	Aa	AA	Aa → aa	AA → aa	AA → Aa
Multiplicativo	f_0	$f_0\gamma$	$f_0\gamma^2$	γ	γ^2	γ
Aditivo	f_0	$f_0\gamma$	$f_02\gamma$	γ	2γ	2
Recesivo	f_0	f_0	$f_0\gamma$	1	γ	γ
Dominante	f_0	$f_0\gamma$	$f_0\gamma$	γ	γ	1

f_0 : penetrancia en el genotipo de menor riesgo (a); γ : aumento del RR con la presencia de cada alelo de riesgo (A). Basado en ¹⁴⁵.

Tabla 6. Penetrancias, razón entre los RR de los genotipos y valor de γ en las series iniciales de LES y AR

Series	Modelo	Penetrancia ¹			Riesgo Relativo ²			$\bar{\gamma}$ ³
		2424 (aa)	24X (Aa)	XX (AA)	24X → 2424	XX → 2424	XX → 24X	
LESC1	Multiplicativo	$2,7 \cdot 10^{-4}$	$5,8 \cdot 10^{-4}$	$8,4 \cdot 10^{-4}$	2,18	3,15	1,45	1,80
ARC1	Multiplicativo	$3,3 \cdot 10^{-3}$	$7,8 \cdot 10^{-3}$	$1,1 \cdot 10^{-2}$	2,36	3,40	1,44	1,88

¹ Frecuencia de enfermedad en el genotipo indicado (cálculo en Materiales y Métodos). ² Razón entre los RR de los genotipos indicados en las series LES C1 y ARC1 frente a DSC1. ³ Promedio de los tres valores γ estimados con los RR de los tres genotipos ($\gamma_1 = RR_{24X \rightarrow 2424}$; $\gamma_2 = \sqrt{RR_{XX \rightarrow 2424}}$; $\gamma_3 = RR_{XX \rightarrow 24X}$)

· Nuevas series de sanos y pacientes de LES

Posteriormente, y con el objeto de aumentar el tamaño de la muestra de estudio, obtuvimos cuatro nuevas series de muestras de DNA genómico: dos de pacientes de LES canarias y peninsulares (LESC2 y LESP1 en la Tabla 7) y otras dos de donantes sanas, canarias y peninsulares (DSC2 y DSP2 en la Tabla 7). En materiales y métodos se detallan los criterios de selección de muestras y procedencia de las mismas. Para el presente trabajo se realizó el tipaje del MSCD40LG (PCR MSCD40LG en Tabla3) en las nuevas muestras y revisamos la asociación a enfermedad en el conjunto total de muestras disponibles.

Tabla 7. Series de DNA de controles y pacientes de LES y AR

Series iniciales	N	Procedencia	Series nuevas	N	Procedencia	Series Ampliadas	N
DSC1	80	2	DSC2	73	2	DSC12	153
LESC1	80	2	LESC2	85	2	LESC12	165
ARC1	140	2	-				
DSP1	241	1	DSP2	54	1	DSP12	295
			LESP1	26	3		

N: número de mujeres; DSP: donantes sanas peninsulares; DSC: donantes sanas canarias; LESP: pacientes de LES peninsulares; LES C: pacientes de LES canarias; ARC: pacientes de AR canarias.

Primero comparamos las nuevas series DSP2, DSC2 y LES C2 con sus correspondientes series iniciales (DSP1, DSC1 y LES C1) para comprobar que eran homogéneas entre sí. Todas las series de controles estaban en equilibrio de *Hardy-Weinberg* para las variantes 24CA/X (datos no mostrados). No encontramos diferencias significativas

en las frecuencias genotípicas del MSCD40LG en ninguna de las tres parejas de muestras cuando se compararon con el programa Genepop v 4.2 (datos no mostrados). Aún así, en las series LESC1 y LESC2 era llamativa la diferencia en la frecuencia de homocigotas 2424 (cercana a la significación estadística, $p=0,056$), que en LESC2 era similar a la de la población sana (Figura 9). Por ello ampliamos la comparación entre las series LESC1 y LESC2 utilizando tres polimorfismos de sustitución de base (SNP) situados en el promotor del gen de la IL-10: rs1800896 (-1087), rs1800871 (-824) y rs1800872 (-597) que habían sido genotipados previamente por nuestro grupo en la serie LESC1 ¹⁴⁶. Para el presente trabajo genotipamos esos tres SNP en la serie LESC2 y se realizaron pruebas de diferenciación génica y genotípica con el programa Genepop v4.2. utilizando los cuatro marcadores (Tabla 8). Los resultados indicaban que ambas series eran homogéneas, por tanto, unimos las series nuevas con las iniciales correspondientes formando las series ampliadas: DSP12 (DSP1+DSP2), DSC12 (DSC1+DSC2) y LESC12 (LESC1+LESC2).

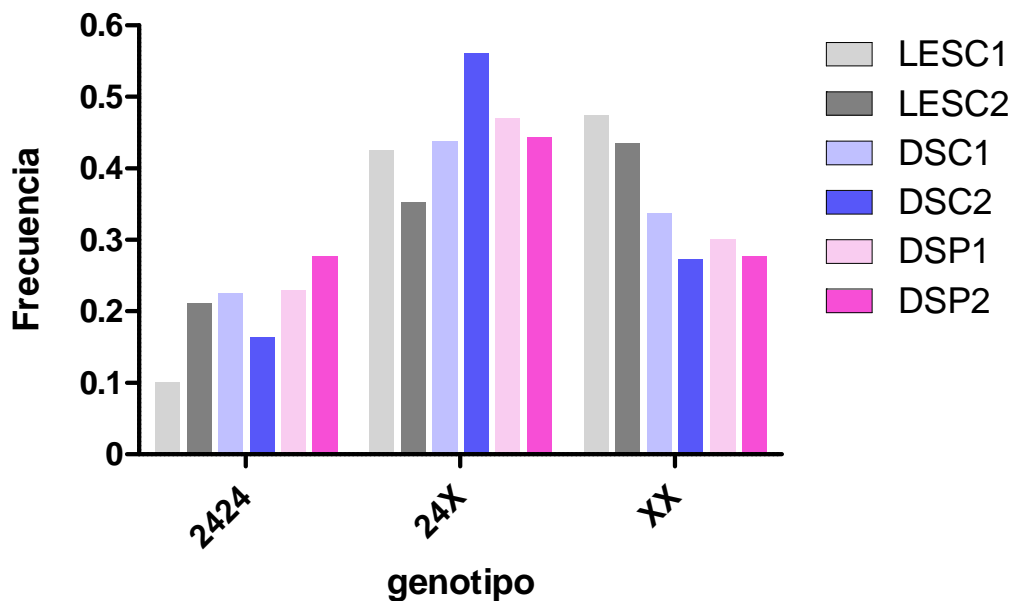


Figura 9: Frecuencias genotípicas en las series iniciales y nuevas en población canaria. Se muestran las frecuencias de los genotipos 2424, 24X y XX formados uniendo los alelos distintos del alelo 24CA en un grupo (X). DSC1. Se muestra la p en la prueba de χ^2 .

Tabla 8. Comparación de frecuencias alélicas y genotípicas entre LESC1 y LESC2

Localización	Gen	Marcador	Grupo	alelo 1/ alelo 2	1	2	1/1	1/2	2/2	p ¹	P ³
Cromosoma X 135.730.352	CD40LG	MSCD40LG	LESC1	24CA / X	50	110	8	34	38	0,163	0,195
			LESC2	24CA / X	66	104	18	30	37		
Cromosoma 1 206.946.897	IL-10	rs1800896	LESC1	A / G	88	72	26	36	18	0,377	0,376
			LESC2	A / G	100	66	30	40	13		
Cromosoma 1 206.946.634	IL-10	rs1800871	LESC1	T / C	47	113	7	33	40	0,902	0,902
			LESC2	T / C	46	120	6	34	43		
Cromosoma 1 206.946.407	IL-10	rs1800872	LESC1	A / C	45	115	6	33	41	0,900	0,901
			LESC2	A / C	45	119	6	33	43		
Todos los marcadores										0,648 ³	0,687 ⁴

Se muestran los valores p en el test exacto de Fisher para comparación de

¹ frecuencias alélicas de cada marcador

² frecuencias alélicas de los 4 marcadores en conjunto

³ frecuencias genotípicas de cada marcador

⁴ frecuencias genotípicas de los 4 marcadores en conjunto

Con las series ampliadas primero estudiamos la asociación a enfermedad del MSCD40LG en los pacientes de LES peninsulares (LESP1) y comprobamos que se replicaban los resultados encontrados previamente en población canaria, con una menor frecuencia del alelo 24CA en pacientes de LES (0,29) que en controles (0,45) ($p=0.0196$, OR (IC95%) = 0.48 (0.26-0.89); Figura 10). Las frecuencias de los genotipos 2424, 24X y XX también eran similares a las encontradas en pacientes de LES canarias con una frecuencia significativamente menor de homocigotas 2424 en pacientes respecto a controles (0.04 vs 0.24; $p=0.014$ OR (IC95%) = 0.13 (0.017-0.97)). A pesar de que en las series de población peninsular la asociación del MSCD40LG resulta similar a lo encontrado anteriormente en población canaria, el tamaño de muestra de la serie LESP1 es bajo, lo que limita la fiabilidad del resultado.

La frecuencia casi idéntica del alelo 24CA (y por tanto de los genotipos 2424, 24X y XX) encontrada en ambas series de controles (sanas canarias y peninsulares) invitaba en principio a unir las muestras en dos únicas series de controles sanas y enfermas de LES. Esta unión permitiría obtener unos tamaños de muestra considerables (LESP1+LESC12 = 191 enfermas, DSP12+DSC12 = 448 sanas) especialmente en el grupo control, con los que teóricamente se podría detectar con buena fiabilidad estadística ($\alpha=\beta=0,05$) una asociación con una OR $\leq 0,65$ para el alelo 24CA, y con una OR $\geq 1,90$ para el genotipo XX.

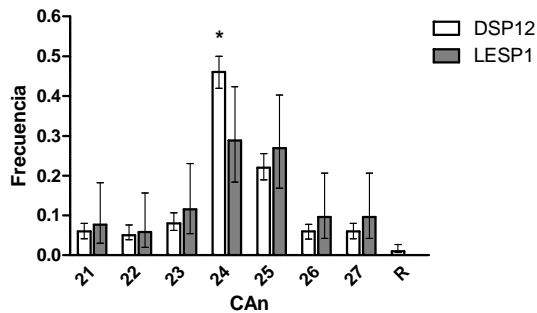
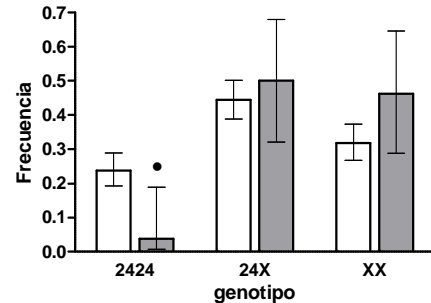
A**B**

Figura 10. Distribución alélica y genotípica del MSCD40LG en controles y pacientes de LES peninsulares. (A) Frecuencias alélicas en las series DSP12 (N=590 alelos) y LESP1 (N=52 alelos). (B) Frecuencias genotípicas en las series DSP12 (N=295) y LESP1 (N=26). R: Alelos minoritarios agrupados. Las barras representan el intervalo de confianza al 95% de la frecuencia alélica. * $p < 0,05$ y • $p < 0,05$ en la prueba de χ^2 . Las barras representan el intervalo de confianza al 95% de la frecuencia alélica.

Sin embargo, aunque al agrupar los alelos en función del alelo 24CA las frecuencias genotípicas entre sanos canarios y peninsulares son muy similares, previamente en nuestro laboratorio habíamos comprobado que la distribución alélica de este polimorfismo difiere entre ambas poblaciones, en concreto en la frecuencia de dos alelos, 22 CAs y 27 CAs¹²⁸. Por otra parte se ha descrito que las poblaciones canaria y peninsular tienen diferencias en la proporción de ancestría¹⁴⁷ que podrían potencialmente confundir el resultado de estudios de asociación genética con muestras de ambas procedencias¹⁴⁸. Por tanto, revisamos la asociación genética del MSCD40LG sólo en población canaria utilizando las series ampliadas de controles sanas (DSC12, N=153) y pacientes de LES (LESC12, N=165) y la serie inicial de AR (ARC1, N=140) (Tabla 7). En la Figura 11 se muestran las frecuencias alélicas y genotípicas en ambas enfermedades frente a las de controles sanas. Los valores numéricos y la comparación estadística se recogen en la Tabla 9.

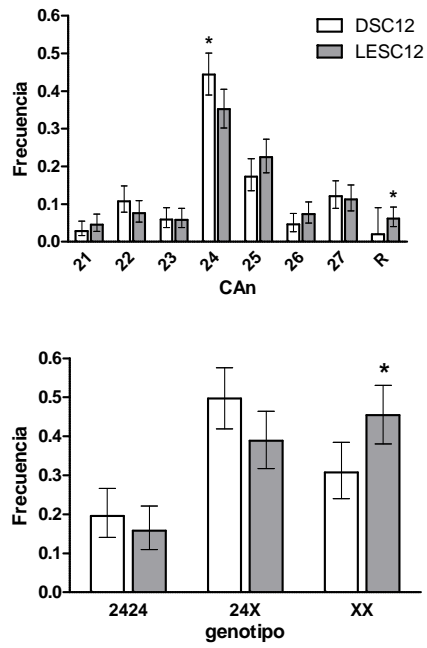
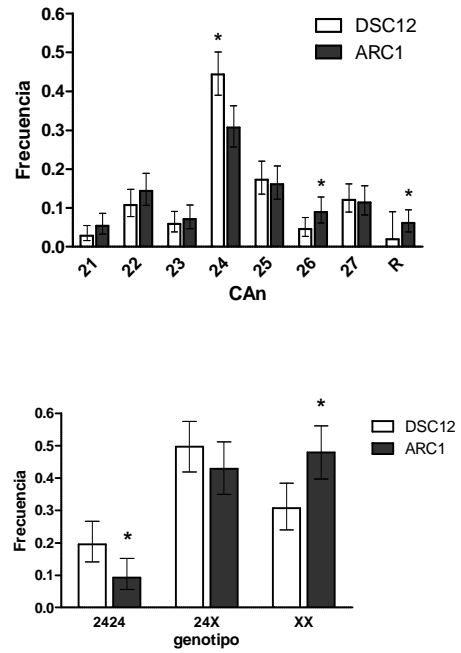
A**B**

Figura 11. Distribución de frecuencias alélicas y genotípicas del MSCD40LG en controles sanas y pacientes de LES y AR en las series ampliadas. (A) frecuencias alélicas (arriba) y genotípicas (abajo) en las series DSC12 (N=153 mujeres) y LESC12 (N=203 mujeres). (B) frecuencias alélicas (arriba) y genotípicas (abajo) en las series DSC12 (N=153 mujeres) y ARC1 (N=140 mujeres). *: $p < 0,05$ en la prueba de χ^2 . Las barras representan el intervalo de confianza al 95% de la frecuencia alélica.

Tabla 9. Comparación de frecuencias alélicas y genotípicas del MSCD40LG en las series ampliadas de controles y pacientes de LES y AR

	DSC12		LESC12						ARC1					
	N ¹	F ²	N	f	p ³	(1-β) ⁴	RR ⁵	OR (IC 95%) ⁶	N	f	p	(1-β)	RR	OR (IC 95%)
Alelo														
21	9	0,03	15	0,05	ns				15	0,05	ns			
22	33	0,11	25	0,07	ns				40	0,14	ns			
23	18	0,06	19	0,06	ns				20	0,07	ns			
24	136	0,44	116	0,39	0,02	0,64	0,68	0,68 (0,49-0,93)	86	0,31	0,0007	0,92	0,56	0,55 (0,39-0,78)
25	53	0,17	74	0,19	ns				45	0,16	ns			
26	14	0,05	24	0,09	ns				25	0,09	0,046	0,49	2,03	2,05 (1,04-4,02)
27	37	0,12	37	0,08	ns				32	0,11	ns			
R ⁷	6	0,02	20	0,06	0,009	0,73	3,22	3,22 (1,28-8,14)	17	0,06	0,02	0,57	3,18	3,23 (1,26-8,32)
Genotipo														
2424	30	0,20	26	0,16	ns				13	0,09	0,013	0,72	0,42	0,42 (0,21- 0,84)
24X	76	0,50	64	0,39	ns				60	0,43	ns		0,76	
XX	47	0,31	75	0,45	0,008	0,76	1,88	1,88 (1,18-2,98)	67	0,48	0,003	0,84	2,06	2,07 (1,28-3,34)
Total	153		165		0,023 ⁸				140		0,004 ⁸			

ns: no significativo. ¹Número de alelos/genotipos. ²Frecuencia del alelo/genotipo. ³Valor p de la prueba χ^2 en la comparación individual de frecuencias de alelos o genotipos en tablas 2x2. ⁴ Poder estadístico. ⁵ Riesgo Relativo (cálculo en Materiales y Métodos). ⁶OR con el intervalo de confianza al 95% entre paréntesis. ⁷Alelos minoritarios unidos en el grupo R. ⁸Valor p de la prueba χ^2 en la comparación global de frecuencias genotípicas en tablas 3x2.

Con la serie ampliada de sanas canarias DSC12, los datos de la comparación de frecuencias alélicas y genotípicas en AR resultan muy parecidos a los previos. Por alelos, el 24CA sigue apareciendo con menor frecuencia en AR que en sanas. Por genotipos, observamos también una frecuencia menor de mujeres homocigotas 2424 y mayor de mujeres homocigotas XX en AR, es decir, se mantienen las diferencias previas, ahora con mayor poder estadístico (0.92 frente a 0.80 para la diferencia en la frecuencia del alelo 24CA 0.84 frente a 0.52 para la diferencia en la frecuencia del genotipo XX). Las penetrancias de los genotipos 2424, 24X y XX siguen encajando en un modelo multiplicativo, con un mejor ajuste que se refleja en estimaciones muy parecidas del aumento de RR con cada alelo X (γ) calculadas a partir de las razones entre los RR de los tres genotipos (1,82; 1,80 y 1,80 en la Tabla 10 frente a 2,36; 1,88 y 1,44 en la Tabla 6). La frecuencia del alelo 26CA, al igual que en nuestro estudio previo ⁹⁸, es significativamente mayor en AR ($p=0.046$), aunque el poder estadístico de esta comparación es bajo (Tabla 9).

Tabla 10. Penetrancias y riesgos relativos con las series ampliadas DSC12 y LESC12

Series	Modelo	Penetrancia ¹			Riesgo Relativo ²			$\bar{\gamma}$ ³
		2424 (aa)	24X (Aa)	XX (AA)	24X → 2424	XX → 2424	XX → 24X	
LESC12	Recesivo	$4,8 \cdot 10^{-4}$	$4,7 \cdot 10^{-4}$	$8,9 \cdot 10^{-4}$	0,97	1,84	1,89	1,87
ARC1	Multiplicativo	$3,8 \cdot 10^{-3}$	$6,9 \cdot 10^{-3}$	$1,2 \cdot 10^{-2}$	1,82	3,26	1,80	1,81

¹ Frecuencia de enfermedad en el genotipo indicado (cálculo en Materiales y Métodos). ² Riesgos relativos entre los genotipos indicados en las series LESC12 y ARC1 frente a DSC12. ³ $\bar{\gamma}$ medio, promedio de los valores γ calculados a partir de las razones entre los RR de los tres genotipos en AR: $\gamma_1 = RR_{24X \rightarrow 2424}$; $\gamma_2 = \sqrt{RR_{XX \rightarrow 2424}}$; $\gamma_3 = RR_{XX \rightarrow 24X}$ y en LES: $\gamma_2 = RR_{XX \rightarrow 2424}$; $\gamma_3 = RR_{XX \rightarrow 24X}$

En el caso de la asociación a LES, los resultados con las series ampliadas presentan algunas diferencias con los anteriores. En cuanto a frecuencias alélicas, el alelo 24CA sigue teniendo menor frecuencia en LES, pero la diferencia con controles es menor que en la serie inicial (Tabla 9). Por genotipos, la diferencia en las frecuencias del genotipo 2424 entre enfermas y controles deja de ser significativa, y sólo el genotipo XX muestra una frecuencia significativamente mayor en LES que en controles (Figura 11). Esto supone que en las series ampliadas DSC12 y LESC12 se pierde la gran similitud que había en la asociación del MSCD40LG a LES y AR. En LES, la relación entre los RR para los tres genotipos se ajusta mejor ahora a un modelo dominante

de penetrancia (Tabla 10) con el alelo 24CA como alelo dominante de carácter protector, reflejado en un RR similar para los genotipos 2424 y 24X y un mayor RR sólo para el genotipo XX.

Por último, quisimos analizar la contribución de los alelos minoritarios que habíamos agrupado en la categoría “R” (todos aquellos menores de 21CA y mayores de 27CA), y comparamos sus frecuencias en las series ampliadas (Figura 11, Tabla 9). En total, entre las series de sanos, LES y AR aparecen ocho alelos minoritarios diferentes, cuatro menores de 21CA y cuatro mayores de 27CA. El conjunto de estos alelos supone un 2% del total de alelos en la población sana, pero en pacientes tanto de LES como de AR aparecen con una frecuencia significativamente mayor ($\approx 6\%$), similar en las dos enfermedades, con un RR de 3,22 en LES y 3,18 en AR (Tabla 9). Teniendo en cuenta sólo los alelos minoritarios largos ($>27CA$), aparecen 5 alelos en ARC1 y 11 en LESC12, pero ninguno en DSC12. Esta diferencia resulta significativa en LES aunque debido a los bajos números el intervalo de confianza para la OR es muy amplio ($p = 0,001$; OR (IC95%) = 21,3 (1,25 - 363,8)). Curiosamente esta mayor frecuencia de alelos $>27CA$ también se observa en la serie de LES peninsulares comprada con la de controles (5/52 alelos en LESP1 *vs* 4/482 alelos en DSP12; $p=0,001$; OR (IC95%) = 5,70 (3,01-10,80)). En conjunto, encontramos que los alelos minoritarios del MSCD40LG se asocian tanto a AR como a LES, y que los alelos largos $>27CA$ se asocian claramente a LES en dos series independientes.

· MSCD40LG y estabilidad del mRNA

A continuación quisimos determinar si el número de repeticiones CA en el MSCD40LG era el responsable de las diferencias en la estabilidad del mRNA y en la expresión en superficie de CD40LG que habíamos observado previamente en linfocitos T CD4⁺ activados *in vitro* y procedentes de individuos portadores de distintos alelos del MSCD40LG^{98,132}. Para ello, expresamos en una línea celular mRNAs quiméricos con una parte de la región 3'UTR del mRNA de CD40LG conteniendo diferentes alelos del MSCD40LG para medir la estabilidad del mRNA en función del alelo presente.

· Clonaje de variantes alélicas comunes del MSCD40LG en vectores de expresión

Para ello, amplificamos mediante PCR (PCR 3'UTRCD40LG, Tabla 3) parte de la región 3'UTR del gen (Figura 12) a partir de DNA genómico de individuos homocigotos para el alelo 24CA y otros tres alelos comunes (23CA, 26CA y 27CA). Los productos de PCR se clonaron en el vector de expresión en células de mamífero phMGFP a continuación de la secuencia codificante de la proteína verde fluorescente (GFP), generando las construcciones GFP-3UTR.CAn (Figura 12). La región 3'UTR de CD40LG clonada contiene, además del microsatélite, los elementos que modulan la estabilidad del mRNA de CD40LG, situados entre las dianas XbaI y HaeIII^{94,104,105}. Contiene además el final de la región 3'UTR hasta la diana ClaI, que se encuentra en el DNA genómico de CD40LG pasado el sitio de poliadenilación, a 186 pb. De esta forma en las construcciones generadas el mRNA que se transcribe utiliza la señal de poliadenilación propia del gen CD40LG, que sustituye en el clonaje a la señal *SV40 Late* que lleva el vector. Además de los plásmidos GFP-UTR.CAn, generamos una construcción adicional con una secuencia de 49 pares de bases insertada en la diana XbaI del vector phMGFP. Esta construcción (GFP-Δ3UTR) serviría de control para comparar la caída de los mRNA GFP-3UTR.CAn con un mRNA GFP sin la región 3'UTR de CD40LG insertada.

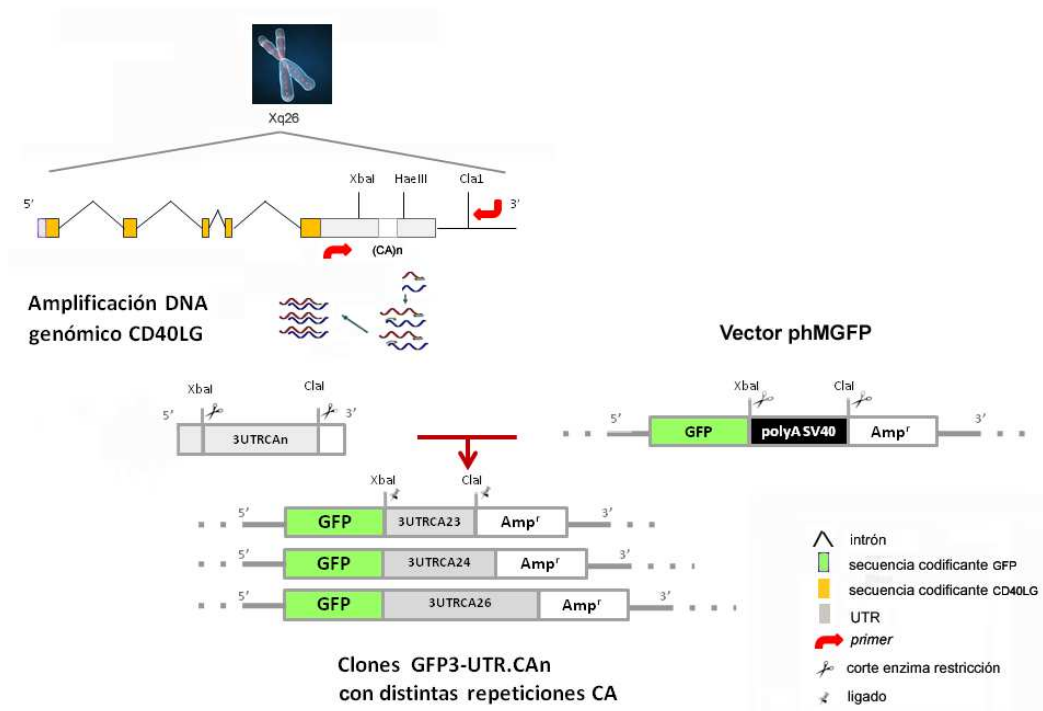


Figura 12. Clonaje de la región 3'UTR del gen CD40LG. Las secuencias clonadas son las situadas entre las dianas de restricción XbaI y ClaI, en el extremo 3'UTR del gen de CD40LG obtenidas a partir de DNA genómico de diferentes individuos, generando así las construcciones GFP-3UTR.CAn con los alelos 23CA, 24CA, 26CA y 27CA del microsatélite. La secuencia 3'UTR de CD40LG queda situada justo a continuación de la secuencia codificante para la proteína verde fluorescente (GFP) y la fusión GFP-3UTR se expresa a través del promotor de citomegalovirus del vector phMGFP (CMV I.E. *Enhancer/promoter*).

Para conseguir células con expresión estable del DNA quimérico, insertamos la secuencia que codifica para la enzima neomicina fosfotransferasa a continuación de un promotor del virus SV40 (SV40 Enhancer/Early Promoter) (Figura 13). Los vectores resultantes, PCR GFP3UTR.CAn-neo (en adelante GFP-3UTR.CAn), expresan de forma constitutiva la fusión de GFP con la región 3'UTR de CD40LG y la resistencia al antibiótico G418, que permite el aislamiento de transfectantes estables.

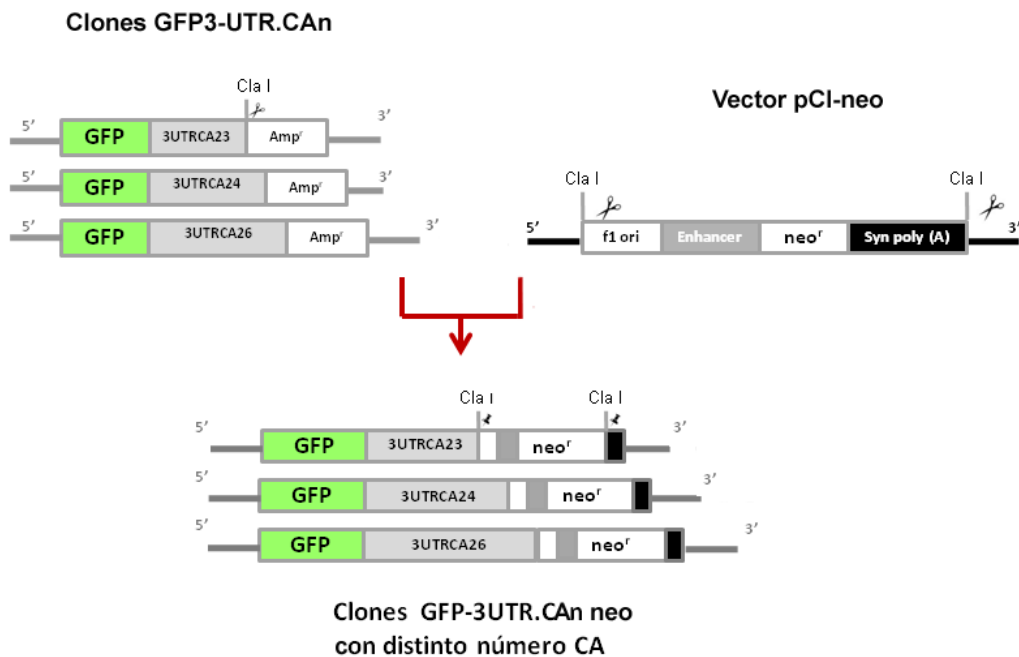


Figura 13. Clonaje de la resistencia a neomicina en los vectores GFP-3UTR.CAn. En cada construcción GFP-3UTR.CAn con los alelos del MSCD40LG se insertó la secuencia de 1969 pb del vector pCI-neo que contienen la región codificante para la enzima neomicina fosfotransferasa (neo^r) que confiere resistencia al antibiótico G418. La resistencia queda insertada en la diana ClaI donde finaliza la secuencia 3'UTR de CD40LG clonada, justo a continuación de la fusión GFP-3'UTR.CAn.

Se comprobó que las construcciones generadas contenían las secuencias insertadas mediante cortes con enzimas de restricción y análisis del tamaño de los fragmentos por electroforesis en geles de agarosa y también mediante PCR (PCRMSCD40LG, Tabla 3). Éstas PCR se hicieron de dos maneras, la primera igualando el número de copias iniciales del vector en la reacción de PCR al número de copias del gen presentes en 10 ng de DNA genómico (unas 1500 copias de cada juego cromosómico por ng de DNA genómico), equivalente a la cantidad que usábamos para los tipajes en los DNAs genómicos de controles y pacientes. De esta manera, el perfil de picos en la electroforesis de los productos

de la PCR era comparable al de los tipajes que habíamos hecho a partir de muestras de DNA genómico de pacientes y controles (Figura 14). Cada PCR produce la banda del tamaño correspondiente al fragmento amplificado y bandas extra producidas por errores de la polimerasa al replicar las repeticiones del dinucleótido CA del microsatélite, que generalmente tienen menor tamaño debido a la pérdida de repeticiones del dinucleótido, un fenómeno conocido y bien estudiado ^{149,150}. La segunda forma consistió en utilizar las mismas condiciones de la de PCR MSCD40LG, pero partiendo de una gran cantidad de copias ($\approx 0,01$ ng/ μ l, unas 10^5 - 10^6 moléculas iniciales por microlitro en la reacción) y amplificando en pocos ciclos (10-12), con lo que se consigue una banda mayoritaria que se distingue claramente de las bandas extra producidas por error de la polimerasa (Figura 14).

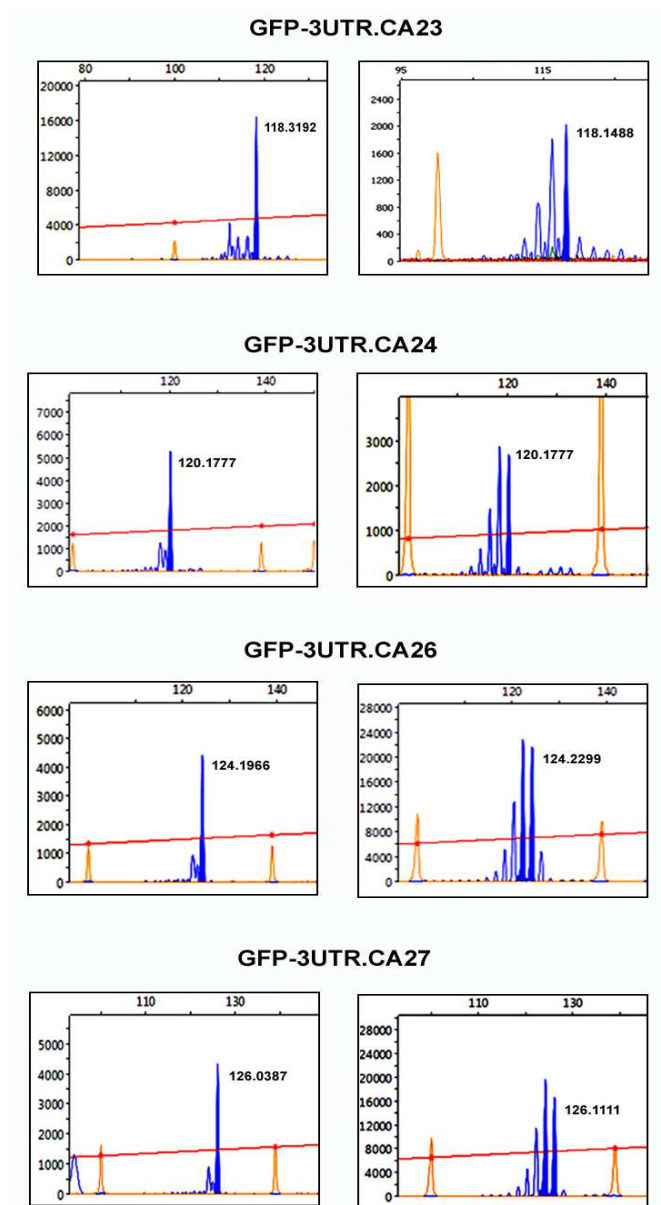


Figura 14. Tipajes del MSCD40LG en los vectores de expresión. En las distintas construcciones generadas se comprobó el alelo del microsatélite presente por PCR (PCR MSCD40LG) y electroforesis capilar. Se muestran los perfiles de electroforesis de los productos amplificados (picos azules) cuyo tamaño se calculó por comparación con un patrón de longitud conocida (picos naranjas). En los gráficos se muestra la longitud en pares de bases (eje X) frente a la intensidad de fluorescencia (eje Y). A la izquierda, imágenes de las PCR realizadas en pocos ciclos partiendo de gran número de copias iniciales, que dan un pico mayoritario claro. A la derecha, PCR con los DNA plasmídicos diluidos para tener un número de copias iniciales semejante al de las PCR con las que se tipan los DNA genómicos. Los tamaños de las bandas en los tipajes realizados de las dos maneras coincidían en todos los casos.

Para comprobar que la diferencia entre las distintas construcciones residía únicamente en el número de repeticiones CA del microsatélite, secuenciamos la región 3'UTR clonada en los plásmidos con las construcciones GFP-3UTRCA_n usando el *primer* izquierdo de la PCR GFP-MSCD40LG (Tabla 3). Las secuencias obtenidas (Figura 15A) se compararon con la secuencia de referencia del DNA genómico de CD40LG (NG_007280.1) comprendida entre la secuencia poli-C y la diana Cla I. Aparte de las pocas indeterminaciones en las lecturas (“n” en las secuencias en Figura 15), todas las secuencias, hasta donde alcanzaban con buena resolución (en dos de los vectores la secuencia no alcanzaba hasta la diana Cla I, pero sí llegaba hasta pasado el microsatélite), eran idénticas a la de referencia (Figura 15B y C), y se diferenciaban sólo en la longitud del microsatélite, con la excepción de GFP-3UTRCA₂₇, que tenía una citosina (C) menos que el resto en la secuencia poli-C, con (C)₁₃ frente a (C)₁₄ en el resto.

C

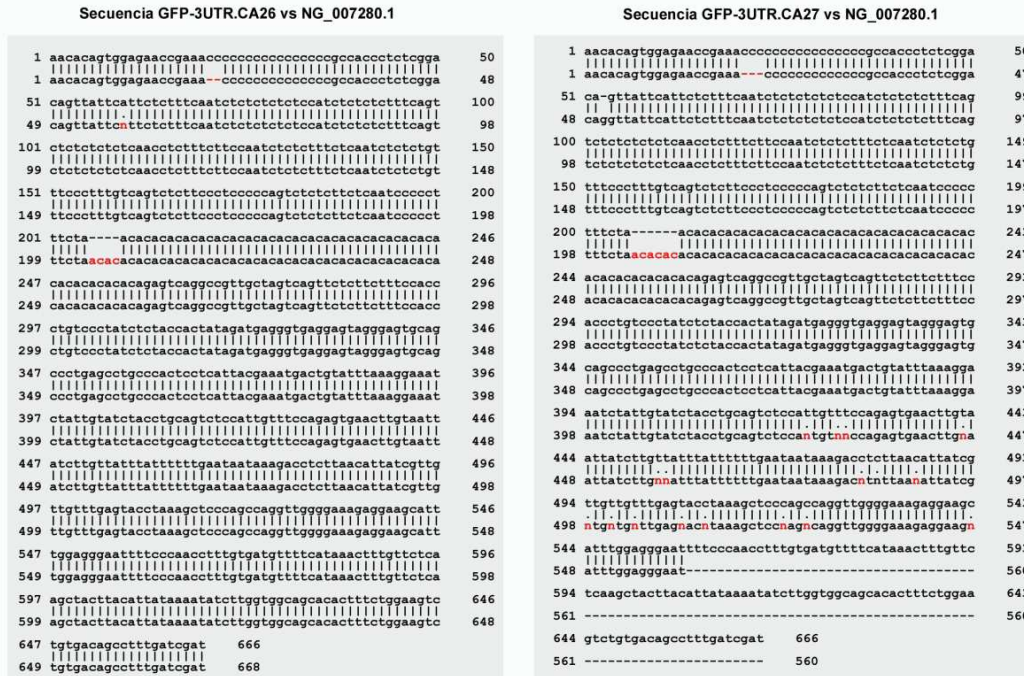


Figura 15. Secuenciación de los fragmentos de la región 3'UTR de CD40LG clonados en los vectores GFP-3UTR.CAn. (A) Se secuenciaron 100 ng / μ l de cada vector a partir del *primer* izquierdo de la PCR GFP-MSCD40LG. En la parte superior se muestra la secuencia de referencia del DNA genómico de CD40LG (NG_00728.1). Debajo, las secuencias obtenidas en los vectores GFP-3UTR.CAn con los alelos 23CA, 24CA, 26CA y 27CA. Están resaltadas las secuencias poli-C (en verde), CU-rich (en azul) y el microsatélite (en rojo). (B) y (C) Comparación de la secuencia de referencia con las secuencias para los vectores, con el programa EMBOSS merger. Las diferencias aparecen señaladas en rojo.

Para verificar que la menor longitud de la poli-C en el vector GFP-3UTRCA27 no era debida a un error en la secuenciación, amplificamos mediante PCR la secuencia poli-C (PCR poli-C, Tabla 3) en los cuatro vectores, partiendo de $\approx 0,01$ ng/ μ l de vector al igual que habíamos hecho en las PCR para el tipaje del microsatélite. Estas PCR confirmaron la longitud (C)₁₃ en la secuencia poli-C del vector GFP-3UTR.CA27 (Figura 16).

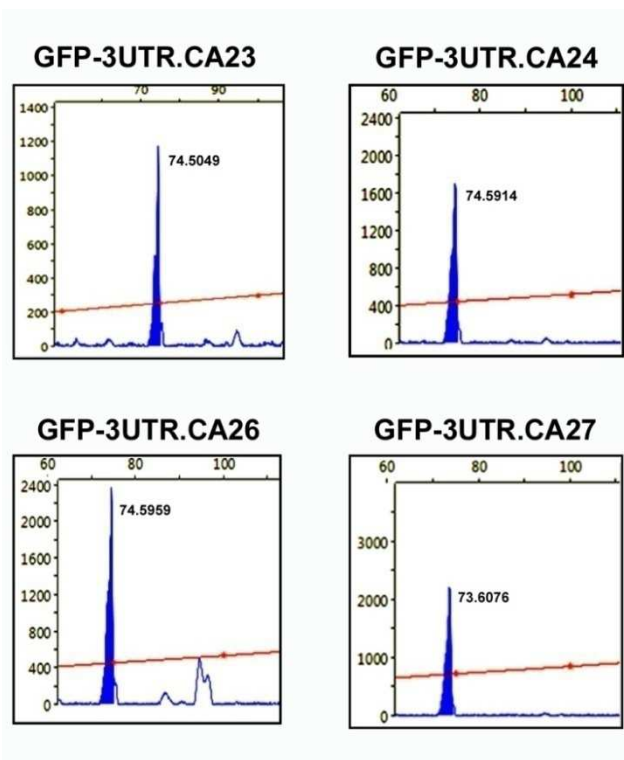


Figura 16. Amplificación de la secuencia poli-C. En las distintas construcciones generadas se amplificó la región conteniendo la secuencia poli-C (PCR poli-C, Tabla 3), en pocos ciclos y partiendo de un gran número de copias del vector. Los productos de PCR se resolvieron por electroforesis capilar, y a partir del tamaño estimado se calculó el número de repeticiones (C)_n por comparación con el de la secuencia de referencia de CD40LG en Genbank. La construcción GFP-3UTR.CA27 tenía una C menos que el resto. Se muestran los perfiles de electroforesis del fragmento amplificado (picos azules) señalando el tamaño en pares de bases (eje X).

El número de repeticiones del dinucleótido (CA)_n del microsatélite y el de citosinas (C)_n de la poli-C obtenidos por secuenciación no coincidente con el calculado a partir del tamaño del fragmento amplificado por PCR, pero siempre se mantiene la diferencia de repeticiones (CA)_n y (C)_n entre los vectores (Tabla 11). Las secuencias y los tipajes en los vectores GFP-3UTR.CAn indican que las distintas construcciones GFP-3UTR.CAn tienen idéntica secuencia salvo por el número de repeticiones (CA)_n, con la excepción de una C menos en la secuencia poli-C del vector GFP-3UTR.CA27 respecto al resto.

Tabla 11. Número de repeticiones (CA)_n en el MSCD40LG y (C)_n en la secuencia poli-C en las construcciones.

Construcción	Nº repeticiones CA		Nº repeticiones C	
	PCR ¹	Secuenciación ²	PCR ³	Secuenciación
GFP-3UTR.CA23	23	25	9	14
GFP-3UTR.CA24	24	26	9	14
GFP-3UTR.CA26	26	28	9	14
GFP-3UTR.CA27	27	29	8	13

¹ Repeticiones calculadas a partir del tamaño de la banda amplificada con la PCR MS40LG. ² Número de repeticiones (CA)_n o (C)_n en las secuencias. ³ Repeticiones calculadas a partir del tamaño de la banda amplificada con la PCR poli-C.

· Obtención de transfectantes estables de las construcciones GFP-3UTR.CAn

Los vectores GFP-3UTR.CAn se transfectaron en la línea celular humana HEK293 y se cultivaron en presencia de 400 µg/ml del antibiótico G418. Tras aproximadamente una semana en presencia del antibiótico la mayoría de las células morían y aparecían algunas colonias de células resistentes, que habían integrado el vector en el genoma y expresaban la resistencia a G418. La expresión de la fusión GFP-3UTR.CAn se detectaba por microscopía de fluorescencia (Figura 17). En las placas de transfecciones se incluyeron pocillos con células no transfectadas como control de G418. En estos pocillos todas las células morían antes de dos semanas de cultivo con G418 (Figura 17A). Se aislaron varias colonias resistentes de cada vector transfectado, y se crecieron de forma independiente para establecer clones de células con las construcciones GFP-3UTR.CAn insertadas de forma estable (Figura 17B-).

Los alelos del MSCd4LG se volvieron a comprobar en los clones HEK293 con las construcciones GFP-3UTR.CAn insertadas. Para ello, se aisló RNA total, y se obtuvo cDNA mediante retrotranscripción. Se realizaron dos amplificaciones en estos cDNAs, la primera con la PCR GFP-MSCD40LG (tabla 3), que tiene un *primer* en el final de la secuencia codificante de la GFP y el otro pasado el MS (Figura 18). Con esta PCR comprobamos que el mRNA que expresaban las construcciones insertadas era el esperado. Alícuotas de los productos de PCR GFP-MSCD40LG purificados se amplificaron 10-12 ciclos adicionales con los *primers* para el tipaje del MS (PCR MSCD40LG), y se asignaron los alelos calculando el tamaño del producto de PCR por electroforesis capilar. En la Figura 18 se muestra un esquema de las PCR realizadas en las construcciones insertadas en los clones HEK293. En todos los clones, el alelo del MSCD40LG coincidía con el del vector introducido (Figura 19). Estos clones, una vez tipados, se utilizaron para medir la estabilidad de los mRNA GFP-3UTR.CAn.

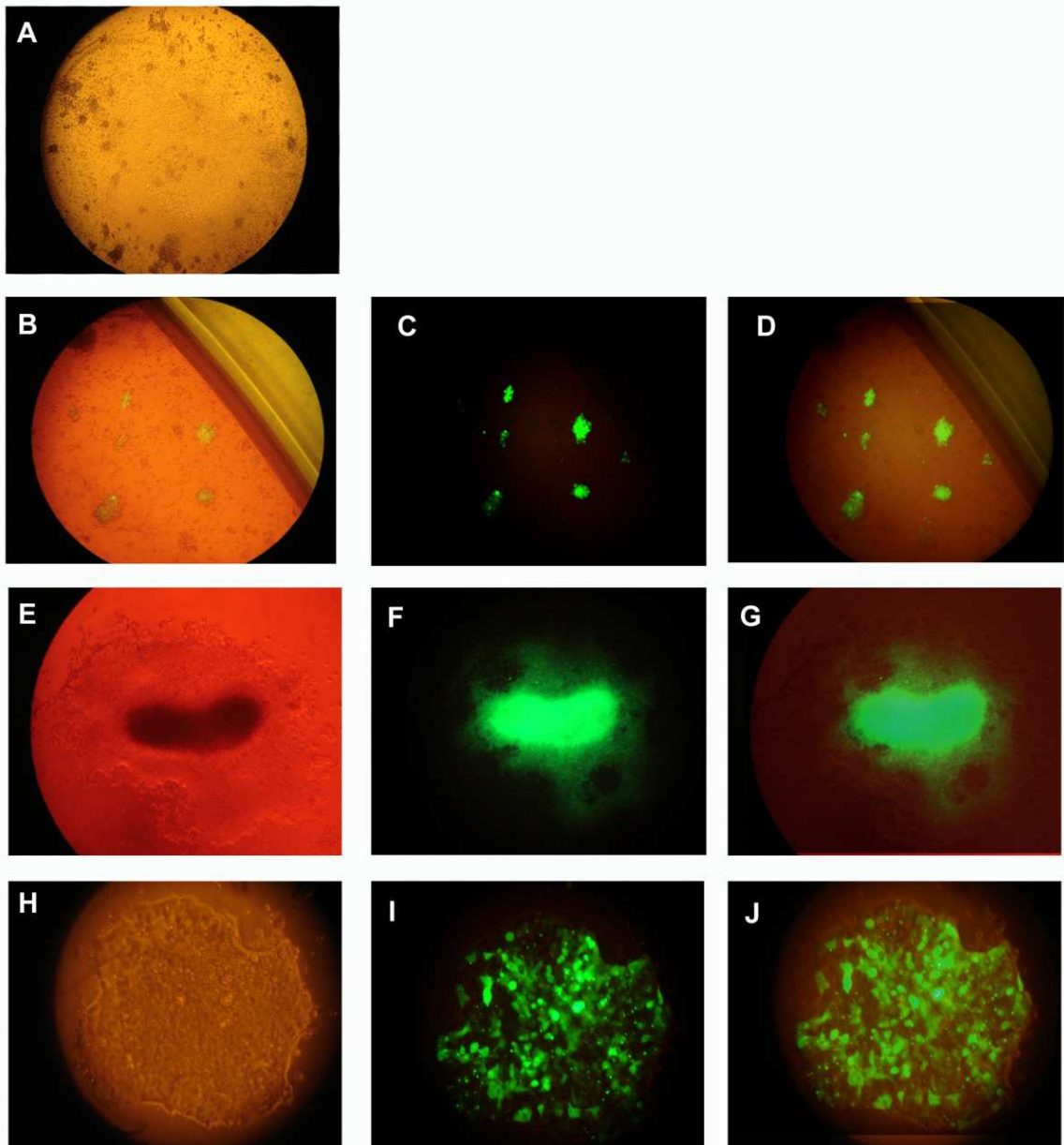


Figura 17. Imágenes de los cultivos de la línea HEK293 transfectada con los vectores GFP-3UTR.CAn. (A) Células no transfectadas usadas como control del antibiótico G418, donde tras 8-10 días morían todas las células. (B-D) Aparición de colonias resistentes al tratamiento con 400 $\mu\text{g}/\text{ml}$ del antibiótico G418 en los pocillos con células transfectadas. (E-J) Imágenes del crecimiento de colonias resistentes aisladas para generar los clones GFP-3UTR.CAn.

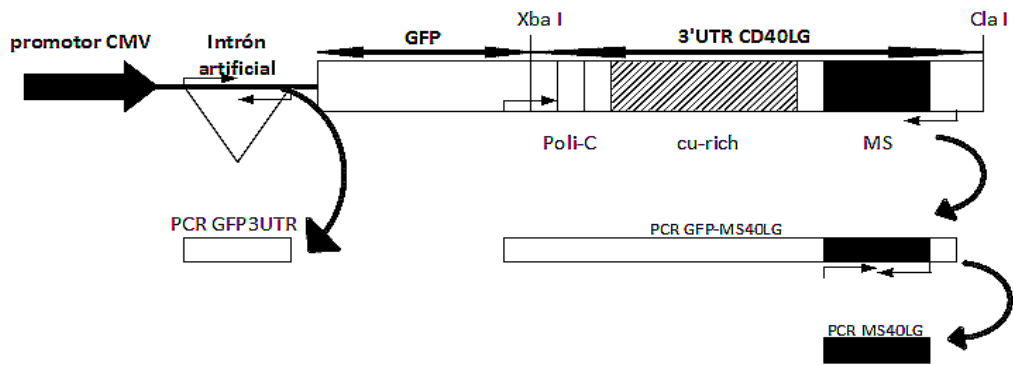


Figura 18. Esquema de las PCR realizadas en las construcciones GFP-3UTR.CAN transfectadas en la línea HEK293. Está representada la porción de los vectores que contiene las construcciones GFP-3UTR.CAN con el promotor de citomegalovirus (CMV), el intrón artificial al comienzo de la secuencia transcrita a mRNA, la secuencia codificante de la GFP y el fragmento clonado de la región 3'UTR de CD40LG (entre Xba I y Cla I), donde se indica la situación del microsatélite. Superpuestas en el esquema están las parejas de *primers* de las PCR realizadas en estas construcciones. Para comprobar que los mRNA expresaban las construcciones completas y que los alelos del MSCD40LG se mantenían en los clones HEK293 se realizaron dos PCR seriadas a partir de cDNA (PCR GFP-MSCD40LG y PCR MSCD40LG, derecha). La PCR GFP3UTR se utilizó para medir la caída del mRNA GFP-3UTR.CAN en las cinéticas.

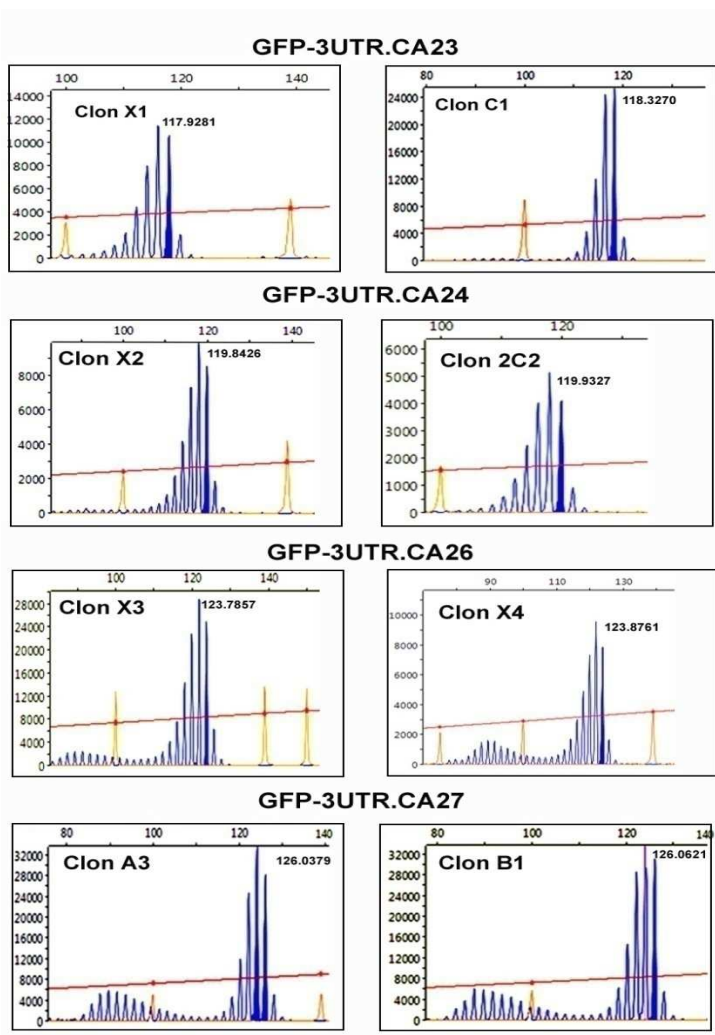


Figura 19. Tipajes en los clones HEK293 GFP-3UTR.CAN. A partir de células HEK293 transfectadas con las distintas construcciones GFP-3UTR.CAN se seleccionaron dos clones portadores de cada uno de los alelos: 23CA (118 pb), 24CA (120 pb), 26CA (124 pb) y 27 CA (126 pb). En ellos se volvió a tipar el MSCD40LG para comprobar que se mantenía el alelo introducido. Se muestran los perfiles de electroforesis de los fragmentos amplificados en cada clon (picos azules) y el patrón de tamaño (picos naranjas) añadido a cada muestra. En el eje X se señala la longitud en pares de bases. En todos los casos el alelo obtenido en ambos clones se correspondía con el esperado.

· Cinéticas de caída de los mRNA GFP-3UTR.CAn

Para medir la estabilidad de los mRNA GFP-3UTR.CAn en función del alelo del MSCD40LG presente, se utilizaron dos clones de HEK293 transfectados con cada uno de los alelos. Las células se cultivaron en placas de 24 pocillos en medio completo y se crecieron hasta llegar a una confluencia aproximada del 90%. Para parar la transcripción se añadió actinomicina D (Act-D) a 10 ug/ml, y se recogieron muestras a los 0, 30 60 y 90 min de tratamiento. A partir de estas muestras se aisló el RNA total. Una alícuota de 600 ng de cada muestra se retrotranscribieron para obtener cDNA. Todas las reacciones incluían como controles negativos muestras sin retrotranscriptasa en la mezcla, para descartar amplificación a partir de DNA genómico. Las muestras de cDNA se amplificaron mediante PCR a tiempo real en placas de 384 pocillos. Además de amplificar el cDNA GFP-3UTR.CAn de las construcciones insertadas, se amplificaron los cDNAs de otros dos genes expresados por HEK293, GADPH, como gen de referencia de expresión estable para controlar las variaciones en las distintas etapas del ensayo, y de c-fos, que tiene un mRNA de vida media corta, como control del ensayo de caída del mRNA.

Las PCR GFP3UTR, GADPH y c-fos se realizaron en pocillos independientes. Se hicieron dos cinéticas de Act-D por cada clon, al menos dos RT por muestra de cada cinética, y al menos dos PCR con duplicados por cada muestra de RT. En repetidas ocasiones no se obtuvo ningún producto de PCR a partir del clon C1 (GFP-3UTR.CA23), y se descartó para el análisis posterior de las cinéticas por lo que los resultados para este alelo provienen de un solo clon. Para construir rectas patrón con las que transformar los valores de umbral de ciclo (Ct) obtenidos en las PCR cuantitativas (qPCR) en valores de logaritmo del número de moléculas se hicieron diluciones seriadas ($1/2$) a partir de muestras de cDNA obtenidas en el tiempo inicial (t_0) de las cinéticas con Act-D. Como las cantidades de moléculas al inicio de las cinéticas (t_0) son desconocidas, se eligió un valor arbitrario de logaritmo de moléculas de partida y se asignó a cada punto de la recta un valor en función de la dilución con respecto a ese valor de partida. Las pendientes de todas las rectas patrón obtenidas para cada PCR (GFP-3UTR.CAn, GADPH y c-fos) se compararon en función del alelo, clon, RT, cinética o placa de la que procedían. No había correlación con ninguno de estos parámetros, y tanto las pendientes como los valores de corte con el eje de ordenadas seguían una distribución

normal, por lo que construimos rectas patrón con los valores promedio de pendiente y corte con el eje de ordenadas de todas las rectas de cada PCR (Figura 20). Con estas tres rectas patrón “medias” (GFP-3UTR.CAn, GADPH y c-fos) transformamos todos los valores de Ct obtenidos en las PCR en valores de logaritmo del número de moléculas de mRNA (log molec).

nº CLON	CLON	Nº Cinética	nº RT	nº placa	PCR	Pendiente	Corte con eje Y
4	2C2	1	1	4	GFP3UTR	-0,287	9,948
2	A3	1	2	3	GFP3UTR	-0,305	8,971
3	B1	1	2	3	GFP3UTR	-0,298	8,995
1	D1	1	2	3	GFP3UTR	-0,297	9,106
5	X0	2	3	1	GFP3UTR	-0,314	10,149
5	X0	1	4	2	GFP3UTR	-0,322	10,525
5	X0	1	6	7	GFP3UTR	-0,314	9,996
6	X1	2	3	1	GFP3UTR	-0,332	9,874
6	X1	1	4	2	GFP3UTR	-0,318	9,652
6	X1	2	6	7	GFP3UTR	-0,331	9,591
7	X2	1	4	2	GFP3UTR	-0,310	9,495
7	X2	2	3	4	GFP3UTR	-0,328	10,193
7	X2	2	5	5	GFP3UTR	-0,343	10,095
7	X2	1	5	6	GFP3UTR	-0,308	9,089
8	X3	1	4	2	GFP3UTR	-0,355	10,541
8	X3	2	3	4	GFP3UTR	-0,314	9,636
8	X3	2	5	5	GFP3UTR	-0,346	9,845
8	X3	1	5	6	GFP3UTR	-0,314	9,320
9	X4	2	3	1	GFP3UTR	-0,332	9,874
9	X4	1	2	3	GFP3UTR	-0,348	10,172
media						-0,323	9,743

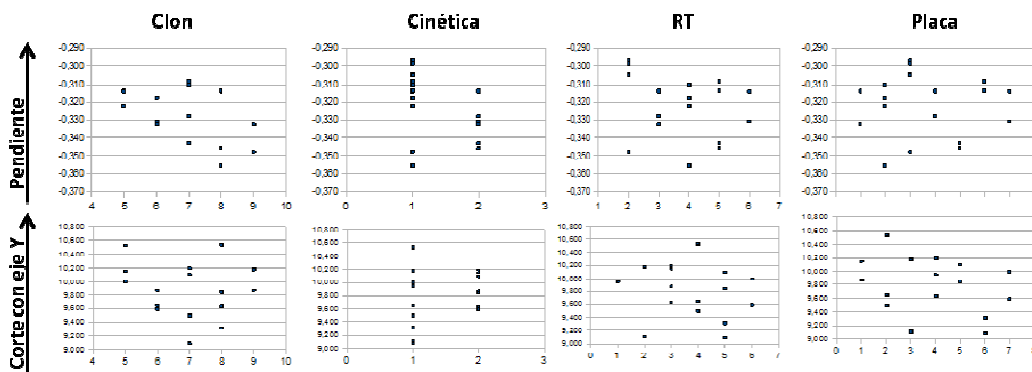


Figura 20. Rectas patrón de la PCR GFP3UTR para la transformación de valores de Ct en logaritmo de moléculas. (A) Rectas patrón de las placas de PCR para la PCR GFP3UTR, con los valores de la recta media (en verde) empleada para transformar todos los valores de Ct. (B) Ni las pendientes ni los valores de corte con el eje Y de las rectas correlacionaban con las variables clon, cinética, RT o placa. Lo mismo ocurría en las PCR GADPH y c-fos, en las que se calcularon rectas patrón medias de igual manera.

Los valores de log molec se clasificaron en función del alelo del microsatélite y del tiempo de tratamiento con Act-D para estimar la cantidad de mRNA a cada tiempo y calcular su vida media. Para corregir las variaciones en el ensayo, los valores individuales de log molec PCR GFP3UTR.CAn y c-fos en cada muestra se corrigieron con el valor de log molec GADPH correspondiente de la siguiente manera:

$$\frac{\log \text{ molec GFP-3UTR.CAn}}{1000 \text{ molec GADPH}} = \left(\log \text{ molec GFP-3UTR.CAn} - \log \text{ molec GADPH} \right) + 3$$

El cálculo es equivalente a expresar el número de moléculas de mRNA GFP-3UTR.CAn (o c-fos) por cada 1000 moléculas de mRNA GADPH, pero operando con los valores en logaritmos. Los valores corregidos se agrupaban en función del clon y/o la cinética de donde procedían (Figura 21). Además, en el caso de GFP-3UTR.CAn estas diferencias parecían reflejar diferencias reales, ya que en los cultivos de los clones HEK293 GFP-3UTR.CAn también se apreciaban diferencias de brillo de la proteína GFP entre clones con la misma construcción insertada, e incluso en un mismo clon dependiendo del estado de crecimiento en el pocillo, lo que podría explicar las diferencias entre las dos cinéticas en un mismo clon. Para poder comparar las pendientes de caída de los distintos alelos entre sí se trasladaron todos los valores de cada serie (t_0 a t_{90} para cada clon en una cinética concreta), de forma que en el tiempo inicial (t_0) la cantidad media de moléculas fuera log molec = 3, tanto para GFP-3UTR.CAn como para c-fos:

$$\frac{\text{Norm log molec GFP-3UTR.CAn}}{1000 \text{ GADPH}} = \frac{\log \text{ molec GFP-3UTR.CAn}}{1000 \text{ GADPH}} - \hat{X}_{\text{clon, cinética}} \left(\frac{\log \text{ molec GFP-3UTR.CAn}}{1000 \text{ GADPH}} \right) + 3$$

Esta normalización agrupa de forma homogénea los valores a cada tiempo de la cinética, sin afectar al valor de la pendiente de la recta calculada con la media de los valores a cada tiempo. En la Figura 21 se muestran los valores de las caídas de los mRNA PCR GFP3UTR.CAn y c-fos para los clones con el alelo 24 alelo del MSCD40LG antes y después de su normalización.

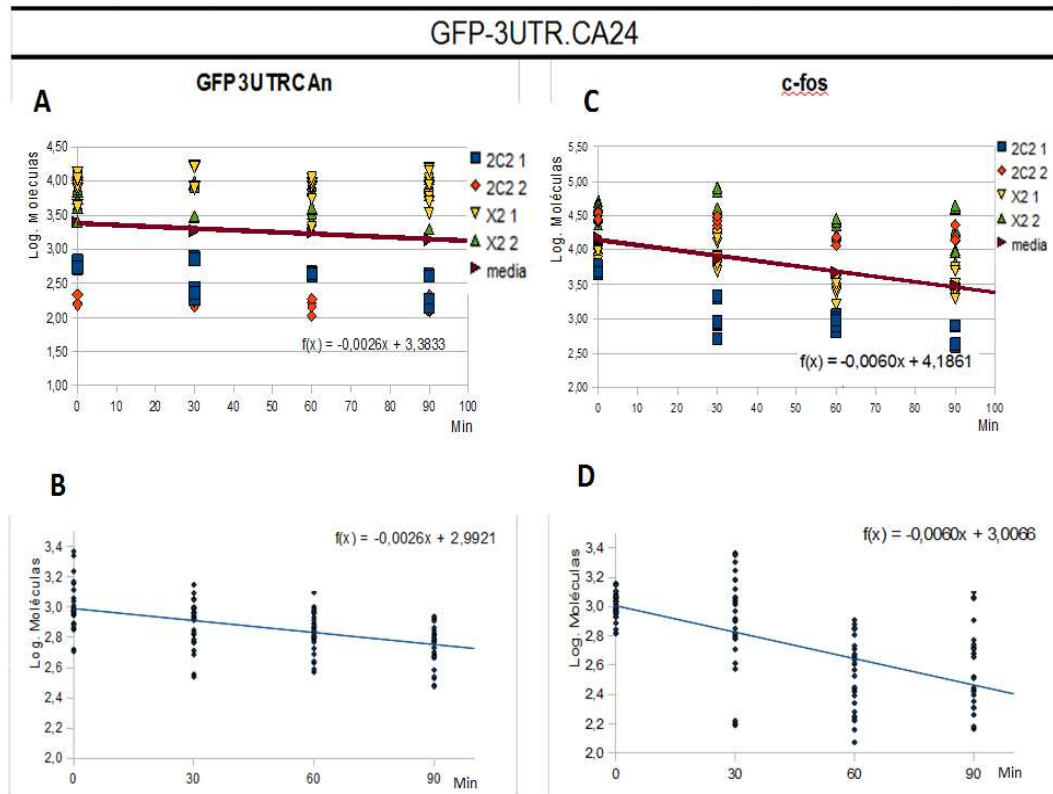


Figura 21. Normalización de los valores log molec GFP-3UTR.CAn y c-fos. Los valores de log molec GFP-3UTR.CAn y c-fos (corregidos con el mRNA control GADPH) se agrupaban en función del clon y/o la cinética de procedencia. Para igualar el origen de todas las series se normalizaron los valores de cada una de forma que a tiempo t_0 todas las series tuvieran la misma media (log molec = 3). Las gráficas muestran los valores previos a la normalización para los mRNA GFP-3UTR.CAn (A) y c-fos (C) del alelo 24CA y después de su normalización (B, D). En los mRNA GFP-3UTR.CAn se aprecia claramente la diferencia entre los dos clones con el alelo 24CA (X2 y 2C2). Sse apreciaban también diferencias entre las dos cinéticas realizadas en cada clon (X2,1 y X2,2). En las gráficas aparecen las ecuaciones de las rectas GFP-3UTR.CAn y c-fos, que se mantienen tras la normalización.

• Comparación de la estabilidad del mRNA en función del alelo del MSCD4oLG

Los valores normalizados de log molec GFP-3UTR y c-fos se utilizaron para calcular las rectas de caída para cada alelo mediante regresión lineal. Como era de esperar, las pendientes de caída de las rectas c-fos son parecidas en todos los alelos (Figura 22). En comparación, las caídas de los mRNA GFP-3UTR.CAn tienen mayores diferencias entre sí. En general, los tiempos de vida media ($t_{1/2}$) de los mRNA GFP-3UTR.CAn oscilan entre 114 y 392

min, menores que el mRNA control GFP- Δ 3UTR, que no tiene una caída apreciable en los 90 min de las cinéticas, lo que indica que la región 3'UTR de CD40LG clonada en las construcciones actúa como un elemento que desestabiliza los mRNA. Llama la atención la recta de GFP-3UTR.CA23, que tiene la menor caída con una pendiente incluso ligeramente positiva.

Las rectas GFP-3UTR.CAn y c-fos para los distintos alelos se compararon mediante el estadístico F (*Extra Sum of Squares F test*) que estima la probabilidad de los datos bajo la hipótesis nula de igualdad de pendientes entre los diferentes alelos frente a la hipótesis alternativa de pendientes distintas para cada alelo. En la Tabla 11 se recogen los parámetros de las rectas c-fos y GFP-3UTR.CAn, con el cálculo de la vida media ($t_{1/2}$), y las comparaciones entre los distintos alelos.

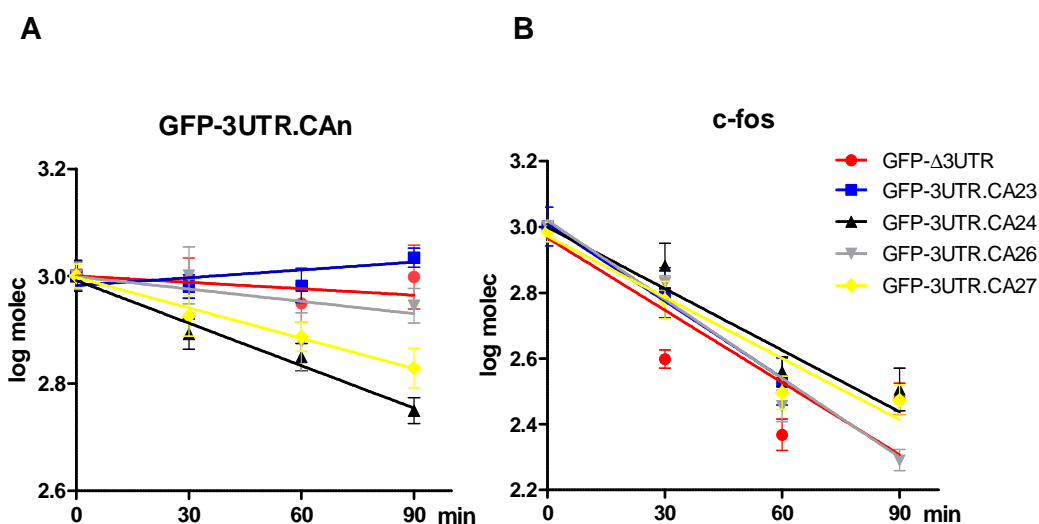


Figura 22. Rectas de caída de los mRNA GFP-3UTR.CAn y c-fos. Caída de los mRNA GFP-3UTR.CAn (A) y c-fos (B) en log molec frente al tiempo tras la parada de la transcripción con Act-D en función del alelo del MSCD40LG.

Las rectas de caída del mRNA de control c-fos son muy homogéneas, con $t_{1/2}$ entre 38 y 49 min, salvo para el alelo 23CA (83 min). Esta última se aparta del resto, aunque incluida en el conjunto total de rectas c-fos no produce diferencias estadísticamente significativas. En conjunto, las rectas c-fos indican una buena homogeneidad en el ensayo entre las distintas construcciones.

Tabla 11. Comparación de rectas c-fos y GFP-3UTR.CAn en función del alelo de MSCD40LG

	Rectas c-fos				
	GFP-Δ3UTR	GFP-3UTR.CA23	GFP-3UTR.CA24	GFP-3UTR.CA26	GFP-3UTR.CA27
Corte con eje de ordenadas	2,967	2,943	3	3,02	2,972
Pendiente	-0,00733	-0,003626	-0,006245	-0,008005	-0,006204
t ½ (min) ¹	41	83	48	38	49
p Δ3UTR,23CA,24CA,26CA,27CA	0,515				
p 23CA,24CA,26CA,27CA	0,161				
p 24CA,26CA,27CA	0,303				
p 24CA,26CA	0,239				
p 24CA,27CA	0,977				
p 26CA,27CA	0,243				
	Rectas GFP-3UTR.CAn				
	GFP-Δ3UTR	GFP-3UTR.CA23	GFP-3UTR.CA24	GFP-3UTR.CA26	GFP-3UTR.CA27
Corte con eje de ordenadas	3,001	2,983	2,992	2,999	2,998
Pendiente	-0,0003996	0,000485	-0,002645	-0,000767	-0,001887
t ½ (min) ¹	753	-	114	392	160
p 23CA,24CA,26CA,27CA	0,0004²				
p 24CA,26CA,27CA	0,0198				
p 24CA,26CA	0,0268				
p 24CA,27CA	0,0631				
p 26CA,27CA	0,0707				

¹ Vida media (- log₂ / Pendiente). ² En negrita valores p <0,05

En cuanto a las rectas de caída del mRNA PCR GFP3UTR.CAn, la comparación del conjunto de rectas (excluyendo el control GFP-Δ3UTR) indica que las caídas no son homogéneas. Nuevamente la recta del alelo 23CA se aparta claramente del resto, y no se aprecia caída a lo largo de los 90 min de la cinética (la recta tiene una pendiente incluso ligeramente positiva). Este resultado, unido una caída de c-fos que también se aparta del resto en este alelo, y al hecho de que solo pudieron analizar los datos de un clon con el alelo 23CA (por la eliminación del clon C1), lleva a poner en duda la fiabilidad de este resultado. Respecto al resto de alelos, la comparación de los alelos 24CA, 26CA y 27CA de nuevo indica que las caídas no son homogéneas (p = 0,02). Comparando estos tres alelos dos a dos, sólo la diferencia entre 24CA y 26CA resulta estadísticamente significativa, aunque las comparaciones 24CA, 27CA y 26CA, 27CA se acercan a la significación (p = 0,06 y p = 0,07 respectivamente). En todas estas comparaciones, las rectas c-fos resultan homogéneas

(Tabla 11), lo que indica que las diferencias entre las caídas del mRNA GFP-3UTR.CAn se deben a la diferencia en el alelo del MSCD40LG que llevan las construcciones.

En conjunto, estos resultados indican que el MSCD40LG actúa de forma autónoma como un elemento que afecta a la estabilidad del mRNA de forma dependiente del número de repeticiones (CA)_n presentes, aunque no permiten determinar con precisión la magnitud de las diferencias en la estabilidad en función del alelo.

MSCD40LG, estructura secundaria del mRNA y estabilidad del mRNA

Para tratar de explorar posibles mecanismos mediante los cuales el MSCD40LG pudiera ejercer un efecto en la estabilidad del mRNA generamos predicciones de las estructuras secundarias de la región 3'UTR de CD40LG clonada en las construcciones GFP-3UTR.CAn, y buscamos en la secuencia del mRNA potenciales dianas para las proteínas PTB y hnRNPL, que forman parte del complejo regulador que se une a la región donde se sitúa el MSCD40LG (*Complex II*, Figura 5). Las predicciones de las estructuras fueron hechas con el programa mFold ¹⁴⁰ que calcula una estructura óptima con el mayor descenso en energía interna (ΔG) y un conjunto de estructuras subóptimas con ΔG mayores que la estructura óptima hasta un límite definido. El conjunto de estructuras se representa en un gráfico llamado *Energy Plot* en el que ambos ejes X e Y representan la secuencia de bases del RNA. Cada punto del *Energy Plot* representa un par de bases (X_i,Y_j) con una gradación de colores en función de la energía libre del plegamiento en el que se encuentran representados, con los pares de bases de la estructura óptima representados en solitario en el triángulo de la parte inferior de la diagonal del gráfico. Relacionados con el *Energy Plot* el programa ofrece tres parámetros (p-num, h-num y s-num) que describen la propensión de cada base a aparearse con el resto de bases de la secuencia. El valor p-num indica el número de bases diferentes con que cada base concreta se encuentra apareada en los plegamientos representados, h-num es el promedio de valores p-num en las bases que forman una hélice determinada y s-num indica la frecuencia con que cada base se encuentra desapareada. Las predicciones de la estructura secundaria con una buena definición según estos parámetros tienen más probabilidad de coincidir con la estructura real cuando se comparan las predicciones con los de moléculas de RNA de estructura conocida ¹⁵¹. Por tanto, estos parámetros ayudan a valorar la probabilidad de que una estructura generada sea parecida a la real.

La predicción de la estructura de GFP-3UTR.CA24 desde XbaI hasta la cola de poli-A (con 50 adeninas añadidas al final de la secuencia) se muestra en la Figura 24 A como representativa de las estructuras GFP-3UTR.CAn, junto con el *Energy Plot* y dos gráficas que representan los valores de p-num y s-num (eje Y) para cada base en la secuencia de bases (eje X), en paralelo a la representación en el *Energy Plot* (Figura 24 B). Según el *Energy Plot* la región con mejor definición dentro de esta secuencia está entre las bases 200 y 380 (delimitada por el cuadrado rojo). De forma correspondiente la gráfica de p-num tiene valores bajos en esa zona (marcada con flecha roja de doble punta). Esta región abarca en la estructura óptima (Figura 24 A) desde 16 bases antes del comienzo del MSCD40LG hasta 131 bases pasado éste, e incluye el MSCD40LG y un *stem* largo situado justo a continuación (MSCD40LG y el *stem* señalados en azul y gris en las gráficas p-num y s-num). Las regiones anterior y posterior a ésta tienen peor definición, y se reflejan en el *Energy Plot* en las dos zonas con mayor densidad de apareamientos de bases en las estructuras subóptimas (círculos en el *Energy Plot* de la Figura 24 B). La gráfica s-num (con la secuencia poli-C resaltada en verde y el MSCD40LG en azul) indica que tanto la secuencia poli-C como el MSCD40LG se encuentran casi en toda su longitud en forma de hebra sencilla en el 100% de ocasiones en los distintos plegamientos (s-num=1).

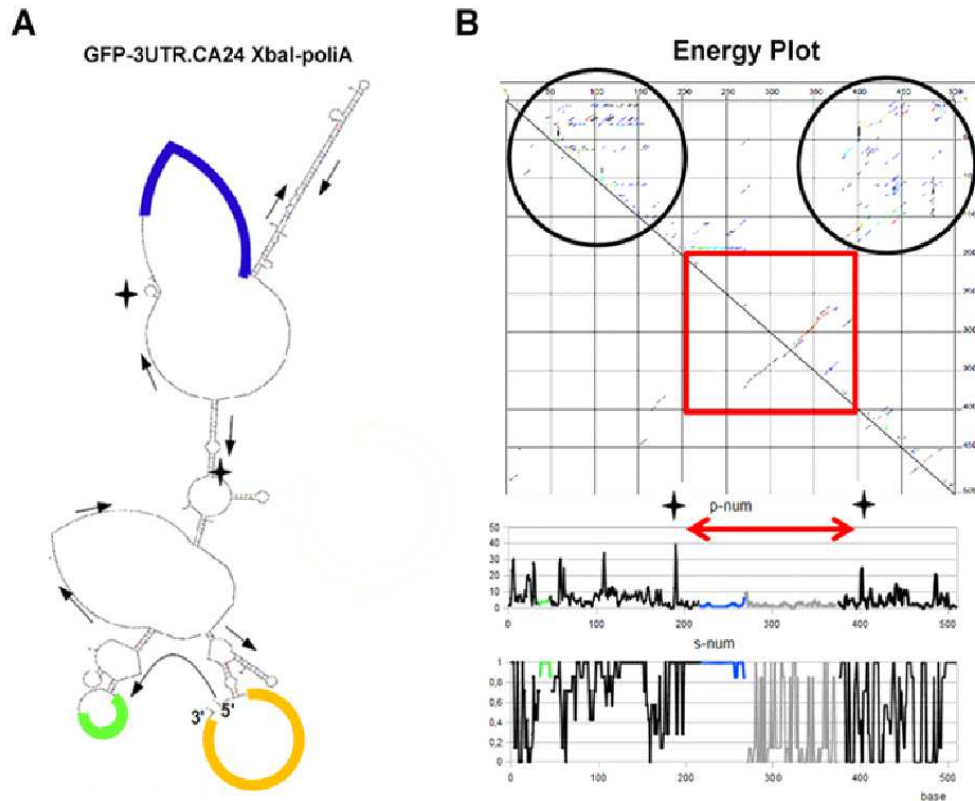


Figura 24. Estructura secundaria de la región 3'UTR del mRNA de CD40LG clonada en GFP-3UTR.CA24. Predicciones de estructuras secundarias generadas en mFold a partir de la secuencia del mRNA GFP-3UTR.CA24 entre XbaI y la cola de poli-A, con 50 adeninas añadidas al extremo 3' para formar una cola de poli-A (naranja). (A) La estructura óptima donde están señaladas la secuencia poli-C (verde) y el MSCD40LG (azul). Las flechas indican la dirección de la secuencia de 5' a 3'. (B) Energy Plot (arriba) y gráficas con los parámetros p-num y s-num a lo largo de la secuencia (ejes X en las gráficas, en paralelo al eje horizontal en la parte superior del Energy Plot). En este último se señalan las regiones con mejor (cuadrado rojo) y peor (círculos negros) definición según estos parámetros.

Al igual que la estructura con la secuencia del alelo 24CA de la Figura 24, generamos estructuras con las secuencias que contenían los otros tres alelos de las construcciones GFP-3UTR.CAn (Figura 25). Estas son prácticamente idénticas (se muestra por comparación la estructura del mRNA GFPΔ3UTR). Esto no apoya la idea de que variaciones importantes en la estructura del mRNA GFP-3UTR.CA23 puedan ser la causa de la falta de caída que muestran los resultados de la cinética.

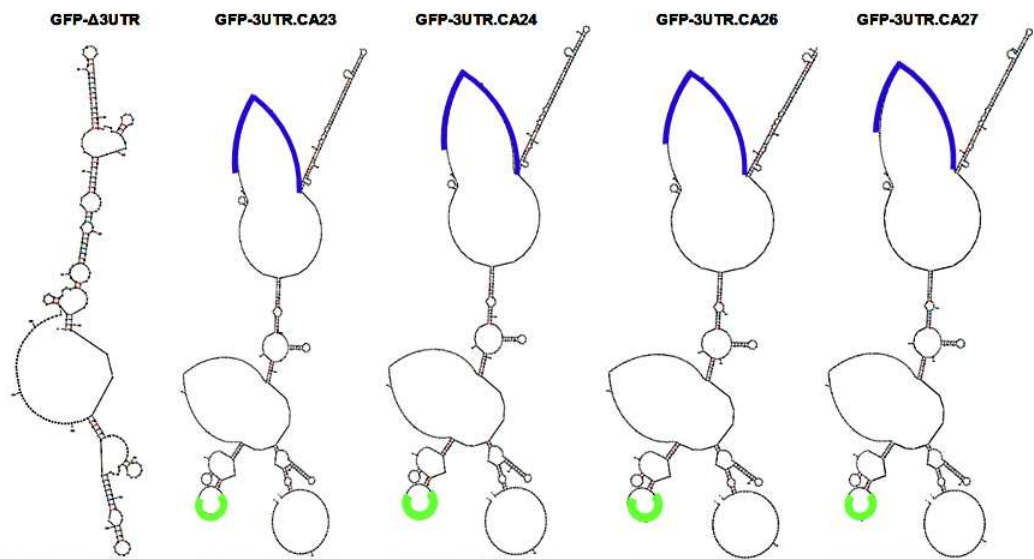


Figura 25. Estructuras secundarias de los mRNA GFP-3UTR.CAn. Las estructuras fueron generadas en el programa mFold con la secuencia 3'UTR de CD40LG clonada en las construcciones GFP-3UTR.CAn, más 50 adeninas añadidas al final como cola de poli-A (Figura 24). Las estructuras mostradas son las calculadas por mFold como óptimas (con el mayor descenso en energía interna (ΔG) en el plegamiento) para las secuencias con los alelos clonados del MSCD40LG para los ensayos de estabilidad del mRNA. El número de nucleótidos (C)_n en la secuencia poli-C y de dinucleótidos (CA)_n en el MSCD40LG está basado en los datos de secuenciación (Tabla 11).

Además de las estructuras generadas a partir de las secuencias con los alelos del MSCD40LG que llevaban las construcciones GFP-3UTR.CAn, generamos estructuras en las mismas condiciones con secuencias con números muy altos o muy bajos de repeticiones (CA)_n, para explorar la resistencia de la configuración desapareada del MSCD40LG y de la estructura donde se encuentra al cambio en la longitud del MSCD40LG (Figura 26). Las estructuras con 14CA y 33CA se corresponden con los alelos más corto y largo que habíamos encontrado en las muestras de DNA tipadas para los estudios de asociación. Éstas se mantienen idénticas a las generadas con los alelos de las construcciones GFP-3UTR.CAn, con el MSCD40LG formando un *loop* más o menos amplio según el número de repeticiones (CA)_n. Ampliando aún más el rango de longitudes, el aumento hasta 40CA sigue produciendo idéntica estructura, mientras que disminuyendo la longitud a 6CA la

estructura óptima cambia (aunque la segunda estructura con menor ΔG (recuadro 6CA S2) es nuevamente idéntica al resto).

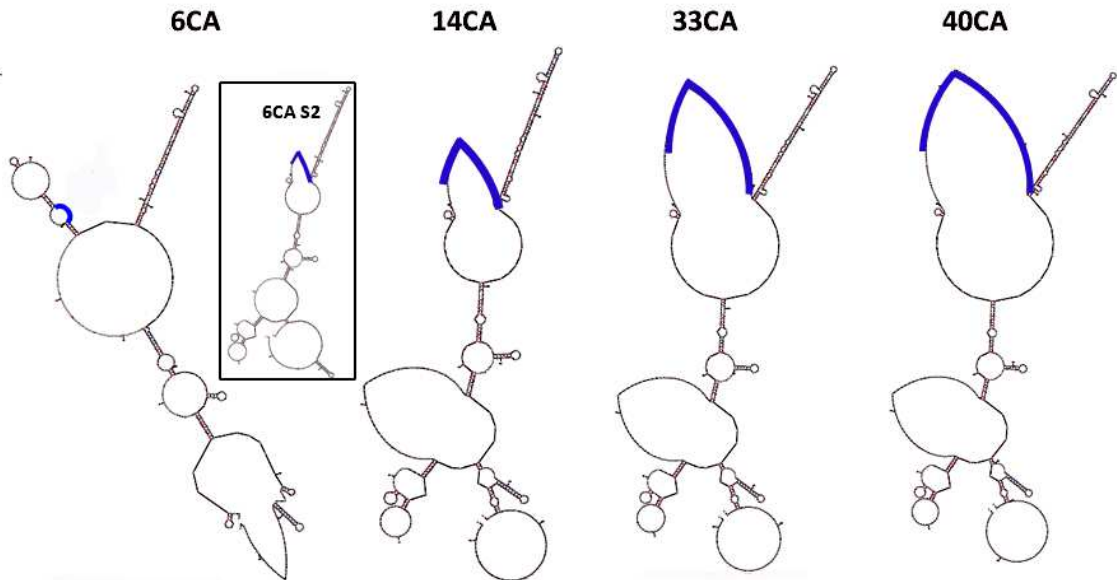


Figura 26. Resistencia de la estructura donde se encuentra el MSCD40LG al cambio en el número de repeticiones (CA)_n. Disminuyendo la longitud a 6CA la estructura óptima cambia, aunque 3 de las 6 repeticiones CA siguen estando desapareadas. La segunda estructura con menor ΔG (6CA S2) es idéntica al resto.

Las estructuras generadas para las construcciones GFP-3UTR.CA_n que aparecen en las figuras 23 a 25 están obtenidas mostrando posibles plegamientos con un ΔG como máximo un 5% mayor que el ΔG de la estructura óptima (el porcentaje se maneja con un parámetro llamado *percent suboptimality*, que varía la diferencia en el ΔG del plegamiento respecto a la estructura óptima de los plegamientos que son incluidos en el *Energy Plot* hasta un límite de 12 kcal mol⁻¹. En estas secuencias, el 5% del ΔG óptimo supone plegamientos a una distancia máxima de unas 4 kcal/mol del ΔG óptimo (de unas -75 kcal/mol). Relajando este parámetro al 25% se incluyen todas las estructuras hasta la diferencia máxima de 12 kcal/mol, y las estructuras generadas con las secuencias GFP-3UTR entre XbaI y la cola de poli-A pasan de 7 a 14. Con estas condiciones el MSCD40LG sigue apareciendo prácticamente en toda su longitud en hebra sencilla en el 100% de las estructuras generadas. El MSCD40LG también aparece en forma de hebra sencilla las

predicciones de estructuras obtenidas a partir de las secuencias completas del mRNA de CD40LG. Todos estos datos indican que hay una alta probabilidad de que el MSCD40LG se encuentre desapareado prácticamente en su totalidad en la estructura del mRNA formando un *loop* amplio, accesible para la unión de proteínas reguladoras.

•MSCD40LG y procesamiento del pre-mRNA

Durante la preparación de los ensayos de caída del mRNA, tras la obtención de los clones de la línea HEK293 con las construcciones GFP-3UTR.CAn y antes de realizar las cinéticas de caída del mRNA, hicimos pruebas de las tres PCR (GFP-3UTR.CAn, GADPH y c-fos) con las que íbamos a realizar las medidas de los mRNA GFP-3UTR.CAn, GADPH y c-fos. Aislamos RNA de células de los clones HEK293 en cultivo y tras tratar con actinomicina D a diferentes tiempos y comprobamos en geles de agarosa que las tres PCR daban una banda única del tamaño esperado. Así fue en todos los clones para las PCR GADPH y c-fos. En la PCR GFP3UTR que medía los c-DNA GFP-3UTR.CAn y GFP- Δ 3UTR, aparecía una banda en las PCR realizadas sobre los c-DNA GFP-3UTR.CAn, de tamaño correspondiente a la PCR sobre la secuencia del mRNA sin el intrón artificial que lleva en el extremo 5' transcrito, es decir, sobre la secuencia correspondiente al mRNA maduro, que ha perdido el intrón artificial por *splicing* en el procesamiento en el núcleo (Figura 27).

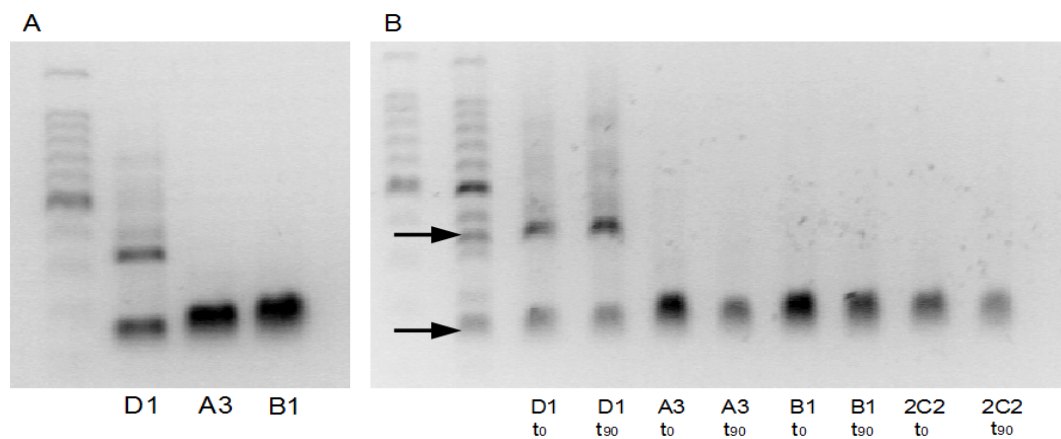


Figura 27. PCR GFP₃UTR realizada en los cDNAs de los clones HEK293 con distintas construcciones GFP-3UTR.CAn y GFP Δ 3UTR. (A) En la primera calle se muestra el patrón de tamaño, con la banda de 500 pb más intensa que el resto y bandas sucesivas con diferencia de 100 pb. En el resto de las calles se muestra el producto de las PCR GFP₃UTR en los clones D₁ (GFP- Δ 3UTR), A₃ y B₁ (los dos GFP-3UTR.CA₂₇). (B) En la primera calle se muestra el patrón de tamaño, en la segunda el patrón de tamaño mezclado con el producto de la PCR GFP₃UTR del clon D₁ y en las siguientes la PCR GFP₃UTR en los clones D₁, A₃, B₁ y 2C₂ (GFP-3UTR.CA₂₄) tras 0 min (t₀) ó 90 min (t₉₀) de tratamiento con Act-D. Las flechas superior e inferior corresponden al tamaño esperado de la PCR sobre la secuencia con y sin intrón (293 y 161 pb respectivamente).

Sin embargo, en las PCR GFP3UTR realizadas sobre los c-DNAs de la construcción GFP- Δ 3UTR aparecían dos bandas, de intensidad similar y con tamaños que se correspondían con los esperados para las PCR realizadas sobre la secuencia de mRNA GFP- Δ 3UTR con y sin el intrón artificial (Figura 27). Esto supone que el c-DNA tenía copias de los pre-mRNAs GFP- Δ 3UTR (secuencia con el intrón) y del mRNA maduro (sin el intrón). Las tres RNP descritas que se unen a la región 3'UTR de CD40LG en los mRNA GFP-3UTR.CAn (Nucleolin, PTB y hnRNPL) se sabe que participan en varias funciones relacionadas con el procesamiento y la regulación de la expresión de mRNA. Dado que en las células en cultivo el mRNA GFP-3UTR.CAn se transcribe de forma continua, la presencia de una de pre-mRNA GFP- Δ 3UTR en equilibrio con una cantidad similar de mRNA procesado, mientras que en los mRNA GFP-3UTR.CAn no aparece una cantidad apreciable de pre-mRNA, sugiere que la velocidad de procesamiento del mRNA GFP-3UTR.CAn aumenta de forma apreciable con la presencia de la región 3'UTR de CD40LG clonada en estas construcciones.

DISCUSIÓN

El componente genético que define la predisposición a desarrollar enfermedades autoinmunes como LES o AR es complejo, está compuesto de múltiples factores con escasa penetrancia y aún no está completamente definido a pesar del gran esfuerzo que se ha dedicado a esta tarea desde hace varias décadas. Estudios realizados en nuestro laboratorio describieron asociación del MSCD40LG a LES y AR, en la que en ambas enfermedades el alelo mayoritario en la población (24CA) se encontraba con menor frecuencia en pacientes que en controles. Además, linfocitos T CD4 de individuos con alelos 24CA tenían una expresión diferente CD40LG que aquellos con alelos diferentes, lo que apoyaba la idea de que el MSCD40LG puede representar una variante causal en la predisposición genética a desarrollar estas enfermedades. En este trabajo hemos mostrado que la asociación del MSCD40LG descrita en los trabajos anteriores tenía gran similitud y se ajustaba a un modelo codominante donde el alelo 24CA confiere un menor riesgo en función de la presencia de 0,1 o 2 copias. Hemos revisado la asociación del MSCD40LG a LES y AR tras la inclusión de nuevas muestras de controles sanos y pacientes de LES. Además, hemos realizado un ensayo para tratar de ver si las variantes alélicas comunes del MSCD40LG influyen en la estabilidad del mRNA de este gen.

Los resultados de este trabajo en cuanto a la asociación marcada por el alelo 24CA frente al resto confirman en lo esencial lo descrito en los trabajos previos, con el alelo 24CA como alelo que confiere menor riesgo. Los resultados de asociación a AR son muy parecidos, como es de esperar ya que sólo ha cambiado respecto al estudio inicial el incremento en la serie de controles con la inclusión de la serie DSC2. El aumento de la serie de controles se traduce en un aumento del poder estadístico en las comparaciones, y en un ajuste más claro a un modelo de penetrancia multiplicativo. En cuanto a los resultados de asociación a LES, en la pequeña serie nueva de LES peninsulares la asociación se replica respecto al estudio inicial en población canaria. En población canaria presentan mayores diferencias con los anteriores con la inclusión de las nuevas muestras de pacientes de LES como de controles. Se mantiene la asociación marcada por el alelo 24CA como alelo de menor riesgo, pero en los genotipos contruidos agrupando el alelo 24CA frente al resto de alelos (2424, 24X y XX), hay un RR similar para los genotipos 2424 y 24X, y mayor para el genotipo XX, lo cual implica dominancia del alelo 24CA. A pesar de que según los resultados de este trabajo la asociación del alelo 2CA a AR y LES se ajusta de forma clara a modelos diferentes de penetrancia, los tamaños muestrales no permiten distinguir con precisión si esa diferencia es real, ya que las series de LES y AR son homogéneas cuando se comparan las frecuencias de los genotipos 2424, 24X, XX, por lo que estadísticamente las

series ARC1 y LESC12 podrían surgir de unas frecuencias genotípicas comunes. Lo mismo sucede con las series LESC1 y LESC2, que resultan estadísticamente homogéneas aunque en LESC2 se aprecia una mayor frecuencia de genotipos 2424 que conduce al cambio en el modelo de penetrancia al que se ajusta la asociación en la serie ampliada LESC12. La limitación del tamaño muestras impide distinguir si otras causas aparte de la distribución en el muestreo pueden estar detrás de estas diferencias. La medida del exceso de homocigotos se utiliza para comprobar si una población se encuentra en equilibrio de *Hardy-Weinberg*. En una muestra de pacientes de una enfermedad, un marcador asociado se encuentra en equilibrio de *Hardy-Weinberg* cuando la asociación sigue un modelo multiplicativo ¹⁴⁵. Este es el caso de las series ARC1 y LESC1, pero no de LESC2, que tiene una mayor frecuencia de homocigotos de lo esperado bajo equilibrio de *Hardy-Weinberg*. Una causa que puede estar detrás de un aumento de homocigotos sobre lo esperado es la presencia de consanguineidad en la muestra. Aunque las pacientes de LES de las series LESC1 y LESC2 no están emparentadas en principio, hay una mayor probabilidad de compartir haplotipos con una ancestría común entre pacientes de una enfermedad con un componente hereditario que entre la población sana (REF: Am. J. Hum. Genet. 2000, 67:258), que es más mayor cuanto menor es la incidencia de la enfermedad, mayor la asociación familiar y menor es la población. Por otro lado, al tratarse de un marcador asociado a enfermedad, el aumento de homocigotos podría deberse a que, como muestran los resultados, la asociación a LES no se rija por un modelo multiplicativo. La limitación en la capacidad de discriminación debida al tamaño muestral impide determinar con precisión la posible contribución de estos fenómenos al resultado y si el alelo 24CA se comporta como alelo dominante en LES a diferencia de en AR.

Además de la asociación marcada por el alelo 24CA, en este trabajo describimos que los alelos minoritarios (<21CA y >27CA) agrupados confieren un mayor RR que el conjunto de alelos distintos del 24CA. Este RR es similar para AR y LES en el conjunto de alelos minoritarios, aunque en LES parecen estar especialmente representados los alelos largos (>27CA). Los alelos X (distintos del alelo 24CA) siguen mostrando asociación sin tener en cuenta a los alelos minoritarios, lo que indica que estos últimos no son los responsables de toda la asociación. En conjunto los resultados parecen indicar que el riesgo conferido por alelos diferentes al 24CA no es homogéneo, aunque no es posible comparar de forma fiable el RR conferido por alelos individuales, que tienen una baja frecuencia.

Posiblemente la principal limitación a la fiabilidad de la asociación del MSCD40LG que hemos encontrado esté en la falta de réplica en estudios independientes, aparte de la pequeña serie de LES peninsulares que describimos en este trabajo, que replica la asociación del alelo 24CA aunque con un tamaño escaso, y de los mencionados en la introducción ^{129,130}. En los años 90 aparecieron los primeros estudios de asociación de marcadores a lo largo del genoma completo a diversas enfermedades, entre las que se encuentran el LES y la AR (GWAS, *Genome Wide Association Studies*). Los primeros GWAS utilizaban microsatélites como marcadores, aunque pocos de ellos incluían los cromosomas sexuales, y en ninguno que sepamos el MSCD40LG formaba parte de los marcadores en el cromosoma X. Los marcadores asociados encontrados en estos primeros GWAS fallaban con gran frecuencia en ser replicados en estudios independientes ^{152,153}. La limitación en el tamaño muestral unida a la débil fuerza de la asociación de las variantes individuales que componen la predisposición genética en estas enfermedades parecía ser la causa principal de la falta de reproducibilidad en los estudios de asociación genética. Para solventar esta limitación, grandes consorcios (como NARAC y BRASS en EEUU, EIRA en Suecia o WTCC en Reino Unido) crearon colecciones con un elevado número de muestras de controles y pacientes de diversas enfermedades. En paralelo, los microsatélites fueron sustituidos como marcadores por polimorfismos de cambio de base (SNP *Single Nucleotide Polymorphisms*), que aunque menos informativos uno a uno por su menor variedad alélica, son mucho más numerosos que los microsatélites y han permitido la selección de un gran número de ellos (10^5 - 10^6) a lo largo del genoma completo, disminuyendo en gran medida la distancia entre marcadores y por tanto la distancia teórica entre un marcador asociado y genes candidatos cercanos, lo que aumenta el poder de detección de asociaciones de carácter débil. Una recopilación de estos estudios GWAS de asociación a numerosas enfermedades y caracteres fenotípicos se encuentra en el NHGRI *GWAS Catalog* (www.genome.gov/gwastudies). Los estudios GWAS de asociación a enfermedades autoinmunes han ido creciendo en número, y a su vez manejando muestras de controles y pacientes cada vez mayores, alcanzado tamaños del orden de 10^3 - 10^4 individuos. Los grandes tamaños muestrales, junto con la réplica de GWAS en muestras independientes y el desarrollo de criterios comunes de análisis estadístico de la asociación han llevado a la identificación de decenas de marcadores asociados a enfermedades como LES o AR ¹⁵⁴. En los estudios GWAS publicados no se ha encontrado ningún SNP en el gen CD40LG asociado a AR ni a LES. Este hecho no desmiente la asociación del MS40LG encontrada por nosotros, ya que no está descrita la estructura haplotípica de las variantes de SNPs junto con las del MS y no se

conoce la distribución de las variantes del MS en relación a los haplotipos formados por los SNPs ¹⁵⁵.

A pesar de que los estudios GWAS tienen una gran capacidad de detección de asociación, y han revelado un buen número de variantes comunes con asociación débil a enfermedades autoinmunes el componente hereditario de enfermedades autoinmunes sigue estando definido de forma incompleta ¹⁵⁶ y se ha estimado que el total de variantes asociadas a este tipo de enfermedades con un componente genético complejo podría estar en el orden de 10^2 - 10^3 ¹⁵⁷. Para completar en la medida de lo posible el cuadro de la asociación genética a enfermedades autoinmunes será necesario identificar con precisión el mayor número posible de variantes asociadas y la magnitud de su efecto en la probabilidad de desarrollar enfermedad, variantes que además de SNPs consisten en otros tipos como son microsatélites ¹⁵⁸, CNVs (Copy Number Variants), deleciones, inversiones, etc, y que incluyen tanto las variantes comunes, que son las representadas en los GWAS, como variantes raras con escasa frecuencia. Un nivel adicional de complejidad lo componen las posibles interacciones entre variantes situadas en distintos genes que hoy en día representan un gran reto para su detección dado el gran número de potenciales interacciones, y requieren del desarrollo de herramientas informáticas con una gran capacidad de cálculo, desarrollo que en algún caso ha conseguido la identificación de epistasia entre varios marcadores ¹⁵⁹.

Hasta el momento, una gran parte de los SNP asociados a enfermedad detectados en los GWAS se sitúan en regiones intergénicas o intrónicas de función desconocida ¹⁶⁰¹⁶¹. Aunque esto no significa que no tengan funcionalidad, ya que aún falta mucho por conocer acerca de la regulación de la expresión génica y el papel que las secuencias no codificantes juegan en ésta, se presupone que algunas de las variantes asociadas son “acompañantes” de variantes causales, ligadas a ellas por su cercanía en el genoma. En este sentido, y dada la complejidad de la tarea de averiguar cómo funciona la composición genética en la predisposición a enfermedad, es importante encontrar las variantes causales y poder describir cómo éstas afectan a la expresión génica, de forma que pueda explicarse el mecanismo mediante el cual modifican en una pequeña cantidad la probabilidad de desarrollar enfermedad.

La asociación del MS40LG a AR y LES vista en nuestros trabajos resulta interesante por la situación del microsatélite en la región 3'UTR del mRNA, en una zona importante para la regulación de la expresión transitoria que se da en linfocitos T en respuesta a la

activación a través del receptor de antígeno. Dada la situación del microsatélite en esta región importante para la regulación de la estabilidad del mRNA y que el control de la estabilidad forma parte de la regulación de la expresión de este gen, pensamos que las variantes alélicas del MS40LG podrían tener un significado funcional a través de un efecto diferencial en la estabilidad del mRNA, que podría formar la base de la asociación genética a AR y LES que habíamos encontrado. En los trabajos previos vimos diferencias en la expresión de CD40LG en la superficie de linfocitos T CD4 activados in vitro con PHA o anticuerpos anti CD3 y CD28 ^{98,132}, y también diferencias en la estabilidad del mRNA entre muestras homocigotas para el alelo 24CA y portadoras de alelos diferentes. Para aislar el efecto que pudiera tener el MS40LG en la estabilidad del mRNA y tratar de definirlo con mayor precisión, clonamos en vectores de expresión la secuencia de la región 3'UTR de CD40LG que regula la estabilidad del mRNA, región que contiene también el MSCD40LG, de forma que las construcciones sólo diferían entre sí en el alelo del microsatélite. Insertamos éstas de forma estable en una línea celular para tratar de conseguir un sistema homogéneo donde medir las diferencias en la estabilidad de los mRNA en función de los diferentes alelos.

De acuerdo con los resultados del presente trabajo la diferencia en la longitud entre variantes alélicas comunes de 24CA, 26CA y 27CA del MSCD40LG es suficiente para variar la estabilidad de un mRNA que contiene la región 3'UTR del mRNA CD40LG. Al contrario de lo que habíamos visto en un estudio anterior ⁹⁸, según los resultados de este trabajo el alelo 24CA es el que confiere menor estabilidad a los mRNA GFP-3UTR.CAn. Aunque diferencias importantes entre los dos trabajos que hay que tener en cuenta, el resultado en sentido contrario a lo que habíamos visto previamente parece en principio difícil de conciliar. La medida de estabilidad en función de los alelos del MSCD40LG en este trabajo tiene la ventaja de estar realizada en un sistema homogéneo donde los mRNA se expresan en una misma línea celular de forma constitutiva, de forma que es de esperar que el mRNA GFP-3UTR.CAn se encuentre en equilibrio y la medida del $t_{1/2}$ de los mRNA no esté influida por posibles variaciones en el tiempo de estabilización del mRNA de CD40LDG en el curso de la activación de linfocitos T, diferencias que pueden darse entre muestras de diferentes individuos. En este sentido el ensayo realizado en este trabajo podría en principio tener mayor capacidad para aislar el efecto del MSCD40LG en la estabilidad del mRNA, y comparar mejor el efecto de los alelos individuales en el caso de los alelos distintos al 24CA.

Por otro lado, el ensayo mediante el clonaje de una parte del gen en una línea celular obviamente supone la expresión de la región 3'UTR del mRNA de CD40LG fuera de su contexto natural en la molécula nativa de mRNA de CD40LG y del tipo celular donde se expresa de forma normal, y no podemos asegurar que alguno de los factores que afecten a la regulación de la expresión del mRNA CD40LG, que son muy numerosos teniendo en cuenta todos los procesos que ocurren desde su transcripción hasta su degradación final, estén presentes de forma específica en linfocitos T activados pero no en la línea HEK293 donde hemos realizado los ensayos. En relación con esto, las RNPs PTB, Nucleolin y hnRNPL, que se unen a la región 3'UTR del mRNA de CD40LG, se expresan en todas las células y están también en la línea HEK293. En ensayos similares al realizado en este trabajo, un mRNA con la región 3'UTR de CD40LG clonada junto a la luciferasa reduce la expresión de ésta en igual proporción en la línea celular Jurkat D1.1, que proviene de una leucemia T (LT CD4⁺) y en la línea celular HeLa, proveniente de un tumor de cuello del útero, lo que indica que la maquinaria que controla la expresión y el recambio de mRNA parece estar conservada en los diferentes tipos celulares. Parece, por tanto, que el ensayo en la línea HEK293 es válido para medir el cambio en la estabilidad del mRNA en función del alelo del MSCD40LG, aunque obviamente las diferencias encontradas entre los alelos del MSCD40LG, aunque puedan dar una medida del efecto de las variantes alélicas comunes en la estabilidad del mRNA, no son directamente trasladables a las que pueda haber en el mRNA CD40LG completo.

En los ensayos de estabilidad del mRNA realizados en este trabajo, aunque las medidas de la caída de los mRNA no han conseguido gran precisión, los $t_{1/2}$ de los mRNA control de c-fos son bastante similares, y la variabilidad en esta medida da una idea de la reproducibilidad del ensayo. Los $t_{1/2}$ de los mRNA GFP-3UTR.CAn tienen diferencias notablemente mayores, lo que indica un efecto de las variantes alélicas del MSCD40LG por encima de la variabilidad del ensayo que estima la medida de los m-RNAs c-fos, y apoya la idea de que el número de repeticiones (CA)_n del MSCD40LG afecta a la estabilidad del mRNA. El mRNA control GFP-Δ3UTR tiene una gran estabilidad, y la región 3'UTR de CD40LG clonada en los mRNA mRNAGFP-3UTR.CAn, con la excepción del alelo 23CA, cuyo resultado parece anómalo, reduce la estabilidad del mRNA al igual que en otros trabajos que han utilizado sistemas similares con la región 3'UTR de CD40LG⁹³. Los tiempos medios obtenidos, con un rango de 114-390 min, también son similares a los encontrados en estudios con un diseño similar en líneas celulares, y están en el orden de los $t_{1/2}$ que se encuentra en linfocitos T activados cuando el mRNA de CD40LG se estabiliza

(Tabla 2). El complejo de proteínas que se une a la región donde se sitúa el MSCD40LG (Complex II, Figura 5) se une al mRNA de CD40LG en su fase estable, por lo que el mRNA estabilizado parece el más idóneo para medir el efecto del MSCD40LG en la estabilidad. El hecho de que el alelo 24CA sea según los resultados de este trabajo el de menor estabilidad encaja con la menor expresión de CD40LG en la superficie de linfocitos T CD4⁺ con alelos 24CA frente a alelos distintos que vimos en nuestros trabajos previos, suponiendo igualdad en la velocidad de transcripción y una relación directa entre la cantidad de mRNA y la de proteína producida, relación que se encuentra de forma general ¹⁶². Esto sucede en la línea Jurkat D1.1, en la que la disminución de la expresión de PTB mediante RNAs de interferencia disminuye la estabilidad del mRNA de CD40LG, y en paralelo la cantidad de mRNA y de proteína en la superficie ¹⁶³. En nuestro trabajo previo, encontrábamos un número de moléculas de mRNA que era menor en las muestras con el alelo 24CA, tanto a 6h como a 24h tras activación de los LT. Sin embargo, comparando las 6h con las 24h, tanto en muestras 24CA como con alelos distintos tras 6h, con el mRNA inestable, había más cantidad de mRNA CD40LG que a 24h. La mayor cantidad de mRNA de CD40LG en linfocitos T a tiempos cortos tras la activación, cuando el mRNA se encuentra en la fase inestable, se ha descrito en otros trabajos ¹⁶⁴. Esto parece en principio chocar con una correlación directa entre la cantidad de mRNA y de proteína, pero teniendo en cuenta que la expresión de CD40LG se inicia con la activación del LT, la menor cantidad de proteína a 6h que a 24h puede deberse al lapso de tiempo que transcurre entre la transcripción del mRNA, la traducción a proteína y el transporte de ésta a la membrana. Este tipo de correlación con retardo entre la cantidad de mRNA y de proteína se da en un gran número de genes ¹⁶⁵.

En cuanto al posible mecanismo mediante el cual el número de repeticiones (CA)_n del MSCD40LG puede modificar la estabilidad del mRNA, el conocimiento que se ha obtenido en los últimos años acerca de la RNP hnRNPL sitúa a ésta como la candidata más probable para conectar diferencias en la longitud del MSCD40LG con cambios en la estabilidad del mRNA. Las predicciones de la estructura secundaria de la región 3'UTR de CD40LG que se muestran en este trabajo indican que con gran probabilidad la secuencia del MSCD40LG se encuentra en el mRNA de CD40LG desapareada en casi toda su longitud, formando un loop amplio que la hace accesible a la unión de proteínas reguladoras. hnRNPL se une a RNA en forma de cadena sencilla, pero no de doble cadena ¹⁶⁶, y con gran afinidad a repeticiones (CA)_n ^{167,168}. El mRNA de CD40LG se encuentra entre los mRNA más representados entre los que se unen a esta RNP en linfocitos T activados ¹⁶⁸.

Se ha descrito cómo hnRNPL tiene un papel en la regulación de la estabilidad de algunos mRNA a través de la unión a repeticiones (CA)_n ¹⁶⁹, regulando la expresión de genes importantes para la función inmune como bcl-2 ¹⁷⁰ o VEGF ¹⁷¹. En el mRNA de CD40LG hnRNPL forma, junto con nucleolín y PTB, un complejo que aparece unido al mRNA cuando se encuentra en la fase estable. Se ha descrito (Hahm et al., 1998) la unión entre PTB y hnRNPL, y los dominios de PTB necesarios para esta unión ¹⁷² no se encuentran en la isoforma PTB-T, que se asocia con el mRNA de CD40LG en la fase inestable ⁹³. Esto apunta a la interacción entre PTB y hnRNPL en la regulación de la estabilidad del mRNA, interacción que se ha descrito en el mRNA de CD40LG de ratón ⁹⁴.

Aunque con los datos de los que disponemos en este momento no es posible afirmar que el en la expresión de CD40LG que hemos visto en función de las variantes alélicas del MSCD40LG puedan dar lugar a cambios en la probabilidad de desarrollar AR o LES, los resultados de este trabajo y de los anteriores realizados en nuestro laboratorio apuntan en esa dirección. Cambios en la función mediada por hnRNPL mediados por una interacción diferencial en función del número de repeticiones (CA)_n de un microsatélite se han descrito en otros genes ¹⁶⁹. Teniendo en cuenta la probable disposición en un *loop* desapareado del MSCD40LG en un rango amplio de repeticiones (CA)_n, cambios en la afinidad por hnRNPL podrían ser la base del mecanismo de acción del MSCD40LG en la expresión del mRNA CD40LG. La cepa murina congénica FVB proporciona un ejemplo de cómo la un fenómeno así puede darse ¹⁷³. Esta cepa expresa aproximadamente la mitad de integrina $\alpha 2\beta 1$ en la superficie de las plaquetas que otras cepas como C57/BL6, Balb/c, C3H/He o 129/sv, que tienen una expresión muy parecida entre ellas. La causa de la menor expresión de la integrina $\alpha 2\beta 1$ en la cepa FVB reside en un microsatélite (CA)_n situado en el gen de la integrina $\alpha 2$ en el intrón 1. La cepa FVB tiene 6 repeticiones CA frente a 21 repeticiones en el resto de cepas, y esta diferencia lleva a una menor afinidad por hnRNPL, que reduce la cantidad de mRNA maduro productivo por una alteración en el splicing del pre-mRNA.

Por último, hnRNPL tiene varias funciones en el procesamiento de pre-mRNAs en el núcleo. Regula de esta forma la expresión de genes importantes en el desarrollo y función de linfocitos T como CD45 ¹⁷⁴. La diferente proporción entre pre-mRNA y mRNA maduro de los mRNAs GFP Δ 3UTR y GFP-3UTR.CAn que hemos visto en este trabajo apuntan a que la región 3'UTR del gen CD40LG podría tener un papel en el procesamiento del pre-mRNA, aumentando la eficiencia del mismo, probablemente a través de la unión a hnRNPL al MSCD40LG.

Aún queda mucho por hacer para poder definir con precisión la regulación de la expresión de C40LG, cómo las variantes alélicas del MSCD40LG pueden modificar la expresión de este gen en el contexto del desarrollo de linfocitos T en el timo o en la respuesta inmune, y cómo el cambio en la expresión en función de las variantes alélicas puede llevar a un cambio en la probabilidad de desarrollar una enfermedad autoinmune como LES o AR.

CONCLUSIONES

- 1- El microsatélite del gen CD40LG muestra, de acuerdo con nuestros resultados previos, asociación genética a artritis reumatoide y a lupus eritematoso sistémico en mujeres en población canaria, con el alelo mayoritario 24CA como el alelo de menor riesgo comparado con el resto de alelos.
- 2- Considerando la presencia de dos, uno o ningún alelo 24CA, a diferencia de nuestros resultados previos, la asociación genotípica no parece ajustarse al mismo modelo de penetrancia en artritis reumatoide y lupus eritematoso sistémico.
- 3- Los alelos minoritarios del microsatélite, considerados en un único grupo, aparecen con mayor frecuencia tanto en enfermas de artritis como de lupus eritematoso sistémico, con un riesgo relativo similar en las dos enfermedades, mayor que el conferido por las variantes comunes.
- 4- Los ensayos cinéticos de caída del mRNA indican que la longitud del dinucleótido $(CA)_n$ que forma el microsatélite de CD40LG afecta de modo directo a la estabilidad del RNA mensajero.
- 5- El análisis de la estructura secundaria del microsatélite de CD40LG indica que éste se encuentra con gran probabilidad desapareado, formando una diana para la ribonucleoproteína hnRNPL, cuya afinidad variable en función de las repeticiones $(CA)_n$ podría ser la base del cambio en la estabilidad del mRNA.
- 6- El microsatélite de CD40LG podría tener un papel en el procesamiento del pre-mRNA de CD40LG a través de la interacción con la ribonucleoproteína hnRNPL.

BIBLIOGRAFÍA

1. Pietravalle F, Lecoanet-Henchoz S, Blasey H, Aubry J-P, Elson G, Edgerton MD, et al. Human native soluble CD40L is a biologically active trimer, processed inside microsomes. *J Biol Chem*. 1996;271(11):5965-7.
2. Reinboldt S, Wenzel F, Rauch BH, Hohlfeld T, Grandoch M, Fischer JW, et al. Preliminary evidence for a matrix metalloproteinase-2 (MMP-2)-dependent shedding of soluble CD40 ligand (sCD40L) from activated platelets. *Platelets*. 2009;20(6):441-4.
3. Yacoub D, Benslimane N, Al-Zoobi L, Hassan G, Nadiri A, Mourad W. CD154 is released from T-cells by a disintegrin and metalloproteinase domain-containing protein 10 (ADAM10) and ADAM17 in a CD40 protein-dependent manner. *J Biol Chem*. 2013 Dec 13;288(50):36083-93.
4. Karpusas M, Hsu YM, Wang JH, Thompson J, Lederman S, Chess L, et al. 2 A crystal structure of an extracellular fragment of human CD40 ligand. *Structure*. 1995 Dec 15;3(12):1426.
5. Peitsch MC, Jongeneel CV. A 3-D model for the CD40 ligand predicts that it is a compact trimer similar to the tumor necrosis factors. *Int Immunol*. 1993;5(2):233-8.
6. Bartlett WC, McCann J, Shepherd DM, Roy M, Noelle RJ. Cognate interactions between helper T cells and B cells. IV. Requirements for the expression of effector phase activity by helper T cells. *J Immunol*. 1990;145(12):3956-62.
7. Yellin MJ, Lee JJ, Chess L, Lederman S. A human CD4+ T cell leukemia subclone with contact-dependent helper function. *J Immunol*. 1991;147:3389-95.
8. Lederman S, Yellin MJ, Krichevsky A, Belko J, Lee JJ, Chess L. Identification of a novel surface protein on activated CD4+ T cells that induces contact-dependent B cell differentiation (help). *J Exp Med*. 1992;175(4):1091-101.
9. Armitage RJ, Fanslow WC, Strockbine L, Sato TA, Clifford KN, Macduff BM, et al. Molecular and biological characterization of a murine ligand for CD40. *Nature*. 1992;357:80-2.
10. Noelle RJ, Roy M, Shepherd DM, Stamenkovic I, Ledbetter JA, Aruffo A. A 39-kDa protein on activated helper T cells binds CD40 and transduces the signal for cognate activation of B cells. *Proc Natl Acad Sci U S A*. 1992;89(14):6550-4.
11. Graf D, Korthäuer U, Mages HW, Senger G, Kroczeck RA. Cloning of TRAP, a ligand for CD40 on human T cells. *Eur J Immunol*. 1992 Dec;22(12):3191-4.
12. Aruffo A, Farrington M, Hollenbaugh D, Li X, Milatovich A, Nonoyama S, et al. The CD40 ligand, gp39, is defective in activated T cells from patients with X-linked hyper-IgM syndrome. *Cell*. 1993;72:291-300.
13. Fuleihan R, Ramesh N, Loh R, Jabara H, Rosen RS, Chatila T, et al. Defective expression of the CD40 ligand in X chromosome-linked immunoglobulin deficiency with normal or elevated IgM. *Proc Natl Acad Sci USA*. 1993;90:2170-3.
14. Sayle RA, Milner-White EJ. RASMOL: biomolecular graphics for all. *Trends Biochem Sci*. 1995;20(9):374.

15. Foy TM, Laman JD, Ledbetter JA, Aruffo A, Claassen E, Noelle RJ. gp39-CD40 interactions are essential for germinal center formation and the development of B cell memory. *J Exp Med*. 1994;180:157-63.
16. Han S, Hathcock K, Zheng B, Kepler TB, Hodes R, Kelsoe G. Cellular interaction in germinal centers. Roles of CD40 ligand and B7-2 in established germinal centers. *J Immunol*. 1995;155:556-67.
17. Van den Eertwegh AJ, Noelle RJ, Roy M, Shepherd DM, Aruffo A, Ledbetter JA, et al. In vivo CD40-gp39 interactions are essential for thymus-dependent humoral immunity. I. In vivo expression of CD40 ligand, cytokines, and antibody production delineates sites of cognate T-B cell interactions. *J Exp Med*. 1993;178(5):1555-65.
18. Grewal IS, Xu J, Flavell RA. Impairment of antigen-specific T-cell priming in mice lacking CD40 ligand. *Nature*. 1995;378(6557):617-20.
19. Kawabe T, Naka T, Yoshida K, Tanaka T, Fujiwara H, Suematsu S, et al. The immune responses in CD40-deficient mice: impaired immunoglobulin class switching and germinal center formation. *Immunity*. 1994;1(3):167-78.
20. Klaus SJ, Berberich I, Shu G, Clark EA. CD40 and its ligand in the regulation of humoral immunity. *Semin Immunol*. 1994;6:279-86.
21. Levitt D, Haber P, Rich K, Cooper MD. Hyper IgM immunodeficiency. A primary dysfunction of B lymphocyte isotype switching. *J Clin Invest*. 1983;72(5):1650-7.
22. Borrow P, Tishon A, Lee S, Xu J, Grewal IS, Oldstone MB, et al. CD40L-deficient mice show deficits in antiviral immunity and have an impaired memory CD8⁺ CTL response. *J Exp Med*. 1996;183(5):2129-42.
23. Van Kooten C, Banchereau J. CD40-CD40 ligand. *J Leukoc Biol*. 2000;67(1):2-17.
24. Cella M, Scheidegger D, Palmer-Lehmann K, Lane P, Lanzavecchia A, Alber G. Ligation of CD40 on dendritic cells triggers production of high levels of interleukin-12 and enhances T cell stimulatory capacity: T-T help via APC activation. *J Exp Med*. 1996;184:747-52.
25. Kiener PA, Moran-Davis P, Rankin BM, Wahl AF, Aruffo A, Hollenbaugh D. Stimulation of CD40 with purified soluble gp39 induces proinflammatory responses in human monocytes. *J Immunol*. 1995;155(10):4917-25.
26. Ma DY, Clark EA. The role of CD40 and CD154/CD40L in dendritic cells. *Semin Immunol*. 2009;21(5):265-72.
27. Malik N, Greenfield BW, Wahl AF, Kiener PA. Activation of human monocytes through CD40 induces matrix metalloproteinases. *J Immunol*. 1996;156(10):3952-60.
28. Bennett SR, Carbone FR, Karamalis F, Flavell RA, Miller JF, Heath WR. Help for cytotoxic-T-cell responses is mediated by CD40 signalling [see comments]. *Nature*. 1998;393:478-80.

29. Toes RE, Schoenberger SP, van der Voort EI, Offringa R, Melief CJ. CD40-CD40Ligand interactions and their role in cytotoxic T lymphocyte priming and anti-tumor immunity. *Semin Immunol.* 1998;10(6):443–8.
30. Ridge JP, Di Rosa F, Matzinger P. A conditioned dendritic cell can be a temporal bridge between a CD4+ T-helper and a T-killer cell. *Nature.* 1998;393(6684):474–8.
31. Olson MR, Seah SG, Edenborough K, Doherty PC, Lew AM, Turner SJ. CD154+ CD4+ T-cell dependence for effective memory influenza virus-specific CD8+ T-cell responses. *Immunol Cell Biol.* 2014;92(7):605–11.
32. Umeshappa CS, Nanjundappa RH, Xie Y, Freywald A, Xu Q, Xiang J. Differential requirements of CD4(+) T-cell signals for effector cytotoxic T-lymphocyte (CTL) priming and functional memory CTL development at higher CD8(+) T-cell precursor frequency. *Immunology.* 2013;138(4):298–306.
33. Andrade RM, Wessendarp M, Subauste CS. CD154 activates macrophage antimicrobial activity in the absence of IFN-gamma through a TNF-alpha-dependent mechanism. *J Immunol.* 2003;171(12):6750–6.
34. Srivastava N, Sudan R, Saha B. CD40-modulated dual-specificity phosphatases MAPK phosphatase (MKP)-1 and MKP-3 reciprocally regulate *Leishmania major* infection. *J Immunol.* 2011 May 15;186(10):5863–72.
35. Subauste CS, Wessendarp M. CD40 restrains in vivo growth of *Toxoplasma gondii* independently of gamma interferon. *Infect Immun.* 2006;74(3):1573–9.
36. Akiyama T, Shimo Y, Yanai H, Qin J, Ohshima D, Maruyama Y, et al. The tumor necrosis factor family receptors RANK and CD40 cooperatively establish the thymic medullary microenvironment and self-tolerance. *Immunity.* 2008;29(3):423–37.
37. Desanti GE, Cowan JE, Baik S, Parnell SM, White AJ, Penninger JM, et al. Developmentally Regulated Availability of RANKL and CD40 Ligand Reveals Distinct Mechanisms of Fetal and Adult Cross-Talk in the Thymus Medulla. *J Immunol.* 2012;189(12):5519–26.
38. Foy TM, Page DM, Waldschmidt TJ, Schoneveld A, Laman JD, Masters SR, et al. An essential role for gp39, the ligand for CD40, in thymic selection. *J Exp Med.* 1995;182(5):1377–88.
39. Mavroudi I, Papadaki HA. The Role of CD40/CD40 Ligand Interactions in Bone Marrow Granulopoiesis. *Sci World J.* 2011;11:2011–9.
40. Mavroudi I, Papadaki V, Pyrovolaki K, Katonis P, Eliopoulos AG, Papadaki HA. The CD40/CD40 ligand interactions exert pleiotropic effects on bone marrow granulopoiesis. *J Leukoc Biol.* 2011;89(5):771–83.
41. Seijkens T, Engel D, Tjwa M, Lutgens E. The role of CD154 in haematopoietic development. *Thromb Haemost.* 2010;104(4):693–701.

42. Grammer AC, Bergman MC, Miura Y, Fujita K, Davis LS, Lipsky PE. The CD40 ligand expressed by human B cells costimulates B cell responses. *J Immunol.* 1995 May 15;154(10):4996–5010.
43. Grammer AC, McFarland RD, Heaney J, Darnell BF, Lipsky PE. Expression, regulation, and function of B cell-expressed CD154 in germinal centers. *J Immunol.* 1999;163(8):4150–9.
44. Wykes M. Why do B cells produce CD40 ligand? *Immunol Cell Biol.* 2003;81(4):328–31.
45. Henn V, Slupsky JR, Grafe M, Anagnostopoulos I, Forster R, Muller BG, et al. CD40 ligand on activated platelets triggers an inflammatory reaction of endothelial cells. *Nature.* 1998;391:591–4.
46. Gauchat JF, Henchoz S, Fattah D, Mazzei G, Aubry JP, Jomotte T, et al. CD40 ligand is functionally expressed on human eosinophils. *Eur J Immunol.* 1995;25:863–5.
47. Mach F, Schönbeck U, Sukhova GK, Bourcier T, Bonnefoy J-Y, Pober JS, et al. Functional CD40 ligand is expressed on human vascular endothelial cells, smooth muscle cells, and macrophages: implications for CD40–CD40 ligand signaling in atherosclerosis. *Proc Natl Acad Sci.* 1997;94(5):1931–6.
48. Johnson S, Zhan Y, Sutherland RM, Mount AM, Bedoui S, Brady JL, et al. Selected Toll-like receptor ligands and viruses promote helper-independent cytotoxic T cell priming by upregulating CD40L on dendritic cells. *Immunity.* 2009;30(2):218–27.
49. El Fakhry Y, Alturaihi H, Yacoub D, Liu L, Guo W, Leveille C, et al. Functional Interaction of CD154 Protein with $\alpha 5 \beta 1$ Integrin Is Totally Independent from Its Binding to $\alpha \text{IIb} \beta 3$ Integrin and CD40 Molecules. *J Biol Chem.* 2012;287(22):18055–66.
50. Leveille C, Bouillon M, Guo W, Bolduc J, Sharif-Askari E, El-Fakhry Y, et al. CD40 Ligand Binds to $\alpha 5 \beta 1$ Integrin and Triggers Cell Signaling. *J Biol Chem.* 2007;282(8):5143–51.
51. Wolf D, Hohmann J-D, Wiedemann A, Bledzka K, Blankenbach H, Marchini T, et al. Binding of CD40L to Mac-1's I-domain involves the EQLKKSRTL motif and mediates leukocyte recruitment and atherosclerosis--but does not affect immunity and thrombosis in mice. *Circ Res.* 2011 Nov 11;109(11):1269–79.
52. Zirlik A, Maier C, Gerdes N, MacFarlane L, Soosairajah J, Bavendiek U, et al. CD40 ligand mediates inflammation independently of CD40 by interaction with Mac-1. *Circulation.* 2007;115(12):1571–80.
53. Elzey BD, Tian J, Jensen RJ, Swanson AK, Lees JR, Lentz SR, et al. Platelet-mediated modulation of adaptive immunity: a communication link between innate and adaptive immune compartments. *Immunity.* 2003;19(1):9–19.
54. Elzey BD, Ratliff TL, Sowa JM, Crist SA. Platelet CD40L at the interface of adaptive immunity. *Thromb Res.* 2011;127(3):180–3.

55. Elzey BD. Cooperation between platelet-derived CD154 and CD4⁺ T cells for enhanced germinal center formation. *J Leukoc Biol.* 2005;78(1):80–4.
56. André P, Prasad KSS, Denis CV, He M, Papalia JM, Hynes RO, et al. CD40L stabilizes arterial thrombi by a beta3 integrin-dependent mechanism. *Nat Med.* 2002;8(3):247–52.
57. Charafeddine AH, Kim EJ, Maynard DM, Yi H, Weaver TA, Gunay-Aygun M, et al. Platelet-Derived CD154: Ultrastructural Localization and Clinical Correlation in Organ Transplantation. *Am J Transplant.* 2012;12(11):3143–51.
58. Choi W-S, Jeon O-H, Kim D-S. CD40 ligand shedding is regulated by interaction between matrix metalloproteinase-2 and platelet integrin alpha(IIb)beta(3). *J Thromb Haemost.* 2010 Jun;8(6):1364–71.
59. Furman MI, Frelinger AL, Michelson AD. GPIIb/IIIa inhibitor-induced dethrombosis. *J Thromb Thrombolysis.* 2004;18(1):11–7.
60. May AE. Engagement of Glycoprotein IIb/IIIa (alphaIIb beta3) on Platelets Upregulates CD40L and Triggers CD40L-Dependent Matrix Degradation by Endothelial Cells. *Circulation.* 2002 Oct 15;106(16):2111–7.
61. Brenner B, Koppenhoefer U, Grassme H, Kun J, Lang F, Gulbins E. Evidence for a novel function of the CD40 ligand as a signalling molecule in T-lymphocytes. *FEBS Lett.* 1997;417:301–6.
62. Brenner B, Koppenhoefer U, Lepple-Wienhues A, Grassmé H, Müller C, Speer CP, et al. The CD40 ligand directly activates T-lymphocytes via tyrosine phosphorylation dependent PKC activation. *Biochem Biophys Res Commun.* 1997;239(1):11–7.
63. Blotta MH, Marshall JD, DeKruyff RH, Umetsu DT. Cross-linking of the CD40 ligand on human CD4⁺ T lymphocytes generates a costimulatory signal that up-regulates IL-4 synthesis. *J Immunol.* 1996;156:3133–40.
64. Jin R, Yu S, Song Z, Zhu X, Wang C, Yan J, et al. Soluble CD40 ligand stimulates CD40-dependent activation of the β 2 integrin Mac-1 and protein kinase C zeta (PKC ζ) in neutrophils: implications for neutrophil-platelet interactions and neutrophil oxidative burst. *PloS One.* 2013;8(6):e64631.
65. Roy M, Waldschmidt T, Aruffo A, Ledbetter JA, Noelle RJ. The regulation of the expression of gp39, the CD40 ligand, on normal and cloned CD4⁺ T cells. *J Immunol.* 1993;151(5):2497–510.
66. Jaiswal AI, Dubey C, Swain SL, Croft M. Regulation of CD40 ligand expression on naive CD4 T cells: a role for TCR but not co-stimulatory signals. *Int Immunol.* 1996;8(2):275–85.
67. Fayen JD. Multiple cytokines sharing the common receptor γ chain can induce CD154/CD40 ligand expression by human CD4⁺ T lymphocytes via a cyclosporin A-resistant pathway. *Immunology.* 2001;104(3):299–306.

68. Lee BO, Haynes L, Eaton SM, Swain SL, Randall TD. The Biological Outcome of CD40 Signaling Is Dependent on the Duration of CD40 Ligand Expression: Reciprocal Regulation by Interleukin (IL)-4 and IL-12. *J Exp Med.* 2002;196(5):693-704.
69. Nakae S, Asano M, Horai R, Sakaguchi N, Iwakura Y. IL-1 enhances T cell-dependent antibody production through induction of CD40 ligand and OX40 on T cells. *J Immunol.* 2001;167(1):90-7.
70. Peng X, Remacle JE, Kasran A, Huylebroeck D, Ceuppens JL. IL-12 up-regulates CD40 ligand (CD154) expression on human T cells. *J Immunol.* 1998;160:1166-72.
71. Skov S, Bonyhadi M, Odum N, Ledbetter JA. IL-2 and IL-15 Regulate CD154 Expression on Activated CD4 T Cells. *J Immunol.* 2000 Apr 1;164(7):3500-5.
72. Bowen F, Haluskey J, Quill H. Altered CD40 ligand induction in tolerant T lymphocytes. *Eur J Immunol.* 1995;25:2830-4.
73. Li W, Carlson TL, Green WR. Stimulation-dependent induction of CD154 on a subset of CD4+ FoxP3+ T-regulatory cells. *Int Immunopharmacol.* 2011;11(9):1205-10.
74. Casamayor-Palleja M, Khan M, MacLennan IC. A subset of CD4+ memory T cells contains preformed CD40 ligand that is rapidly but transiently expressed on their surface after activation through the T cell receptor complex. *J Exp Med.* 1995;181(4):1293-301.
75. Koguchi Y, Buenafe AC, Thauland TJ, Gardell JL, Bivins-Smith ER, Jacoby DB, et al. Preformed CD40L Is Stored in Th1, Th2, Th17, and T Follicular Helper Cells as Well as CD4+8- Thymocytes and Invariant NKT Cells but Not in Treg Cells. Mosley RL, editor. *PLoS ONE.* 2012;7(2):e31296.
76. Whitmire JK, Eam B, Whitton JL. Tentative T Cells: Memory Cells Are Quick to Respond, but Slow to Divide. Koup RA, editor. *PLoS Pathog.* 2008;4(4):e1000041.
77. Fuleihan R, Ramesh N, Horner A, Ahern D, Belshaw PJ, Alberg DG, et al. Cyclosporin A inhibits CD40 ligand expression in T lymphocytes. *J Clin Invest.* 1994;93:1315-20.
78. Lobo FM, Zanjani R, Ho N, Chatila TA, Fuleihan RL. Calcium-dependent activation of TNF family gene expression by Ca²⁺/calmodulin kinase type IV/Gr and calcineurin. *J Immunol.* 1999;162(4):2057-63.
79. Nüsslein HG, Frosch KH, Woith W, Lane P, Kalden JR, Manger B. Increase of intracellular calcium is the essential signal for the expression of CD40 ligand. *Eur J Immunol.* 1996;26(4):846-50.
80. Tsitsikov EN. The CD40L Promoter Contains Nuclear Factor of Activated T Cells-binding Motifs Which Require AP-1 Binding for Activation of Transcription. *J Biol Chem.* 1996 Feb 16;271(7):3763-70.
81. Suárez A, Mozo L, Gayo A, Zamorano J, Gutierrez C. Requirement of a second signal via protein kinase C or protein kinase A for maximal expression of CD40 ligand.

- Involvement of transcriptional and posttranscriptional mechanisms. *Eur J Immunol.* 1997;27(11):2822-9.
82. Cron RQ, Bandyopadhyay R, Genin A, Brunner M, Kersh GJ, Yin J, et al. Early growth response-1 is required for CD154 transcription. *J Immunol.* 2006;176(2):811-8.
 83. Siddiqa A, Sims-Mourtada JC, Guzman-Rojas L, Rangel R, Guret C, Madrid-Marina V, et al. Regulation of CD40 and CD40 ligand by the AT-hook transcription factor AKNA. *Nature.* 2001;410(6826):383-7.
 84. Schubert LA, Cron RQ, Cleary AM, Brunner M, Song A, Lu L-S, et al. A T Cell-specific Enhancer of the Human CD40 Ligand Gene. *J Biol Chem.* 2002;277(9):7386-95.
 85. Srahna M, Remacle JE, Annamalai K, Pype S, Huylebroeck D, Boogaerts MA, et al. NF- κ B is involved in the regulation of CD154 (CD40 ligand) expression in primary human T cells. *Clin Exp Immunol.* 2001;125(2):229-36.
 86. Shim J, Karin M. The control of mRNA stability in response to extracellular stimuli. *Mol Cells.* 2002;14(3):323-31.
 87. Khabar KSA. Rapid transit in the immune cells: the role of mRNA turnover regulation. *J Leukoc Biol.* 2007;81(6):1335-44.
 88. Lu J-Y, Sadri N, Schneider RJ. Endotoxic shock in AUF1 knockout mice mediated by failure to degrade proinflammatory cytokine mRNAs. *Genes Dev.* 2006;20(22):3174-84.
 89. Taylor GA, Carballo E, Lee DM, Lai WS, Thompson MJ, Patel DD, et al. A pathogenetic role for TNF alpha in the syndrome of cachexia, arthritis, and autoimmunity resulting from tristetraproline (TTP) deficiency. *Immunity.* 1996;4(5):445-54.
 90. Ford GS, Barnhart B, Shone S, Covey LR. Regulation of CD154 (CD40 ligand) mRNA stability during T cell activation. *J Immunol.* 1999;162:4037-44.
 91. Murakami K, Ma W, Fuleihan R, Pober JS. Human endothelial cells augment early CD40 ligand expression in activated CD4+ T cells through LFA-3-mediated stabilization of mRNA. *J Immunol.* 1999;163(5):2667-73.
 92. Vavassori S, Shi Y, Chen C-C, Ron Y, Covey LR. In vivo post-transcriptional regulation of CD154 in mouse CD4+ T cells. *Eur J Immunol.* 2009;39(8):2224-32.
 93. Hamilton BJ, Genin A, Cron RQ, Rigby WFC. Delineation of a Novel Pathway That Regulates CD154 (CD40 Ligand) Expression. *Mol Cell Biol.* 2003;23(2):510-25.
 94. Hamilton BJ, Wang X-W, Collins J, Bloch D, Bergeron A, Henry B, et al. Separate cis-trans Pathways Post-transcriptionally Regulate Murine CD154 (CD40 Ligand) Expression: A NOVEL FUNCTION FOR CA REPEATS IN THE 3'-UNTRANSLATED REGION. *J Biol Chem.* 2008;283(37):25606-16.

95. Nicodemos RAM. Modulation of CD40 Ligand (CD40L) Expression by Polypyrimidine Tract-binding Protein [Tesis doctoral]. University of Medicine and Dentistry of New Jersey; 2010.
96. Matus-Nicodemos R, Vavassori S, Castro-Faix M, Valentin-Acevedo A, Singh K, Marcelli V, et al. Polypyrimidine Tract-Binding Protein Is Critical for the Turnover and Subcellular Distribution of CD40 Ligand mRNA in CD4+ T Cells. *J Immunol.* 2011 Feb 15;186(4):2164–71.
97. Barnhart B, Kosinski PA, Wang Z, Ford GS, Kiledjian M, Covey LR. Identification of a Complex that Binds to the CD154 3' Untranslated Region: Implications for a Role in Message Stability During T Cell Activation. *J Immunol.* 2000;165(8):4478–86.
98. Martin-Donaire T, Losada-Fernandez I, Perez-Chacon G, Rua-Figueroa I, Erausquin C, Naranjo-Hernandez A, et al. Association of the microsatellite in the 3' untranslated region of the CD154 gene with rheumatoid arthritis in females from a Spanish cohort: a case-control study. *Arthritis Res Ther.* 2007;9(5):R89.
99. Raghavan A, Ogilvie RL, Reilly C, Abelson ML, Raghavan S, Vasdewani J, et al. Genome-wide analysis of mRNA decay in resting and activated primary human T lymphocytes. *Nucleic Acids Res.* 2002;30:5529–38.
100. Fan XC, Myer VE, Steitz JA. AU-rich elements target small nuclear RNAs as well as mRNAs for rapid degradation. *Genes Dev.* 1997;11(19):2557–68.
101. Halees AS, El-Badrawi R, Khabar KSA. ARED Organism: expansion of ARED reveals AU-rich element cluster variations between human and mouse. *Nucleic Acids Res.* 2007;36(Database):D137–40.
102. Xu N, Chen CY, Shyu A-B. Modulation of the fate of cytoplasmic mRNA by AU-rich elements: key sequence features controlling mRNA deadenylation and decay. *Mol Cell Biol.* 1997;17(8):4611–21.
103. Rigby WF, Waugh MG, Hamilton BJ. Characterization of RNA binding proteins associated with CD40 ligand (CD154) mRNA turnover in human T lymphocytes. *J Immunol.* 1999;163:4199–206.
104. Laughlin J, Oghlidos S, Porter JF, Matus-Nicodemos R, Siquett FL, Marcelli V, et al. Functional analysis of a tripartite stability element within the CD40 ligand 3' untranslated region. *Immunology.* 2008 Jul;124(3):368–79.
105. Singh K, Laughlin J, Kosinski PA, Covey LR. Nucleolin Is a Second Component of the CD154 mRNA Stability Complex That Regulates mRNA Turnover in Activated T Cells. *J Immunol.* 2004;173(2):976–85.
106. Peters AL, Stunz LL, Bishop GA. CD40 and autoimmunity: the dark side of a great activator. *Semin Immunol.* 2009 Oct;21(5):293–300.
107. Toubi E, Shoenfeld Y. The Role of CD40–CD154 Interactions in Autoimmunity and the Benefit of Disrupting this Pathway. *Autoimmunity.* 2004;37(6-7):457–64.

108. Desai MA, Lu L, Ramsey GR, Datta SK. Hyperexpression of CD40 ligand by B and T cells in human lupus and its role in pathogenic autoantibody production. *J Clin Invest.* 1996;97:2063-73.
109. Katsiari CG, Liossis S-N, Dimopoulos AM, Charalambopoulos DV, Mavrikakis M, Sfrikakis PP. CD40L overexpression on T cells and monocytes from patients with systemic lupus erythematosus is resistant to calcineurin inhibition. *Lupus.* 2002;11(6):370-8.
110. Koshy M, Berger D, Crow MK. Increased expression of CD40 ligand on systemic lupus erythematosus lymphocytes. *J Clin Invest.* 1996;98:826-37.
111. Vakkalanka RK, Woo C, Kirou KA, Koshy M, Berger D, Crow MK. Elevated levels and functional capacity of soluble CD40 ligand in systemic lupus erythematosus sera. *Arthritis Reum.* 1999;42:871-81.
112. Gerritse K, Laman JD, Noelle RJ, Aruffo A, Ledbetter JA, Boersma WJ, et al. CD40-CD40 ligand interactions in experimental allergic encephalomyelitis and multiple sclerosis. *Proc Natl Acad Sci U S A.* 1996;93:2499-504.
113. Danese S, Katz JA, Saibeni S, Papa A, Gasbarrini A, Vecchi M, et al. Activated platelets are the source of elevated levels of soluble CD40 ligand in the circulation of inflammatory bowel disease patients. *Gut.* 2003;52(10):1435-41.
114. Liu MF, Chao SC, Wang CR, Lei HY. Expression of CD40 and CD40 ligand among cell populations within rheumatoid synovial compartment. *Autoimmunity.* 2001;34(2):107-13.
115. MacDonald KP, Nishioka Y, Lipsky PE, Thomas R. Functional CD40 ligand is expressed by T cells in rheumatoid arthritis. *J Clin Invest.* 1997;100(9):2404.
116. Berner B, Wolf G, Hummel KM, Müller GA, Reuss-Borst MA. Increased expression of CD40 ligand (CD154) on CD4+ T cells as a marker of disease activity in rheumatoid arthritis. *Ann Rheum Dis.* 2000;59(3):190-5.
117. Goules A, Tzioufas AG, Manousakis MN, Kirou KA, Crow MK, Routsias JG. Elevated levels of soluble CD40 ligand (sCD40L) in serum of patients with systemic autoimmune diseases. *J Autoimmun.* 2006;26(3):165-71.
118. Herve M, Isnardi I, Ng Y -s., Bussel JB, Ochs HD, Cunningham-Rundles C, et al. CD40 ligand and MHC class II expression are essential for human peripheral B cell tolerance. *J Exp Med.* 2007;204(7):1583-93.
119. Clegg CH, Rulfes JT, Haugen HS, Hoggatt IH, Aruffo A, Durham SK, et al. Thymus dysfunction and chronic inflammatory disease in gp39 transgenic mice. *Int Immunol.* 1997;9(8):1111-22.
120. Kawamura T, Kanai T, Dohi T, Uraushihara K, Totsuka T, Iiyama R, et al. Ectopic CD40 Ligand Expression on B Cells Triggers Intestinal Inflammation. *J Immunol.* 2004;172(10):6388-97.

121. Quezada SA, Eckert M, Adeyi OA, Schned AR, Noelle RJ, Burns CM. Distinct mechanisms of action of anti-CD154 in early versus late treatment of murine lupus nephritis. *Arthritis Rheum.* 2003;48(9):2541-54.
122. Wang X, Huang W, Schiffer LE, Mihara M, Akkerman A, Hiromatsu K, et al. Effects of anti-CD154 treatment on B cells in murine systemic lupus erythematosus. *Arthritis Rheum.* 2003 Feb;48(2):495-506.
123. Stuber E, Strober W, Neurath M. Blocking the CD40L-CD40 interaction in vivo specifically prevents the priming of T helper 1 cells through the inhibition of interleukin 12 secretion. *J Exp Med.* 1996;183(2):693-8.
124. Kalunian KC, Davis JC, Merrill JT, Totoritis MC, Wofsy D, for the IDEC-131 Lupus Study Group. Treatment of systemic lupus erythematosus by inhibition of T cell costimulation with anti-CD154: A randomized, double-blind, placebo-controlled trial. *Arthritis Rheum.* 2002;46(12):3251-8.
125. Kawai T, Andrews D, Colvin RB, Sachs DH, Cosimi AB. Thromboembolic complications after treatment with monoclonal antibody against CD40 ligand. *Nat Med.* 2000;6(2):114.
126. Xie JH, Yamniuk AP, Borowski V, Kuhn R, Susulic V, Rex-Rabe S, et al. Engineering of a novel anti-CD40L domain antibody for treatment of autoimmune diseases. *J Immunol.* 2014;192(9):4083-92.
127. Goris A, Liston A. The Immunogenetic Architecture of Autoimmune Disease. *Cold Spring Harb Perspect Biol.* 2012;4(3):a007260-a007260.
128. Citores MJ, Pérez-Aciego P, Rodríguez-Gallego C, Contreras-Martín B, García-Laorden I, Durántez A. CD154 polymorphism in Spanish populations. Differences in the allelic distribution between Canary islanders and Peninsulars. *Eur J Immunogenet.* 2000 Jun;27(3):141-4.
129. Gomolka M, Menninger H, Saal JE, Lemmel EM, Albert ED, Niwa O, et al. Immunoprinting: various genes are associated with increased risk to develop rheumatoid arthritis in different groups of adult patients. *J Mol Med.* 1995;73:19-29.
130. John S, Myerscough A, Marlow A, Hajeer A, Silman A, Ollier W, et al. Linkage of cytokine genes to rheumatoid arthritis. Evidence of genetic heterogeneity. *Ann Rheum Dis.* 1998;57:361-5.
131. Dai Y, Masterman T, Huang W, Hillert J. Analysis of a CD40 ligand dinucleotide microsatellite in multiple sclerosis. *Eur J Immunogenet.* 2002;29:81-5.
132. Citores MJ, Rúa-Figueroa I, Rodríguez-Gallego C, Durántez A, García-Laorden MI, Rodríguez-Lozano C, et al. The dinucleotide repeat polymorphism in the 3'UTR of the CD154 gene has a functional role on protein expression and is associated with systemic lupus erythematosus. *Ann Rheum Dis.* 2004;63(3):310-7.
133. Chakrabarty S, Snyder JT, Shen J, Azmi H, Hu PQ, Chen Q, et al. Human CD14hi monocytes and myeloid dendritic cells provide a cell contact-dependent costimulatory signal for early CD40 ligand expression. *Blood.* 2011;117(5):1585-94.

134. Tan EM, Cohen AS, Fries JF, Masi AT, McShane DJ, Rothfield NF, et al. The 1982 revised criteria for the classification of systemic lupus erythematosus. *Arthritis Reum.* 1982;25:1271-7.
135. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Reum.* 1988;31:315-24.
136. Blin N, Stafford DW. A general method for isolation of high molecular weight DNA from eukaryotes. *NucleicAcidsRes.* 1976;3:2303-8.
137. Paz I, Kostı I, Ares M, Cline M, Mandel-Gutfreund Y. RBPmap: a web server for mapping binding sites of RNA-binding proteins. *Nucleic Acids Res.* 2014;42(W1):W361-7.
138. Cook KB, Kazan H, Zuberi K, Morris Q, Hughes TR. RBPDB: a database of RNA-binding specificities. *Nucleic Acids Res.* 2011 Jan 1;39(Database):D301-8.
139. Rice P, Longden I, Bleasby A. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet.* 2000;16(6):276-7.
140. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 2003;31(13):3406-15.
141. Faul F, Erdfelder E, Lang A-G, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods.* 2007;39(2):175-91.
142. Moser BK, Halabi S. Estimation and testing of the relative risk of disease in case-control studies with a set of k matched controls per case with known prevalence of disease. *Stat Med.* 2012 Jan 13;31(1):29-44.
143. Schmidt CO, Kohlmann T. When to use the odds ratio or the relative risk? *Int J Public Health.* 2008;53(3):165-7.
144. López P, Mozo L, Gutiérrez C, Suárez A. Epidemiology of systemic lupus erythematosus in a northern Spanish population: gender and age influence on immunological features. *Lupus.* 2003;12(11):860-5.
145. Clarke GM, Anderson CA, Pettersson FH, Cardon LR, Morris AP, Zondervan KT. Basic statistical analysis in genetic case-control studies. *Nat Protoc.* 2011 Feb;6(2):121-33.
146. Rosado S, Rúa-Figueroa I, Vargas JA, Garcia-Laorden MI, Losada-Fernandez I, Martin-Donaire T, et al. Interleukin-10 promoter polymorphisms in patients with systemic lupus erythematosus from the Canary Islands. *Int J Immunogenet.* 2008;35:235-42.
147. Flores C, Maca-Meyer N, Perez JA, Gonzalez AM, Larruga JM, Cabrera VM. A predominant European ancestry of paternal lineages from Canary Islanders. *AnnHumGenet.* 2003;67:138-52.

148. Pino-Yanes M, Corrales A, Basaldúa S, Hernández A, Guerra L, Villar J, et al. North African Influences and Potential Bias in Case-Control Association Studies in the Spanish Population. O'Rourke D, editor. PLoS ONE. 2011;6(3):e18389.
149. Hauge XY, Litt M. A study of the origin of "shadow bands" seen when typing dinucleotide repeat polymorphisms by the PCR. Hum Mol Genet. 1993;2(4):411-5.
150. Hite JM, Eckert KA, Cheng KC. Factors affecting fidelity of DNA synthesis during PCR amplification of d(C-A)_n.d(G-T)_n microsatellite repeats. NucleicAcidsRes. 1996;24:2429-34.
151. Zuker M, Jacobson AB. "Well-determined" regions in RNA secondary structure prediction: analysis of small subunit ribosomal RNA. Nucleic Acids Res. 1995;23(14):2791-8.
152. Altmüller J, Palmer LJ, Fischer G, Scherb H, Wjst M. Genomewide scans of complex human diseases: true linkage is hard to find. Am J Hum Genet. 2001 Nov;69(5):936-50.
153. Hirschhorn JN, Lohmueller K, Byrne E, Hirschhorn K. A comprehensive review of genetic association studies. Genet Med. 2002;4(2):45-61.
154. Cui J, Stahl EA, Saevarsdottir S, Miceli C, Diogo D, Trynka G, et al. Genome-Wide Association Study and Gene Expression Analysis Identifies CD84 as a Predictor of Response to Etanercept Therapy in Rheumatoid Arthritis. PLoS Genet. 2013 Mar;9(3):e1003394.
155. Chadha S, Miller K, Farwell L, Lightstone LB, Daly MJ, Rioux JD, et al. Haplotype structure of TNFRSF5-TNFSF5 (CD40-CD40L) and association analysis in systemic lupus erythematosus. Eur J Hum Genet. 2005;13:669-76.
156. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. Nature. 2009 Oct 8;461(7265):747-53.
157. So H-C, Yip BHK, Sham PC. Estimating the Total Number of Susceptibility Variants Underlying Complex Diseases from Genome-Wide Association Studies. Weedon MN, editor. PLoS ONE. 2010 Nov 17;5(11):e13898.
158. Hannan AJ. Tandem repeat polymorphisms: modulators of disease susceptibility and candidates for "missing heritability." Trends Genet. 2010 Feb;26(2):59-65.
159. Guo X, Meng Y, Yu N, Pan Y. Cloud computing for detecting high-order genome-wide epistatic interaction via dynamic clustering. BMC Bioinformatics. 2014;15(1):102.
160. Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic Acids Res. 2014 Jan 1;42(D1):D1001-6.

161. Schrodin SJ, Mukherjee S, Shan Y, Tromp G, Sninsky JJ, Callear AP, et al. Genetic-based prediction of disease traits: prediction is very difficult, especially about the future. *Front Genet.* 2014;5(162):1-18.
162. De Sousa Abreu R, Penalva LO, Marcotte EM, Vogel C. Global signatures of protein and mRNA expression levels. *Mol Biosyst.* 2009;5(12):1512-26.
163. Matus-Nicodemus R, Vavassori S, Castro-Faix M, Valentin-Acevedo A, Singh K, Marcelli V, et al. Polypyrimidine Tract-Binding Protein Is Critical for the Turnover and Subcellular Distribution of CD40 Ligand mRNA in CD4+ T Cells. *J Immunol.* 2011 Feb 15;186(4):2164-71.
164. Fuleihan R, Ramesh N, Horner A, Ahern D, Belshaw PJ, Alberg DG, et al. Cyclosporin A inhibits CD40 ligand expression in T lymphocytes. *J Clin Invest.* 1994;93(3):1315.
165. Wang H, Wang Q, Pape UJ, Shen B, Huang J, Wu B, et al. Systematic investigation of global coordination among mRNA and protein in cellular society. *BMC Genomics.* 2010;11(1):364.
166. Li Y, Masaki T, Shimakami T, Lemon SM. hnRNP L and NF90 interact with hepatitis C virus 5'-terminal untranslated RNA and promote efficient replication. *J Virol.* 2014 Jul;88(13):7199-209.
167. Hung L-H, Heiner M, Hui J, Schreiner S, Benes V, Bindereif A. Diverse roles of hnRNP L in mammalian mRNA processing: A combined microarray and RNAi analysis. *RNA.* 2007 Dec 14;14(2):284-96.
168. Shankarling G, Cole BS, Mallory MJ, Lynch KW. Transcriptome-Wide RNA Interaction Profiling Reveals Physical and Functional Targets of hnRNP L in Human T Cells. *Mol Cell Biol.* 2014 Jan 1;34(1):71-83.
169. Hui J. Novel functional role of CA repeats and hnRNP L in RNA stability. *RNA.* 2003 Aug 1;9(8):931-6.
170. Lee J-H, Jeon M-H, Seo Y-J, Lee Y-J, Ko JH, Tsujimoto Y, et al. CA Repeats in the 3'-Untranslated Region of bcl-2 mRNA Mediate Constitutive Decay of bcl-2 mRNA. *J Biol Chem.* 2004 Oct 8;279(41):42758-64.
171. Shih S-C, Claffey KP. Regulation of Human Vascular Endothelial Growth Factor mRNA Stability in Hypoxia by Heterogeneous Nuclear Ribonucleoprotein L. *J Biol Chem.* 1999 Jan 15;274(3):1359-65.
172. Hahm B, Cho OH, Kim J-E, Kim YK, Kim JH, Oh YL, et al. Polypyrimidine tract-binding protein interacts with HnRNP L. *FEBS Lett.* 1998;425(3):401-6.
173. Cheli Y, Kunicki TJ. hnRNP L regulates differences in expression of mouse integrin alpha2beta1. *Blood.* 2006 Jun 1;107(11):4391-8.
174. Tong A, Nguyen J, Lynch KW. Differential expression of CD45 isoforms is controlled by the combined activity of basal and inducible splicing-regulatory elements in each of the variable exons. *J Biol Chem.* 2005 Nov 18;280(46):38297-304.

