

UNIVERSIDAD AUTÓNOMA DE MADRID

FACULTAD DE CIENCIAS

DEPARTAMENTO DE BIOLOGÍA MOLECULAR

**SPLICING EN EL MHC DE CLASE III:
CARACTERIZACIÓN Y EXPRESIÓN DE LAS ISOFORMAS DEL
GEN *NFκB1*. ESTUDIO DE SU RELACIÓN CON ARTRITIS
REUMATOIDE**

TESIS DOCTORAL

OLATZ VILLATE BEJARANO

MADRID 2010

Memoria presentada por Dña. Olatz Villate Bejarano para optar al grado de
Doctor en Ciencias por la Universidad Autónoma de Madrid

El trabajo presentado en esta memoria ha sido realizado en el Centro de Biología Molecular Severo Ochoa del Consejo Superior de Investigaciones Científicas y Universidad Autónoma de Madrid bajo la dirección de la Dra. Begoña Aguado, gracias a la beca predoctoral concedida por la Comunidad de Madrid (Referencia: 01/FPI/0811/2005).

SUMMARY

A major discovery in the post-genomic era is that the number of proteins in a human proteome is by no means equivalent to the number of genes in the genome. Alternative splicing (AS) is one of the mechanisms that increases transcriptome and proteome diversification by generating multiple mRNA products from a single gene.

The human major histocompatibility complex (MHC) is located on the short arm of chromosome 6 and is divided into the class I, extended class I, class II, extended class II and class III regions. The MHC contributes to a large number of immune-related disorders including insulin dependent diabetes mellitus, rheumatoid arthritis (RA) and common variable immunodeficiency. The human NF- κ B inhibitor-like protein 1 (*NF κ BIL1*) gene, also known as I κ BL, is located at the telomeric end of the Major Histocompatibility Complex (MHC) class III region on chromosome 6. It encodes a protein resembling members of the I κ B protein family with ankyrin repeat sequences that may interact with members of the nuclear factor κ B (NF- κ B)/Rel family. NF- κ B controls many genes involved in inflammation so it is not surprising that NF- κ B is found to be chronically active in many inflammatory diseases.

Two mRNA variants of *NF κ BIL1* were previously described and we have identified two new isoforms in humans. We have characterised *NF κ BIL1* splicing isoforms in different species as well. In this study, we analysed *NF κ BIL1* isoforms expression levels in different human cell lines and tissues and we investigated the possible role of this gene in the development of RA by studying expression levels of each isoform in RA patients and controls. For precise analyses of *NF κ BIL1* expression, mRNA levels of the four isoforms identified were measured by real-time RT-PCR analysis. *NF κ BIL1* mRNA isoforms present different expression levels in all samples studied, being the canonical isoform *NF κ BIL1* the most expressed in all cases. To describe the distribution of all *NF κ BIL1* protein isoforms, localisation experiments with confocal microscopy were performed. Transient expression of NF κ BIL1 proteins showed that all of them were targeted to the nucleus, specifically in nuclear speckles, suggesting a possible role in the regulation of splicing. We also quantified *NF κ BIL1* mRNA isoforms in blood samples of healthy controls and RA patients by real-time RT-PCR analysis. Different expression levels of *NF κ BIL1* isoforms were also found between RA patients and controls, showing a clear over-expression of each isoform in RA patients. These data suggest that *NF κ BIL1* variants may be related with the regulation of splicing and the overexpression of each isoform in RA patients lend support to a role of this gene in the pathogenesis of RA.

We questioned whether the level of DNA methylation in *NF κ BIL1* promoter region may be relevant to *NF κ BIL1* expression and may be a possible contributory factor in the inflammatory

pathogenesis of RA. A detailed methylation analysis of the CpG sites in samples of controls and RA patients was performed but our results showed that the methylation pattern did not change between cases and controls, suggesting that mechanisms other than methylation are involved in regulation of *NFκB1* expression.

ABREVIATURAS

Ab	Anticuerpo	LPS	Lipopolisacárido
ADN	Ácido desoxirribonucleico	M	Molar
ADNc	Ácido desoxirribonucleico complementario	Mb	Megabase
ARN	Ácido ribonucleico	MHC	Complejo Mayor de Histocompatibilidad
ARNr	Ácido ribonucleico ribosómico	NES	Señales de exportación nuclear
ARNt	Ácido ribonucleico transferente	NFκB	Factor nuclear κB
AR	Artritis reumatoide	NLS	Señales de localización nuclear
BSA	Albúmina de suero bovino	ORF	Fase abierta de lectura
°C	Grado centígrado	pb	Pares de bases
CNV	Variaciones del número de copias	PBMC	Célula mononuclear de sangre periférica
Ct	Ciclo umbral	PBS	Tampón fosfato salino
DIP	Polimorfismo de delección/inserción	PCR	Reacción en cadena de la polimerasa
DMSO	Dimetilsulfóxido	RFU	Unidades relativas de fluorescencia
EDTA	Ácido etilendinitrilotetracético	RIN	Número de integridad del ARN
ESE	Activador del splicing de exones	rpm	Revoluciones por minuto
ESS	Represores del splicing de exones	RT	Transcripción reversa
EST	Marcadores de secuencias expresadas	SDS	Dodecil sulfato sódico
FBS	Suero fetal bovino	snRNA	ARN pequeño nuclear
GAPDH	Deshidrogenasa del gliceraldehído 3 fosfato	snoRNA	ARN pequeño nucleolar
HEP	Proyecto del Epigenoma Humano	snRNP	Ribonucleoproteína pequeña
HLA	Antígeno Humano Leucocitario	SNP	Polimorfismo de un único nucleótido
IκB	Inhibidores del factor κB	STR	Repeticiones cortas en tándem
INDEL	Inserciones/Delecciones	TAD	Dominio de activación de la transcripción
IPTG	Isopropil β-D-tiogalacopiranosido	TNF	Factor de Necrosis Tumoral
ISE	Activadores del splicing de intrones	UTR	Región no traducida
ISS	Represores del splicing de intrones	WB	Western Blot
kDa	Kilodaltons	X-Gal	5-bromo-4-cloro-3-indolil-β-D-galactósido

ÍNDICE

I. INTRODUCCIÓN.....	1
1. LA COMPLEJIDAD DE LOS GENOMAS.....	1
2. EL PROCESO DE SPLICING.....	6
2.1 El código de splicing.....	6
2.1.1 Elementos <i>cis</i>	6
2.1.2 Elementos <i>trans</i>	8
2.2 Splicing alternativo.....	10
2.2.1 Tipos de splicing alternativo.....	11
2.3 Splicing y enfermedades.....	12
2.4 Splicing y desarrollo de nuevas terapias.....	14
3. EL COMPLEJO MAYOR DE HISTOCOMPATIBILIDAD (MHC) HUMANO.....	15
3.1 Clases del MHC.....	16
3.2 Polimorfismos en el MHC.....	18
3.3 El MHC y su función en el sistema inmune.....	19
3.4 MHC y enfermedades.....	20
3.5 Perfil de metilación del ADN en el MHC.....	21
3.6 Splicing en el MHC.....	23
4. EL GEN <i>NFκBIL1</i>	23
4.1 El factor nuclear κB y la familia de inhibidores IκB.....	24
4.1.1 La familia NF-κB/Rel.....	24
4.1.2 La familia IκB.....	27
4.1.3 Sistema NF-κB/ IκB.....	30
4.1.4 Dominios presentes en la proteína NFκBIL1.....	31
4.2 <i>NFκBIL1</i> y splicing.....	32
4.3 Asociación de <i>NFκBIL1</i> con enfermedades.....	33
II. OBJETIVOS.....	35
III. MATERIALES Y MÉTODOS.....	36
1. ANÁLISIS INFORMÁTICOS.....	36
2. CULTIVOS CELULARES.....	37
3. MUESTRAS DE SANGRE DE PACIENTES Y CONTROLES.....	37
4. EXTRACCIÓN, PRECIPITACIÓN Y DETERMINACIÓN DE LA CALIDAD DEL ARN.....	38
5. MUESTRAS DE DIFERENTES ESPECIES.....	39
6. 5´RLM-RACE.....	40
7. SÍNTESIS DE ADNc Y ANÁLISIS MEDIANTE PCR.....	41

8. RT-PCR EN TIEMPO REAL.....	44
8.1 Curvas patrón.....	48
8.2 Ensayos de PCR a tiempo real.....	49
8.3 Análisis de datos.....	49
9. ANÁLISIS DE LA REGIÓN PROMOTORA DE <i>NFκBIL1</i>.....	50
9.1 Extracción de ADN genómico.....	50
9.2 Análisis de mutaciones y SNPs.....	50
9.3 Análisis de metilación en la región promotora.....	50
10. SOBRE-EXPRESIÓN DE LAS ISOFORMAS DE <i>NFκBIL1</i>.....	52
10.1 Construcciones para la expresión de cada isoforma.....	52
10.1.1 Epítos en el extremo carboxilo terminal.....	52
10.1.2 Epítos en el extremo amino y carboxilo terminal.....	53
10.2 Transfecciones y análisis por Western Blot.....	55
10.3 Inmunofluorescencias.....	56
10.4 Microscopía confocal y análisis de imágenes.....	57
IV. RESULTADOS.....	59
1. CARACTERIZACIÓN DE LAS ISOFORMAS DE <i>NFκBIL1</i>.....	59
1.1 En <i>Homo Sapiens</i>.....	59
1.2 En otros mamíferos.....	64
1.2.1 <i>NFBκIL1</i> en <i>Mus musculus</i> y <i>Rattus norvegicus</i> (Orden Rodentia).....	64
1.2.2 <i>NFBκIL1</i> en <i>Sus scrofa</i> y <i>Bos Taurus</i> (Orden Artiodactyla).....	66
1.2.3 <i>NFκBIL1</i> en otros mamíferos del Orden Primates.....	69
1.2.3.1 <i>NFκBIL1</i> en distintos tejidos de <i>Macaca mulata</i>	69
1.2.3.2 <i>NFκBIL1</i> en muestras de sangre de otros primates.....	70
2. ANÁLISIS DE LA EXPRESIÓN DE LAS ISOFORMAS DE <i>NFκBIL1</i> EN LÍNEAS CELULARES Y TEJIDOS HUMANOS.....	74
3. EXPRESIÓN DE LAS PROTEÍNAS <i>NFκBIL1</i>.....	79
3.1 Clonación, transfección y análisis por Western Blot.....	79
3.1.1 Análisis de las construcciones con epítos en el extremo carboxilo terminal.....	79
3.1.2 Análisis de las construcciones con epítos en los extremos amino y carboxilo terminal.....	81
3.2 Inmunofluorescencias.....	82
3.3 Co-localizaciones.....	85
4. ANÁLISIS DE LA EXPRESIÓN DE LAS ISOFORMAS DE <i>NFκBIL1</i> EN MUESTRAS DE PACIENTES CON ENFERMEDADES AUTOINMUNES Y CONTROLES.....	87
4.1 Estudio en pacientes con AR.....	87
4.2 Estudio en pacientes con Lupus Eritematoso Sistémico.....	89

5. ANÁLISIS DE LA REGIÓN PROMOTORA DE <i>NFκBIL1</i>	90
5.1 Análisis de polimorfismos y mutaciones.....	90
5.2 Análisis de metilación.....	91
V. DISCUSIÓN.....	94
1. CARACTERIZACIÓN DE <i>NFκBIL1</i> EN HUMANOS Y OTRAS ESPECIES.....	94
2. EXPRESIÓN Y LOCALIZACIÓN DE LAS ISOFORMAS DE <i>NFκBIL1</i> EN HUMANOS.....	96
3. ANÁLISIS DE LA EXPRESIÓN DE <i>NFκBIL1</i> EN MUESTRAS DE PACIENTES CON AR Y LUPUS.....	100
4. ANÁLISIS DE LA REGIÓN PROMOTORA DE <i>NFκBIL1</i>	101
VI. CONCLUSIONES.....	104
VII. BIBLIOGRAFÍA.....	106
VIII. ANEXOS.....	114
1. Curriculum Vitae.....	114
2. Publicaciones científicas.....	118

1. LA COMPLEJIDAD DE LOS GENOMAS

Nos encontramos en la era post-genómica en la que disponemos de la secuencia completa de nucleótidos del genoma de diversos organismos eucariotas y procariotas. En Junio del año 2000, hace ya una década exactamente, se completó la secuenciación del primer genoma humano y en este período de diez años, se han producido avances espectaculares tanto tecnológicos como intelectuales. A pesar de ello, existe una opinión generalizada entre los científicos sobre la necesidad que existe todavía de utilizar este conocimiento del genoma humano en beneficio de la humanidad, concretamente en el área sanitaria.

Debido a los grandes avances en la tecnología de secuenciación, disponemos de un gran cúmulo de información a la que ahora debemos darle un sentido biológico, tanto funcional como evolutivo. Por ello, hay varias disciplinas clave en este período post-genómico como son la genómica comparativa, que se encarga de comparar los genomas de diversas especies; la genómica individual que estudia variaciones dentro de los genomas de individuos de la misma especie; la genómica funcional encargada de dilucidar la función de los genes; la transcriptómica que analiza la expresión de los genes y la proteómica encargada del estudio de la estructura y función de las proteínas. Con todo ello, se pretende dar un sentido a toda la información de secuencias de nucleótidos de la que actualmente disponemos y que irá aumentando progresivamente a medida que se vayan secuenciando genomas de otros organismos.

Durante años, se pensó que la complejidad de los distintos organismos se debía al tamaño de su genoma y al número de genes que éste contenía basándose en el axioma “un gen, un ARNmensajero, una proteína”. La complejidad del genoma de los organismos sería, por lo tanto, el reflejo del número de proteínas codificadas por sus genes capaces de desarrollar las múltiples funciones requeridas. Sin embargo, el Proyecto Genoma Humano hizo que se cuestionase esta teoría (Lander et al., 2001; Venter et al., 2001) al identificar en *Homo sapiens* un número reducido de genes (~30.000) en su genoma de 3.400 Mb. Este hecho, junto con el conocimiento de diversos proyectos de secuenciación del genoma de otras especies, llevó a la observación de que organismos

con genomas tan pequeños como el nematodo *Caenorhabditis elegans* (100 Mb) o el pez *Fugu rubripes* (393 Mb) poseían un número de genes similar al humano (*C.elegans*, ~20.000; *F.rubripes*, ~30.000) (Figura 1 y Tabla 1). La aparente discrepancia entre el número de genes y la complejidad del organismo indicaba que el número de proteínas funcionales debería de ser mucho mayor que el número de genes. A medida que la tecnología ha ido avanzando, se han ido conociendo diversos mecanismos capaces de dar una explicación al incremento de la complejidad fenotípica.

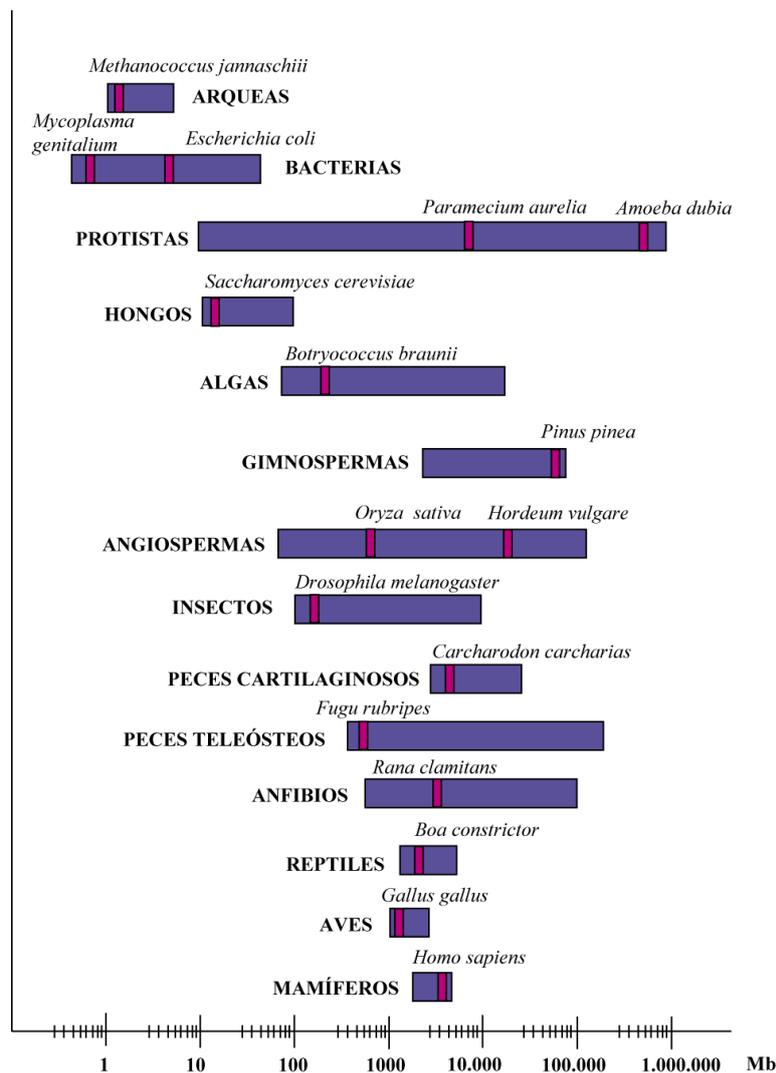


Figura 1. Representación esquemática del rango de tamaños del genoma (en Mb) de distintos organismos. Se observa que no existe una relación directa entre el grado de desarrollo evolutivo y el contenido de ADN del genoma de los organismos.

Nombre científico	Tamaño del genoma (Mb)	Nº de genes	Densidad génica (genes/Mb)	Nº de transcritos
<i>Homo sapiens</i>	3.400	30.000	9	142.000
<i>Mus musculus</i>	3.400	22.900	7	82.000
<i>Bos taurus</i>	3.250	21.000	6,5	30.500
<i>Macaca mulatta</i>	3.000	21.900	7	43.000
<i>Pan troglodytes</i>	2.900	25.000	9	40.000
<i>Rattus norvegicus</i>	2.500	21.000	8	38.100
<i>Fugu rubripes</i>	393	30.000	76	48.600
<i>Drosophila melanogaster</i>	168	15.000	89	22.400
<i>Arabidopsis thaliana</i>	125	27.379	220	32.000
<i>Caenorhabditis elegans</i>	100	20.000	200	42.400
<i>Saccharomyces cerevisiae</i>	12	6.241	520	7.124
<i>Escherichia coli</i>	4.6	4.200	913	4.200
<i>Bacillus subtilis</i>	4.2	4.100	976	4.100

Tabla 1. Comparativa del tamaño del genoma, número de genes y transcritos identificados en las distintas especies. Los datos fueron obtenidos de las bases de datos de NCBI y Ensembl.

El reconocimiento de la regulación del ARN como un punto central en la expresión génica y la generación de complejidad fenotípica comenzó en los años 70 y 80. En esa época, se descubrió a través de análisis de híbridos ARN-ADN que el procesamiento del pre-ARNmensajero de los adenovirus a ARN maduro se realizaba a través de la unión de secuencias expresadas (exones) que estaban separadas por secuencias no codificantes (intrones) en los transcritos primarios (Berget et al., 1977; Chow et al., 1977). Este mecanismo también se reconoció rápidamente como una característica del procesamiento del ARN eucariótico. El descubrimiento del **splicing** (corte y empalme) reveló la complejidad del ARN, mucho mayor que la del ADN. Este potencial se demostró por los hallazgos, en los adenovirus y en las células eucariotas durante la diferenciación celular y en distintos tejidos, de que los productos de ARNm alternativos se podían generar de un sólo precursor de pre-ARNm de una manera regulada. De esta manera, la regulación del **splicing alternativo** permitía a un único gen

producir múltiples ARNs que poseían distintas secuencias reguladoras y codificantes (Licatalosi and Darnell, 2010).

Posteriormente, con el conocimiento de las secuencias completas de genomas de diversos organismos así como del número de genes, empezó a gestarse un nuevo concepto que se basaba en que la complejidad biológica se debía principalmente a la complejidad del ARN. Los papeles conocidos del ARN en la célula se expandieron a lo largo de los años desde el ARN como simple intermediario entre el ADN y las proteínas a su papel regulatorio en el control transcripcional. Se conocieron más mecanismos mediante los cuales los transcritos de ARN se diversificaban desde el punto de la transcripción en adelante contribuyendo a la complejidad transcriptómica como por ejemplo el **procesamiento de ARN (RNA editing)** (Bass, 2002) y las **modificaciones post-transcripcionales como la metilación y la pseudouridilización** (Rottman et al., 1994; Terns and Terns, 2006). Las isoformas maduras de ARN además estaban sujetas a varios niveles de regulación que incluyen la regulación de la traducción mediante factores de regulación, el uso de distintos inicios de traducción, la localización y la estabilidad del ARNm (Filipowicz et al., 2008; Houseley and Tollervey, 2009; Rodriguez et al., 2008).

Otro hecho sorprendente en el estudio de la complejidad de los organismos, fue el descubrimiento de una gran proporción de transcritos que no codificaban proteínas, particularmente en los organismos que se consideraban más complejos (mamíferos y primates) puesto que la proporción de ARN no codificante/ARN codificante incrementaba progresivamente desde las bacterias a los primates (Gustincich et al., 2006). Estos **ARNs no codificantes** que primero se consideraban artefactos, pasaron a describirse como piezas fundamentales en los procesos de regulación de los organismos complejos. En el transcriptoma de mamíferos se encuentran muchos tipos de ARNs no codificantes. Como ejemplos tenemos los ARNs nucleares pequeños (snRNA, *small nuclear RNAs*) de 200 nucleótidos de longitud, que dirigen el complejo del spliceosoma y los ARNs nucleolares pequeños (snoRNAs, *small nucleolar RNAs*) de 60 a 300 nucleótidos de longitud, que conducen a las proteínas asociadas a ejecutar modificaciones específicas de secuencia en las moléculas de ARN transferente y ARN ribosómico (Lindberg and Lundeberg, 2009). Otros ARNs no codificantes presentes en el transcriptoma de mamíferos son los piRNAs (*Piwi-interacting RNAs*), los ARNs no

codificantes largos (lRNAs), o cortos (sRNAs) y los microRNAs. Gracias al descubrimiento de estos ARNs y al estudio de todos los mecanismos citados anteriormente, se demostró, sin duda alguna, que la complejidad del transcriptoma de los mamíferos sobrepasaba de lejos lo que se esperaba sólo unas décadas antes (Lindberg and Lundeberg, 2009). De esta manera, el número de transcritos en organismos superiores superaba al número de genes identificados (Tabla 1).

Mientras que en años previos los avances en la comprensión de la regulación del ARN y la complejidad se realizaban mediante estudios de ARNs de manera individual, el objetivo de las nuevas técnicas como los microarrays y la secuenciación de “nueva generación”, es la caracterización de poblaciones de ARNs enteras en contextos celulares a nivel de nucleótido. Con estas metodologías se ha revelado de qué manera la diversificación del ARNm contribuye a la especificidad celular y los mecanismos mediante los cuales esta diversificación se consigue. Un reto para el futuro consiste en determinar qué isoformas distintas de ARNm contribuyen a la complejidad biológica de los organismos (Licatalosi and Darnell, 2010).

Dentro de los mecanismos que contribuyen al incremento de la complejidad, el splicing alternativo es el proceso más estudiado capaz de conseguir diversidad transcriptómica y, por tanto, proteómica (Barak et al., 2009; Kim et al., 2007). Las primeras estimaciones consideraban que este mecanismo sólo lo presentaban una porción pequeña de genes, como los genes de las inmunoglobulinas (Sharp, 1994). Posteriormente, con los análisis de microarrays y los análisis basados en los alineamientos de los marcadores de secuencias expresadas, ESTs, (*Expressed Sequences Tags*)/ADN complementarios (ADNc) con el ADN genómico, se estableció que dos tercios de los genes humanos contenían al menos una variante de splicing alternativo (Lander et al., 2001). Las tecnologías de secuenciación de “nueva generación” ofrecían el potencial de solucionar esta cuestión y varios estudios recientes han estimado mediante análisis de lecturas cortas de ADNc, que aproximadamente el 95% de los genes humanos presentan splicing alternativo (Pan et al., 2008; Wang et al., 2008).

En estudios comparativos del splicing entre especies basados en ESTs se ha visto que el porcentaje de genes y exones sujetos a splicing alternativo, es mayor en vertebrados que en invertebrados (Kim et al., 2007). Y se ha demostrado, por ejemplo,

que sólo el 23% de los genes de pollo presentan splicing y el 68% de los genes en el caso del ratón (Chacko and Ranganathan, 2009). Estos datos ponen de manifiesto que una fracción significativa de los eventos de splicing es especie-específica, sugiriendo un papel importante de este proceso en el cambio evolutivo (Mendes Soares and Valcarcel, 2006). El hecho de usar ESTs en estos estudios comparativos ha sido cuestionado, puesto que a mayor complejidad del organismo hay mayor número de ESTs disponibles en las bases de datos para analizar el organismo en cuestión, dato a tener en cuenta a la hora de analizar los resultados.

En resumen, la complejidad transcriptómica y proteómica se puede conseguir a través de varios mecanismos, en los que el ARN juega un papel primordial. Dentro de los mecanismos que contribuyen a la complejidad fenotípica, en este estudio nos centraremos en el proceso de splicing alternativo.

2. EL PROCESO DE SPLICING

En eucariotas, la mayoría de los genes codificantes contienen las secuencias codificantes (exones) en segmentos discontinuos, interrumpidos por otras secuencias (intrones) generalmente más largas, y que no formarán parte del ARN maduro. Durante la maduración, los intrones son eliminados del transcrito primario mediante un proceso bioquímico conocido como splicing (Hui, 2009; Modrek and Lee, 2002).

2.1 El código de splicing

En la transcripción del ADN a ARN, el procesamiento de los pre-mensajeros está regulado por la presencia de una serie de elementos, los elementos *cis* y los elementos *trans*, que juntos constituyen el “código de splicing” resumido en la Figura 2 (Faustino and Cooper, 2003; Villate et al., 2008). Los elementos *cis* se pueden definir como pequeñas secuencias necesarias presentes a nivel de ARNm que son identificadas por la maquinaria de splicing (spliceosoma) para producir un correcto proceso de splicing. Los elementos *trans* incluyen las proteínas (spliceosoma) que están implicadas en los procesos requeridos para identificar las uniones intrón/exón y catalizar las

reacciones de corte y empalme que escinden los intrones y unen los exones en el orden correcto.

2.1.1 Elementos *cis*

Un gen humano típico contiene de media 8 exones. Los exones internos tienen una longitud media de 145 nucleótidos, y los intrones una longitud de diez veces mayor que los exones (Faustino and Cooper, 2003; Lander et al., 2001). Los exones se definen con tres pequeñas y degeneradas secuencias de splicing: los sitios de splicing 5' y 3' en las uniones de exón/intrón, y el sitio de ramificación dentro de los intrones (Figura 2). El sitio de splicing 5' presenta una secuencia consenso clásica que en eucariotas superiores presenta más variabilidad: 5'-AGGURAGU-3'. La secuencia del sitio de splicing 3' consta de un tramo de polipirimidinas (Y)_n, seguido de un dinucleótido AG en el extremo 3' del intrón. La secuencia de ramificación se sitúa a una distancia de entre 18-40 nucleótidos del AG del sitio de splicing 3' y es reconocida por complementariedad de bases con la pequeña ribonucleoproteína (snRNP) U2, fenómeno que tiene lugar en etapas precatalíticas de la reacción, generalmente después del reconocimiento del sitio de splicing 5'. La secuencia consenso se encuentra degenerada en eucariotas superiores: YNCURAY, siendo Y: C, U; R: A, G y N: cualquier nucleótido. Los componentes de la maquinaria basal de splicing se unen a los sitios de splicing clásicos y promueven el ensamblaje del complejo de splicing conocido como spliceosoma.

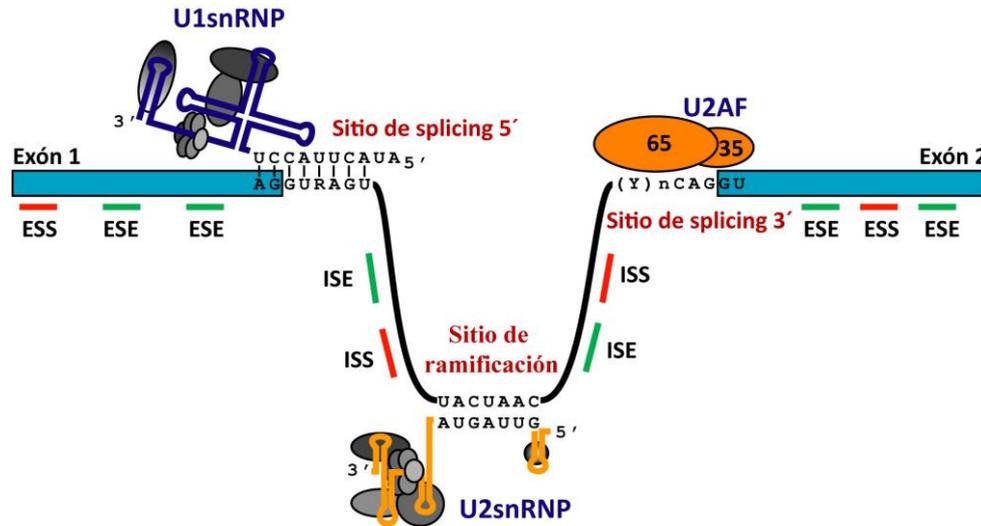


Figura 2. Esquema representativo de los elementos *cis* y *trans* implicados en el proceso de splicing. Podemos ver representados los elementos *cis* básicos, sitios 3' y 5' de splicing y el sitio de ramificación, reconocidos por la maquinaria del spliceosoma (snRNPs y diversas proteínas) y los elementos *cis* regulados (activadores y represores); ESS (*Exonic Splicing Silencers*), ESE (*Exonic Splicing Enhancer*), ISS (*Intronic Splicing Silencers*) e ISE (*Intronic Splicing Enhancer*).

Además de estos sitios de splicing, los exones se definen por otros elementos de regulación que actúan en *cis*, que pueden dividirse en cuatro categorías funcionales: 1) ESEs (*Exonic Splicing Enhancers*), activadores del splicing de exones, 2) ESSs (*Exonic Splicing Silencers*), represores del splicing de exones, 3) ISE (*Intronic Splicing Enhancers*) activadores del splicing de intrones y 4) ISSs (*Intronic Splicing Silencers*), represores del splicing de intrones (Figura 2). A nivel de la secuencia de exones, los ESEs activan el reconocimiento de exones y promueven su inclusión en los transcritos maduros, mientras que los ESSs reprimen su inclusión en los transcritos maduros (Del Gatto-Konczak et al., 1999; Garcia-Blanco et al., 2004). Por otra parte, en la secuencia intrónica, los ISEs activan la inclusión de los exones adyacentes (Del Gatto-Konczak et al., 2000; Garcia-Blanco et al., 2004) mientras que los ISSs inhiben la definición de los exones mediante el secuestro de los represores de splicing que directamente se unen y ocultan elementos críticos en *cis* de los exones regulados o mediante secuestro de los represores a las zonas de unión que flanquean los exones regulados creando una zona de silencio (Del Gatto-Konczak et al., 2000).

2.1.2 Elementos *trans*

Los elementos *trans* incluyen las proteínas que están implicadas en los procesos requeridos para identificar las uniones intrón/exón y catalizar las reacciones de corte y empalme que escinden los intrones y unen los exones en el orden correcto. El spliceosoma está compuesto por cinco pequeñas ribonucleoproteínas (snRNPs) llamadas U1, U2, U4, U5 y U6 y más de 150 proteínas, incluyendo quinasas, fosfatasa y helicasas, muchas de ellas se requieren para la función spliceosomal, así como proteínas asociadas con factores de transcripción (Wang and Cooper, 2007). Los sitios de ramificación y 5' (Figura 2) sirven de sitios de unión para los componentes de ARN, las pequeñas ribonucleoproteínas nucleares U2 y U1 (snRNPs), respectivamente. Esta proporción ARN:ARN determina la unión precisa de los exones en los nucleótidos correctos. A nivel de exón, algunos ESEs se unen a proteínas ricas en serina/arginina (*SR proteins*) y secuestran y estabilizan la unión de componentes del spliceosoma como U2AF, mientras que otros ESEs se unen a otras proteínas activadoras de splicing que contribuyen al correcto splicing. Además, algunos ESSs se unen a componentes proteicos de ribonucleoproteínas heterogéneo-nucleares (hnRNP) que reprimen el uso de exones. A nivel de intrón, algunos ISEs e ISSs se unen a factores de splicing auxiliares que no están normalmente asociados con el spliceosoma para regular el splicing.

La reacción de splicing consta de dos etapas, en cada etapa se da una reacción de transesterificación, catálisis ejercida por los snRNP. También intervienen enzimas con función ATPasa y helicasa (Staley and Guthrie, 1998), que ayudan a llevar a cabo los cambios conformacionales del spliceosoma necesarios para las reacciones bioquímicas. En el primer paso, el grupo hidroxilo 2' de un residuo especial, una Adenina, del sitio de ramificación ataca el fosfato del sitio del splicing 5' intrónico. Esto lleva al corte del exón I a partir del intrón y la ligación del extremo 5' del intrón al hidroxilo 2' de la A del punto de ramificación, formando una estructura de lazo. El segundo paso de la transesterificación es el ataque del hidroxilo-3' del exón I, al fosfato en el extremo 3' del intrón. De esta forma se ligan los dos exones y se libera el intrón, todavía en forma de lazo (Figura 3).

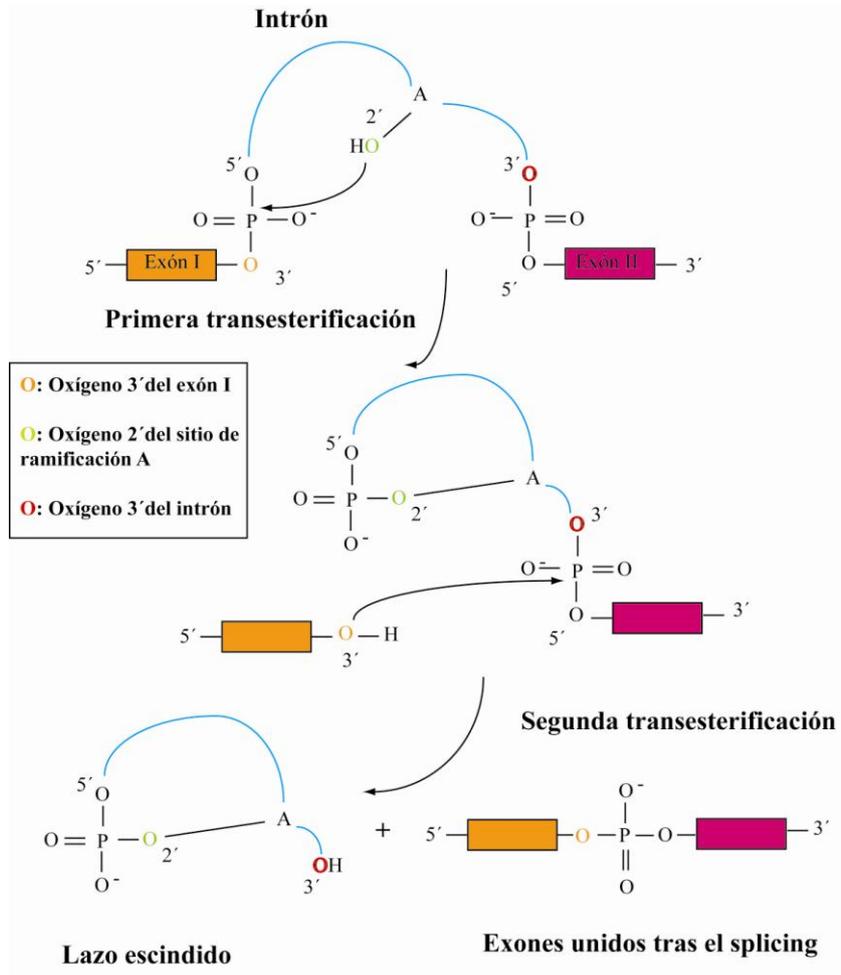


Figura 3. Representación de las etapas de transesterificación del proceso de splicing. Tras dos reacciones de transesterificación, se escinde el intrón y los exones quedan unidos.

2.2 Splicing alternativo

El splicing alternativo posibilita la generación de varios transcritos diferentes a partir de un único gen que, eventualmente, pueden ser traducidos y por tanto producir varias proteínas con funciones totalmente diferentes, o con la misma función pero con propiedades distintas. El splicing alternativo está considerado como uno de los mecanismos más importantes en la regulación génica en eucariotas superiores (Modrek and Lee, 2002).

2.2.1 Tipos de splicing alternativo

Se pueden definir diferentes eventos de splicing (Blencowe, 2006): la posibilidad de alternar el primer exón del mensajero (codificante o no), la utilización de sitios de splicing 5' y 3' alternativos, la retención de un intrón, la exclusión en el mensajero final de ciertos exones o incluso la utilización de sitios de poladenilización y exones finales alternativos (Figura 4) (Villate et al., 2008). Todos estos diferentes eventos de splicing generan una enorme variabilidad en los mensajeros, insertando o excluyendo nucleótidos, cambiando la pauta de lectura, introduciendo codones de terminación en las regiones codificantes o modificando los extremos UTR (*UnTranslated Region*, región no traducida) pudiendo afectar, por tanto, a la expresión génica mediante la inserción o eliminación de elementos reguladores que controlan la traducción, la estabilidad del mensajero o su localización (Smith and Valcarcel, 2000).

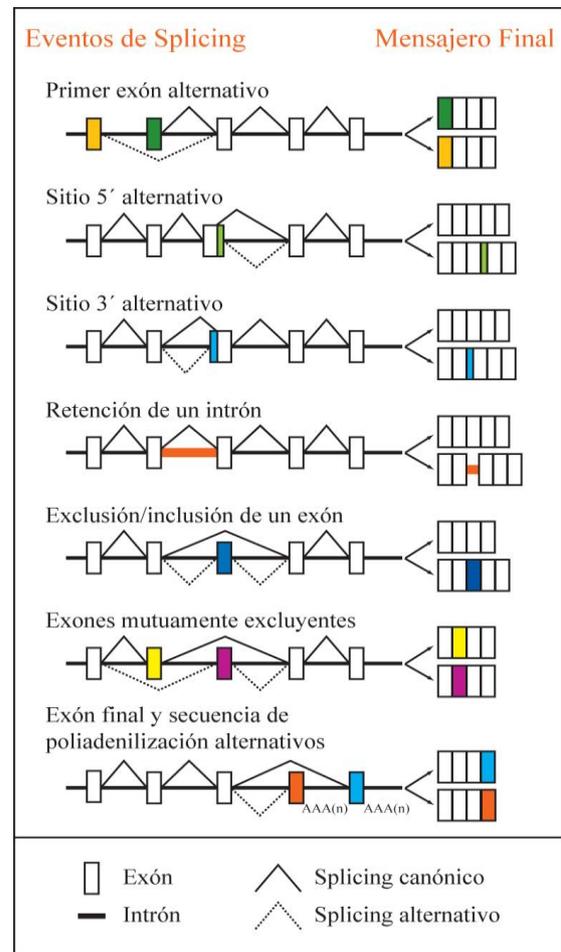


Figura 4. Clasificación de los diferentes eventos de splicing.

La regulación del splicing es, en consecuencia, un mecanismo de control de la expresión génica que permite la generación de diversidad transcriptómica, y por tanto proteómica, lo que juega un papel esencial en muchos procesos biológicos, como el desarrollo embrionario y el crecimiento celular (Hagiwara, 2005). Los procesos de splicing están modulados de acuerdo al tipo celular, el estado de desarrollo o en respuesta a estímulos externos (Faustino, 2007). Por lo tanto, el conocimiento del splicing alternativo y su regulación es necesario para la comprensión de la expresión génica en diferentes tejidos y especies. Este mecanismo es parte de un proceso normal

de variación y cuando este proceso se ve alterado puede conllevar al desarrollo de enfermedades (Villate et al., 2008).

2.3 Splicing y enfermedades

La regulación del splicing juega un papel muy importante en las patologías humanas (Wang and Cooper, 2007). Una alta proporción (60%) de mutaciones causantes de enfermedades en humanos se localizan en las secuencias implicadas en el proceso de splicing del ARNm (Faustino and Cooper, 2003; Villate et al., 2008). Las mutaciones pueden afectar a los elementos *cis*, básicos o reguladores, modificando la función y/o expresión de un sólo gen, cambiando por ello la proporción de las isoformas proteicas naturales de dicho gen. Esto aparece claramente ejemplificado en enfermedades como la fibrosis quística atípica o la β -talasemia (Tabla 2) (Villate et al., 2008; Wood et al., 2007). Algunas mutaciones pueden dañar la función de un elemento *trans* lo que influye sobre múltiples genes, como ocurre en la retinitis pigmentosa y en la distrofia miotónica (Tabla 2) (Blencowe, 2006; Faustino and Cooper, 2003; Garcia-Blanco et al., 2004; Lunn and Wang, 2008; Villate et al., 2008). El splicing aberrante se ha relacionado con procesos cancerígenos tal como se demostró en el estudio del splicing en el gen *CD44*, una glicoproteína de membrana, que presenta múltiples mensajeros diferentes, incluyendo la isoforma CD44v6 que aparece sobre-expresada en múltiples tumores y contra la que se están desarrollando nuevas estrategias terapéuticas (Heider et al., 2004).

Enfermedad	Número OMIM	Elementos alterados	Comentarios	
<i>Porfiria aguda intermitente</i>	#176000	elemento- <i>cis</i>	Sustitución de nucleótidos (C por G) en el exón 3 de la desaminasa porfobilinógena	
<i>β-talasemia</i>	#603902	elemento- <i>cis</i>	Mutaciones en el intrón 2 de la <i>β-globina</i> que genera un sitio aceptor 3' críptico	
<i>Heterotopia nodular bilateral periventricular (BPNH)</i>	#300049	elemento- <i>cis</i>	Mutación de sentido erróneo en el exón 6 (G por C) de la <i>Filamina A (FLNA)</i>	
<i>Cáncer</i>	Cáncer de ovario y mama	+113705	elemento- <i>cis</i>	Mutación sin sentido de <i>BRCA1</i> en el exón 18 (G por T)
	Leucemias y sarcomas	*605221	elemento- <i>trans</i>	La interacción de la proteína FUS 1 con FUS está interrumpida
	Liposarcomas, leucemia aguda mieloide (AML)	*137070	elemento- <i>trans</i>	Translocación del gen <i>FUS (TLS)</i>
	Carcinoma hepatocelular	*604739	elemento- <i>trans</i>	Autoanticuerpos contra el factor de splicing HCC1
	Carcinoma celular renal	*605199	elemento- <i>trans</i>	Translocación de <i>SFPQ</i> (splicing factor proline- and glutamine-rich) y unión con el gen <i>TFE3</i> .
<i>Fibrosis quística (CF)</i>	#219700	elemento- <i>cis</i>	Cuatro mutaciones sin sentido del gen <i>CFTR</i> en los exones 3 (G por U, y C por T), 11 (C por T) y 20 (G por A)	
<i>Anemia de Fanconi (FA)</i>	#227650	elemento- <i>cis</i>	Mutaciones sin sentido del gen <i>FANCG</i> en el exón 8 (C por T)	
<i>Hemofilia A</i>	+306700	elemento- <i>cis</i>	Mutaciones sin sentido del <i>Factor VIII</i> en los exones 19 (G por T) y 22 (C por T)	
<i>Leucodistrofia metacromática</i>	#250100	elemento- <i>cis</i>	Mutación de sentido erróneo de la <i>Arylsulfatasa A</i> en el exón 8 (C por T)	
<i>Distrofia miotónica 1 (DM1)</i>	#160900	elemento- <i>trans</i>	Secuestro de la proteína de unión CUG (CUGBP1) por las repeticiones del trinucleótido (CUG) _n >50 en el gen <i>DMPK</i> .	
		elemento- <i>trans</i>	Secuestro de las proteínas MBNL1.	
<i>Retinitis pigmentosa</i>	#268000	elemento- <i>trans</i>	Más de 30 mutaciones conservadas en genes relacionados con snRNPs.	
<i>Atrofia muscular spinosa (SMA)</i>	#253300	elemento- <i>cis</i>	Deleción completa de <i>SMN1</i>	
		elemento- <i>cis</i>	Mutaciones sin sentido de <i>SMN1</i> en el exón 3 (G por A)	
		elemento- <i>cis</i>	Sustitución de nucleótidos (C por T) en <i>SMN2</i> que interrumpen una secuencia ESE en el exón 7 que causa exclusión de exones	

Tabla 2. Enfermedades debidas a splicing aberrante (Villate et al., 2008).

2.4 Splicing y desarrollo de nuevas terapias

Debido a la fuerte asociación entre splicing aberrante y enfermedad es imprescindible un amplio conocimiento de los diferentes mensajeros generados por cada gen en cada situación fisiológica para así poder desarrollar una terapia eficaz y específica carente de efectos secundarios indeseados. Si se ignora el proceso de splicing en el proceso de desarrollo de nuevos tratamientos para enfermedades, sólo se está estudiando un porcentaje de las proteínas y por lo tanto se pueden perder muchas dianas potenciales. El splicing tiene un gran impacto en el desarrollo de fármacos y en las aplicaciones diagnósticas. Las variantes de splicing pueden tener diversas funciones debido a propiedades regulatorias diferentes y/o cambios estructurales que crean nuevos dominios. Además, las variantes solubles con funciones relacionadas con enfermedades pueden estar produciéndose en tejidos específicos, por lo que pueden ser firmes candidatas para el desarrollo de tratamientos. En este contexto, un fármaco puede tener efecto en varias formas de splicing causando daños colaterales, por lo que se necesitan fármacos específicos que actúen en la variante de splicing de interés (Villate et al., 2008).

Las mutaciones que causan enfermedades pueden afectar a los sitios de splicing así como a las secuencias reguladoras dando lugar a proteínas defectuosas o alteradas. Por lo tanto, las estrategias a seguir para corregir los procesos de splicing aberrantes serían actuar sobre las secuencias mutadas o los factores que se unen a ellas. Se han propuesto muchas aproximaciones, desde pequeñas moléculas que actúen como fármacos a terapias génicas dirigidas al ARN (García-Blanco et al., 2004; Villate et al., 2008). Las estrategias basadas en ARN ofrecen una serie de aplicaciones terapéuticas novedosas, incluyendo la alteración del procesamiento del transcrito de ARNm de interés, reprogramación de los defectos genéticos a través de la reparación del ARNm y el silenciamiento de isoformas específicas (Villate et al., 2008). La reprogramación del proceso de splicing es una nueva forma de terapia basada en la modificación del ARNm sin cambiar la secuencia del gen. Dentro de esta estrategia novedosa nos encontramos con distintas aproximaciones como, la modificación del procesamiento de ARNm, la alteración de la secuencia del ARNm (*trans*-splicing) y fármacos diseñados contra reguladores de splicing. El uso de pequeñas moléculas como fármacos (Figura 5) es una

estrategia atractiva para modificar los patrones de splicing por la facilidad de suministro y del control de dosis (Villate et al., 2008).

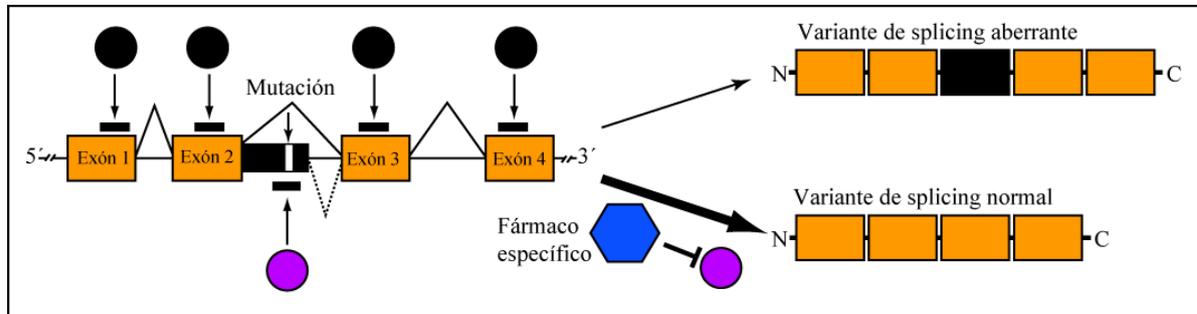


Figura 5. Reprogramación del proceso de splicing mediante pequeños fármacos. Se representa un gen con cuatro exones. Los círculos negros representan los reguladores de splicing que actúan en los sitios correspondientes representados por barras negras. El rectángulo negro representa una parte de un intrón con una mutación (barra blanca) creando un sitio de unión a un nuevo regulador (círculo morado), que activa un sitio de splicing 5' crítico, dando lugar al splicing de una secuencia adicional en el ARNm final y produciendo una proteína defectuosa. La terapia consiste en un fármaco específico que impide la unión del nuevo regulador y se restaura el proceso de splicing normal.

El proceso de splicing alternativo no sólo contribuye a aumentar la diversidad transcriptómica y proteómica, si no que, además, como alteraciones en este proceso provocan múltiples enfermedades, es de gran importancia su estudio para el desarrollo de nuevas aplicaciones terapéuticas. Por lo tanto, el análisis de este complejo mecanismo en regiones del genoma implicadas en enfermedades es de gran relevancia.

3. EL COMPLEJO MAYOR DE HISTOCOMPATIBILIDAD (MHC) HUMANO

El Complejo Mayor de Histocompatibilidad (Major Histocompatibility Complex, MHC) humano se descubrió hace más de 50 años. Inicialmente era conocido por su papel en los trasplantes a través de los antígenos de histocompatibilidad, de ahí su otro nombre comúnmente usado Antígeno Humano Leucocitario (*Human Leukocyte Antigen*, HLA) (Bodmer, 1972; Dausset, 1981; Snell, 1948). Desde su descubrimiento, el MHC ha estado en la vanguardia de la investigación genética. Como consecuencia de su particular estructura genómica, su polimorfismo remarcable, su alta densidad génica

y su implicación en numerosas enfermedades, el MHC está considerado como una región modelo para los estudios genómicos (Vandiedonck and Knight, 2009).

3.1 Clases del MHC

El MHC humano está localizado en el cromosoma 6p21.3 y comprende aproximadamente 7´6 Mb. Clásicamente se ha dividido en tres regiones denominadas clase I (telomérica), clase II (centromérica) y clase III (central). Mediante análisis de ligamiento se ha demostrado que las regiones de clase I y II presentan extensiones más allá de lo anteriormente descrito (Horton et al., 2004). A estas regiones se las conocen como las regiones extendidas de clase I y II (Figura 6) (Horton et al., 2004; Stephens et al., 1999). Como consecuencia de su importancia biológica, conocer la secuencia del MHC humano era una prioridad y se consiguió secuenciar las clases I, II y III en 1999 gracias al Consorcio de Secuenciación del MHC (MHC Sequencing Consortium) dirigido por Stephan Beck, Daniel Geraghty, Hidetosi Inoki y Lee Rowen (1999). Posteriores estudios y la secuenciación del cromosoma 6 (Mungall et al., 2003; Stephens et al., 1999), dieron lugar al conocimiento de las secuencias de las regiones extendidas. Desde que se secuenció el MHC completo, se han identificado 421 loci, el 60% de los cuales se expresan y sólo un 28% tienen un papel definido en el sistema inmune (Horton et al., 2004).

La mayoría de los genes localizados en las **clases I y II** codifican proteínas de superficie celular altamente polimórficas implicadas en la presentación antigénica a las células T, CD8+ y CD4+ respectivamente, durante la respuesta inmune (Benacerraf, 1981). El extraordinario polimorfismo de estas moléculas garantiza una amplia diversidad para el reconocimiento de antígenos y por lo tanto, reactividad contra los patógenos, confiriendo una ventaja selectiva. Se cree que el polimorfismo de los genes de clase I y II se mantiene por la ventaja selectiva conferida por el estado heterocigoto permitiendo una respuesta inmune más diversa a las infecciones de patógenos tales como el HIV-1 (Carrington et al., 1999).

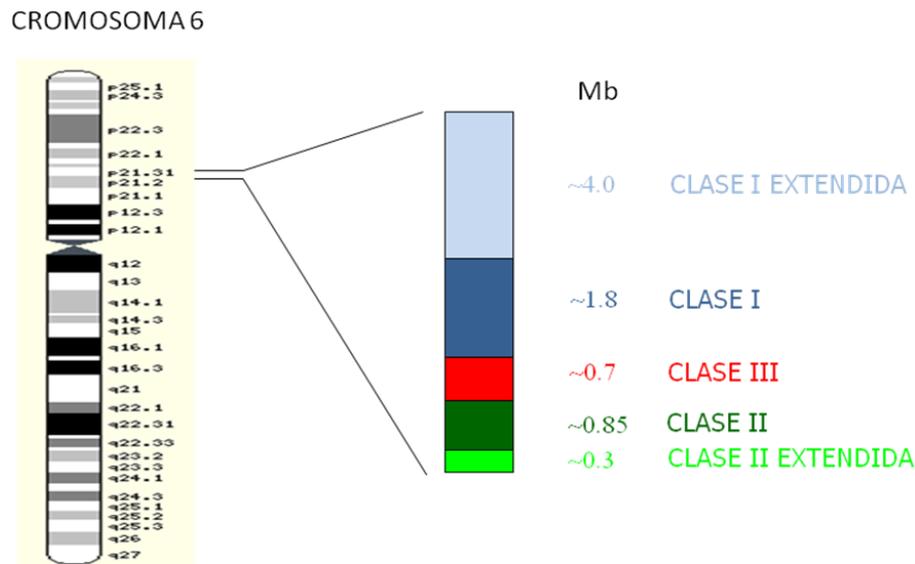


Figura 6. Representación esquemática del MHC humano, situado en el cromosoma 6. Están representadas las clases del MHC así como su tamaño en Mb.

La región central de **clase III** tiene un tamaño de 0.9 Mb y contiene 59-63 genes y 0-2 pseudogenes, dependiendo del haplotipo (Figura 7) (1999). Es la región más densa en genes de todo el genoma humano (Vandiedonck and Knight, 2009; Xie et al., 2003) y podemos encontrar algunos con funciones propias del sistema inmune como los componentes del complemento *C4A*, *C4B* y el *Factor B (BF)*, citoquinas pro-inflamatorias como *Factor de Necrosis Tumoral (TNF)* y las *linfotoxinas α y β (LT α y LT β)* (Xie et al., 2003). Sin embargo, los genes localizados en la clase III tienen un amplio espectro de funciones incluyendo una *valyl tARN sintetasa (VAR2)* (Hsieh and Campbell, 1991), el *receptor de productos de glicosilación avanzada (RAGE)* (Neeper et al., 1992), un posible miembro de la familia IkappaB (*NFKB1*) (Albertella and Campbell, 1994) y enzimas como la *ácido lisofosfatídico acil transferasa (AGPAT1)* (Aguado and Campbell, 1998), la *palmitoil proteína esterasa (PPT2)* (Aguado and Campbell, 1999) y la *sialidasa Neu 1* (Milner et al., 1997). Adicionalmente, hay genes potencialmente implicados en la regulación transcripcional y el procesamiento de ARNm como *PBX2* (Aguado and Campbell, 1995), y *UAP56 (BAT1)* un factor de splicing que está implicado en el transporte del ARNm (Luo et al., 2001).

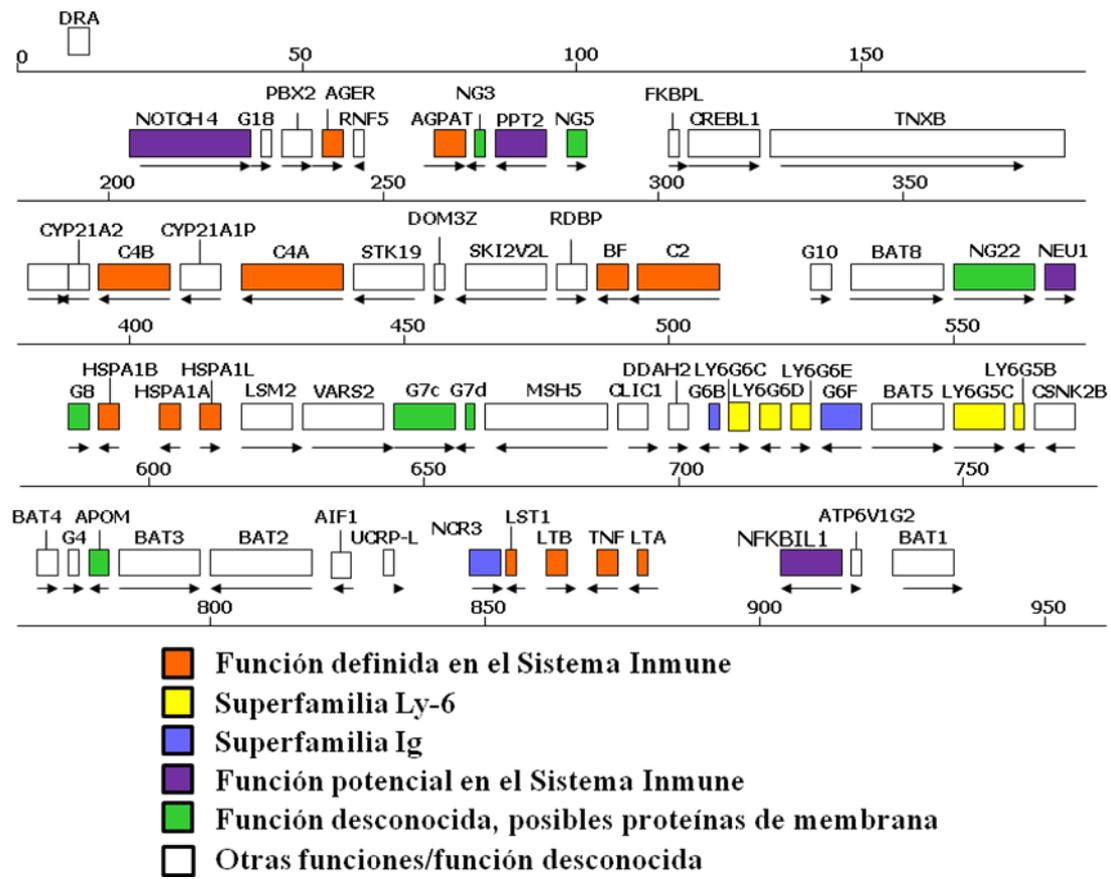


Figura 7. Representación de los genes de la región de clase III del MHC humano.

3.2 Polimorfismos en el MHC

La naturaleza de la variación genética humana engloba desde cambios de un sólo nucleótido a grandes variaciones estructurales genómicas de miles de pares de bases (Scherer et al., 2007). Aquí se engloban los SNPs (*Single Nucleotide Polymorphisms*, polimorfismos de un único nucleótido), los polimorfismos de inserciones y deleciones (DIPs, *Deletion Insertion Polymorphisms*) así como variaciones que implican a más de una base como las repeticiones pequeñas en tándem (STRs, *Short Tandem Repeats*), las grandes inserciones y deleciones (INDELs, *Insertions Deletions*), inversiones y otras repeticiones.

Los genes de clase I y clase II son altamente polimórficos y recientemente se ha confirmado que *HLA-B* es el gen más polimórfico del genoma humano con 1195 alelos

y 86 SNPs por kilobase (Kidd et al., 2008). Dentro de las variaciones, los SNPs son las más comunes y dentro de este tipo, los SNPs codificantes (cSNPs) son los más informativos. Dentro del MHC, nos encontramos con un total de dos inversiones, 63 INDELs (desde 100 pares de bases hasta 1 kilobase) y 181 variantes de número de copias (CNV, *Copy Number Variation*) (Horton et al., 2004), algunas de las cuales de gran importancia funcional como la variante CNV-507 o CNV-505 que solapan los loci de susceptibilidad de enfermedades como sarcoidosis y psoriasis, respectivamente (Traherne, 2008). Además, dentro del MHC nos encontramos con dos regiones afectadas por duplicaciones, la región HLA-DRB1 en clase II y la región del complemento en clase III (Horton et al., 2004; Traherne, 2008).

3.3 MHC y su función en el sistema inmune

En humanos, más del 5% de los genes expresados están relacionados con la defensa inmune (Trowsdale and Parham, 2004). Algunos genes del sistema inmune presentan unas características que reflejan su papel crucial como poligenia, polimorfismo y rápida evolución, rasgos que están presentes en algunos genes del MHC (Horton et al., 2004).

Muchos de los genes del MHC, incluyendo los genes de clase I y clase II de HLA, son parte del sistema inmune adaptado. Este sistema, caracterizado por la memoria inmunológica, se ha desarrollado para proporcionar flexibilidad en la respuesta a numerosos y altamente variables dianas antigénicas. Por el contrario, el sistema innato reconoce microbios patógenos a través de un número limitado de receptores. Esta división del sistema inmune no es precisa y ambos sistemas están entrecruzados. El MHC ayuda a proporcionar el nexo de unión entre el sistema inmune innato y adaptado. Además del papel de presentación antigénica de las moléculas de clase I, éstas también son utilizadas por las células NK (*Natural Killer*), las cuales se engloban dentro del sistema innato, para manifestar la evasión del sistema inmune de un patógeno que ha regulado negativamente la expresión de los antígenos de clase I (Horton et al., 2004).

En las primeras estimaciones (Bodmer, 1972; Trowsdale, 1993) se calculaba que aproximadamente un 40% de los genes expresados del MHC estaban relacionados con el sistema inmune. Horton y colaboradores (2004) recopilaron los genes de todo el

MHC, incluyendo las regiones extendidas, relacionados potencialmente con el sistema inmune y calcularon que representaban un 28% de todos los genes. Estos genes, pertenecientes a todas las clases descritas del MHC, se pueden agrupar en una o más de las siguientes categorías: procesamiento y presentación antigénica, superfamilia de inmunoglobulinas, inflamación, maduración de leucocitos, cascada del complemento, familia de receptores de clase I del MHC no clásica, regulación del sistema inmune y respuesta a estrés. Esta variedad de diferentes funciones inmunes puede ayudar a comprender algunas de las complejas asociaciones de esta región con enfermedades. Sin embargo, hay todavía genes que por su secuencia se encuentran relacionados con el sistema inmune pero de los que todavía se desconoce su función.

3.4 MHC y enfermedades

Se han asociado cientos de enfermedades con el MHC humano, siendo la mayoría enfermedades de carácter autoinmune. Establecer una relación directa entre una enfermedad y un gen del MHC es muchas veces una tarea difícil por dos causas: una, que la mayoría de las enfermedades asociadas al MHC son enfermedades multifactoriales y dos, el desequilibrio de ligamiento que caracteriza a algunos haplotipos del MHC provoca una difícil identificación del locus causante de la enfermedad (Horton et al., 2004).

Mutaciones puntuales, deleciones y otras variantes dentro del MHC son responsables de enfermedades monogénicas de herencia mendeliana como la hemocromatosis (Feder et al., 1996) o la hiperplasia adrenal congénita (Miller and Morel, 1989). Pero más significativo es el papel de esta región del genoma humano en muchas enfermedades multifactoriales (Vandiedonck and Knight, 2009). Muchas enfermedades autoinmunes están asociadas con el MHC, incluyendo la diabetes tipo I (Cudworth and Woodrow, 1975; Todd et al., 1987) y la esclerosis múltiple (Fogdell et al., 1995) entre otras, mientras otras asociaciones incluyen enfermedades infecciosas e inflamatorias (Hill, 2006) así como también cáncer (Howell et al., 2002) y en otros aspectos como la elección de pareja (Jacob et al., 2002). Frecuentemente, estas asociaciones no pueden ser explicadas de forma concluyente por alteraciones en los genes de las regiones de clase I y clase II. Por ejemplo, la deficiencia de IgA y la Inmunodeficiencia Variable Común (CVID) se han ligado a la región de clase III

(Cucca et al., 1998; Schroeder et al., 1998). Otra enfermedad asociada a esta región es la Orquitis, una enfermedad autoinmune de los testículos, de la cual se sabe que el gen de susceptibilidad *Orch-1* se encuentra en la región de clase III entre los genes *G7* y *HSP70* (Snoek et al., 1998). La artritis reumatoide (AR) es un ejemplo de enfermedad cuyo mapeo está definido en el MHC (Bali et al., 1999), teniendo una fuerte asociación con alelos particulares de la región de clase II, aunque se han descrito polimorfismos en la región de clase III que también pueden jugar un papel importante en la susceptibilidad a esta enfermedad (Jenkins et al., 2000; Kilding and Wilson, 2005).

En resumen, el MHC humano está asociado a numerosas enfermedades de las cuales muchas no pueden explicarse por variaciones en las clásicas regiones de clase I y II. Por lo tanto, el estudio de los genes de la región de clase III, especialmente de aquellos genes con un papel potencial en el sistema inmune, pueden ayudar al conocimiento de estas enfermedades.

3.5 Perfil de metilación del ADN en el MHC

Además de la variación genética, existen difentes causas que dan lugar a diferencias en la expresión génica. Los cambios epigenéticos, que son cambios que no afectan a la secuencia del ADN, como las modificaciones de las histonas o la metilación del ADN, están siendo estudiados por su papel en el desarrollo de muchas enfermedades, entre ellas enfermedades de carácter autoinmune.

Una de las principales modificaciones epigenéticas en los organismos multicelulares es la metilación reversible de la citosina (Bird, 2002). En mamíferos, la metilación de las citosinas ocurre casi exclusivamente en los dinucleótidos CpG, que prácticamente sólo aparecen en el genoma en las islas CpG. Estas islas son pequeñas regiones ricas en CpG que, en muchos casos, están asociadas con las regiones promotoras. La metilación de las citosinas resulta en una represión transcripcional o bien interfiriendo con la unión de los factores transcripcionales o bien induciendo una estructura de cromatina represiva (Jaenisch and Bird, 2003). Alteraciones en la metilación del ADN están asociadas a muchas enfermedades y se ha estudiado especialmente su relación con el desarrollo del cáncer (Jones and Baylin, 2002; Weber et al., 2005). Actualmente se están realizando estudios a nivel de metilación del ADN

de las regiones promotoras de genes implicados en patologías con el fin de comprobar si este mecanismo es responsable de las diferencias encontradas en los niveles de expresión del ARNm entre pacientes e individuos sanos (Nile et al., 2008).

El MHC humano fue elegido para estudiar el fenómeno de la metilación dentro del Proyecto del Epigenoma Humano (*Human Epigenome Project*, HEP) (Rakyan et al., 2004) por tres razones principales: (1) el MHC está asociado con más enfermedades que ninguna otra región del genoma humano, y por lo tanto, los datos generados son de interés para los investigadores con distintos intereses biomédicos, (2) es la región más polimórfica del genoma, lo que permite el estudio de la pérdida y ganancia de las islas CpG debidas a los SNPs en los sitios de metilación y (3) en el momento en que este proyecto fue comenzado (Beck et al., 1999), el MHC era una de las pocas regiones secuenciada y anotada dentro del genoma humano. Lo que pretende el HEP es identificar, catalogar e interpretar el fenómeno de la metilación del ADN a través del genoma.

Las modificaciones epigenéticas en las regiones reguladoras se correlacionan con el estado transcripcional del gen en cuestión (Bird, 2002). En nuestro genoma, la mayoría de las islas se encuentran hipometiladas durante el desarrollo (Antequera, 2003), debido al estado transcripcional activo, aunque una pequeña proporción se hipermetila y esto se correlaciona con un silencio transcripcional del gen asociado (Rakyan et al., 2004). Para realizar el perfil de metilación del MHC humano, se eligieron las siguientes regiones de interés (1) una región potencial reguladora para cada gen y (2) la región más densa en islas CpG. Los datos del estudio del MHC humano demostraron que de la mayoría de los amplicones que solapaban las islas CpG estaban hipometilados. Los resultados de este estudio piloto también demostraron que una proporción significativa de los loci analizados dentro del MHC mostraban perfiles de metilación específicos de tejido y que las diferencias de metilación inter-individuales eran bastante comunes. Además, las diferencias específicas de tejido en la metilación del ADN sugieren que los mecanismos epigenéticos están relacionados con el uso de sitios de comienzo de transcripción alternativos (Rakyan et al., 2004). Para elucidar la contribución de la metilación al desarrollo de enfermedades se necesita un estudio a

nivel de genoma de los patrones de metilación específicos de tejido, lo que se conoce como “metiloma”(Butcher and Beck, 2008).

3.5 Splicing en el MHC

El splicing alternativo también es común en genes del MHC humano como se ha demostrado en varios estudios (Kwan et al., 2008; Kwan et al., 2007; Pan et al., 2008). Por ejemplo, las variantes de splicing de las cadenas pesadas de moléculas del MHC de clase I aportan mayor complejidad al mecanismo de presentación antigénica (Belicha-Villanueva et al., 2009).

Estudios transcriptómicos de algunos genes de la región de clase III del MHC indican una alta tasa de diferentes eventos de splicing. Previamente se han definido estos mecanismos de splicing en varios genes como *LY6G6C*, *LY6G6D*, *LY6G6E*, *LY6G5C*, *LY6G5B* (Calvanese et al., 2008; Mallya et al., 2002), *NKp30* (Neville and Campbell, 1999), *G6B* (de Vet et al., 2001) y *RAGE* (Sterenczak et al., 2009) entre otros. Incluso en esta región del genoma se han descrito transcritos quimera entre varios de sus genes (Calvanese et al., 2008). Algunas isoformas solubles del gen *LST1*, otro gen perteneciente a esta región de clase III, están implicadas en enfermedades inflamatorias e infecciosas (Mulcahy et al., 2006) por lo que variantes de splicing están relacionadas directamente con enfermedades.

4 EL GEN *NFκBIL1* (nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor-like 1).

La región telomérica de la clase III del MHC humano, situada junto a la clase I, es particularmente rica en genes conteniendo al menos 11 genes dentro de un intervalo de 82 kb: *BAT1*, *ATP6V1G2*, *NFκBIL1*, *LTα*, *TNF*, *LTβ*, *LST1*, *NCR3*, *AIF-1*, *BAT3* y *BAT2*. La función de alguno de estos genes está pobremente caracterizada, aunque existen evidencias de su relación con respuestas inmunes e inflamatorias (Neville and Campbell, 1999). Por ejemplo, el gen *AIF-1* codifica una proteína específica de macrófagos de respuesta a citoquinas (Utans et al., 1995), y las citoquinas *LTα*, *TNF*, *LTβ* están implicadas en la respuesta inflamatoria (Gruss and Dower, 1995). Esta región

del MHC se ha asociado a enfermedades de carácter autoinmune como AR, lupus eritematoso sistémico y diabetes tipo I entre otras. Dentro de esta región se encuentra el gen *NFκBIL1*, relacionado también con varias enfermedades, algunas de las cuales de carácter autoinmune y de especial relevancia por su homología con la familia de inhibidores del factor nuclear κB.

4.1 El factor nuclear κB y la familia de inhibidores IκB

El gen *NFκBIL1* (*NF-κB inhibitor-like protein 1*) también conocido como *IκBL* (inhibitory κB like) está localizado a 30 Kb del gen *TNF* y debe su nombre a su homología con los miembros de la familia de proteínas de inhibidores del factor nuclear κB (IκB) (Albertella and Campbell, 1994). Estas proteínas son componentes críticos de la vía de señalización del factor nuclear NF-κB, el cual regula la transcripción de muchos mediadores de los procesos inflamatorios y de destrucción de tejidos en la articulación reumatoide, incluyendo TNF, IL-1 y IL-6 (Ghosh and Karin, 2002).

4.1.1 La familia NF-κB/Rel

El factor de transcripción NF-κB fue descubierto como parte de la maquinaria de transcripción de los genes de la cadena κ de las inmunoglobulinas en linfocitos B (de ahí su nombre) (Sen and Baltimore, 1986). En las células no estimuladas, el factor NF-κB existe como homo o heterodímero formando complejos con las proteínas inhibidoras de la familia IκB. Las diferentes subunidades proteicas que componen los dímeros se agrupan dentro de la familia Rel. En mamíferos se han descrito varias proteínas Rel: p65 (también denominada RelA), c-Rel, RelB, p100 (de donde proviene p52) y p105 (la cual es el precursor de p50) (Figura 8). Todas las proteínas de esta familia tienen una región conservada de 300 aminoácidos en el extremo N terminal llamado dominio Rel, que es responsable de la dimerización, la unión al ADN y la interacción con los inhibidores (Muller et al., 1995; Vallabhapurapu and Karin, 2009).

Las subunidades p65, c-Rel y RelB contienen en el extremo C terminal un dominio de activación de la transcripción (TAD) mientras que las subunidades p52 y p50 no poseen este dominio y dependen de la interacción con otros factores para regular positivamente la transcripción (Hayden and Ghosh, 2008). RelB posee una peculiaridad que no comparte con las demás subunidades, y es la presencia de un motivo de

cremallera de leucina en su extremo amino terminal, el cual juega un papel importante en la regulación de la transcripción (Figura 8) (Hayden and Ghosh, 2008). Mientras que RelB heterodimeriza preferentemente con p100 (Dobrzanski et al., 1995) y su forma procesada p52 (Yilmaz et al., 2003), p65 y c-Rel heterodimerizan predominantemente con p50 (Karin and Ben-Neriah, 2000).

Los extremos C terminal de las subunidades p100 y p105 poseen repeticiones de anquirina, lo que les permite actuar también como inhibidores de NF- κ B (Figura 8). A estas formas, p100 y p105, se les conoce como moléculas NF- κ B1 y NF- κ B2 respectivamente. Un proceso mediado por el proteosoma elimina la mitad del extremo C-terminal de p105 para generar la subunidad p50 y de p100 para generar la subunidad p52. El procesamiento de p105 ocurre constitutivamente en células no estimuladas pero cuando se estimulan con TNF- α o IL-1, p105 se fosforila y se degrada sin producción de p50. De esta manera, los precursores p100/p105 pueden funcionar como inhibidores cuando no sufren ninguna modificación o bien como miembros de la familia NF- κ B cuando son procesados por el proteosoma. En el extremo carboxilo terminal contienen además un dominio de muerte (D) que estaría implicado en la activación de la muerte celular (Hacker and Karin, 2002).

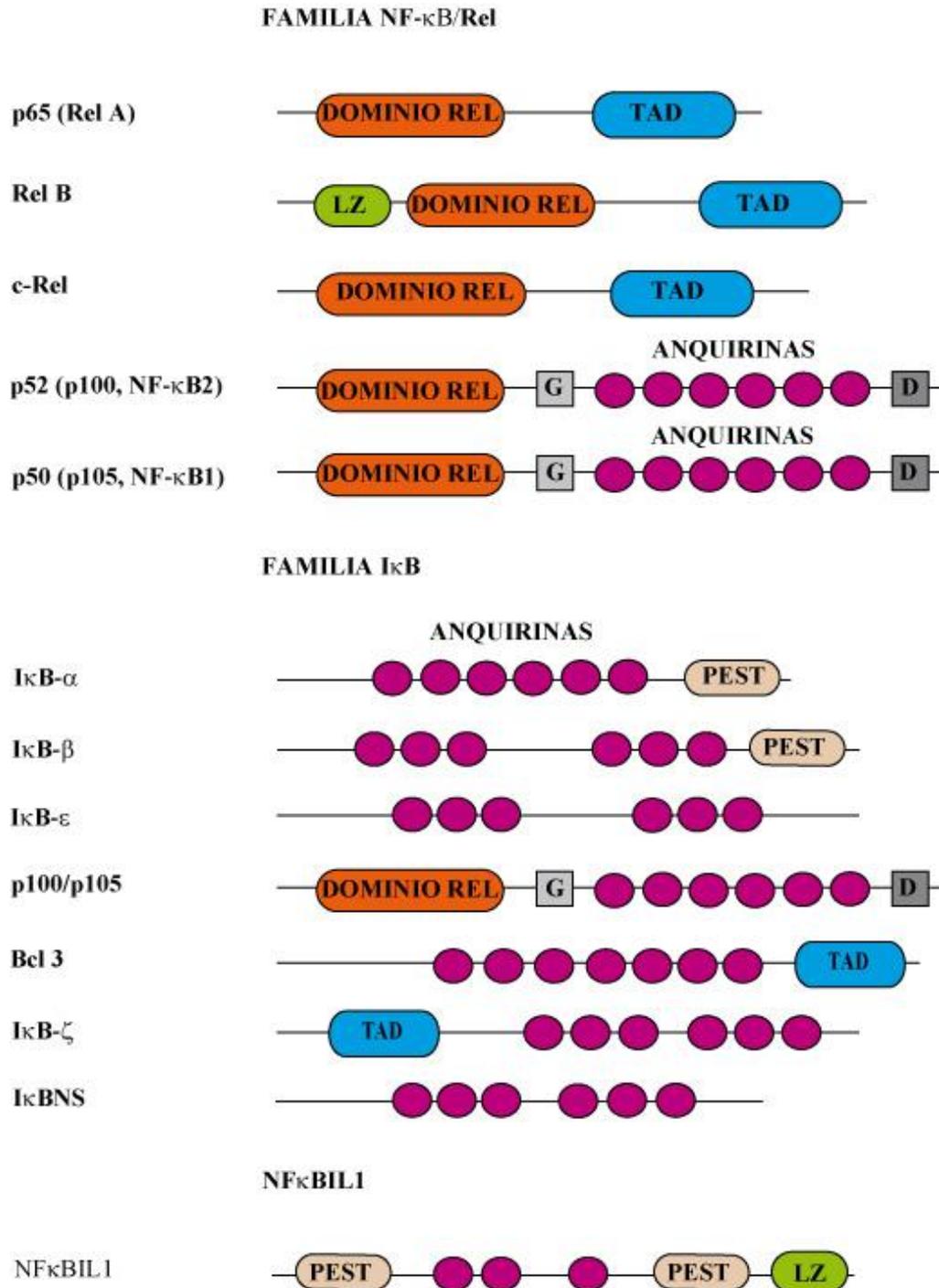


Figura 8. Estructuras de los miembros de las familias NF- κ B e I κ B y de NF κ BIL1. Se indican los dominios proteicos característicos. Los precursores p100/p105 funcionan como inhibidores o como miembros de la familia NF- κ B cuando son procesados por el proteosoma. TAD: Dominio de transactivación; G: Sitio rico en Glicina; D: dominio de muerte; LZ: cremallera de leucina, PEST: secuencias ricas en Prolina, Glutámico, Serina y Treonina.

4.1.2 La familia IκB

La familia de inhibidores IκB (Figura 8) comprende tres grupos funcionales, que son capaces de interactuar a su vez con distintos homo o heterodímeros de NF-κB: las proteínas IκB-α, IκB-β e IκB-ε que responden a los estímulos clásicos que activan NF-κB como citoquinas inflamatorias y antígenos bacterianos; los extremos C-terminales de las proteínas precursoras p105, también llamada IκB-γ y p100, también llamada IκB-δ; y las proteínas IκB atípicas que son IκB-ζ, Bcl3 (linfoma 3 de células B) e IκBNS, que generalmente no se expresan en células sin estimular y median sus efectos en el núcleo (Ghosh and Hayden, 2008) (Vallabhapurapu and Karin, 2009). Todas las isoformas de IκB descritas hasta el momento, son muy semejantes en estructura tridimensional aunque tienen distinto número de dominios repetidos de anquirina que constituyen el sello característico de la molécula y que forman los sitios de interacción con NF-κB. La degradación diferencial de estas isoformas no sólo determina la velocidad de translocación del factor sino que también liberan diferentes cantidades de homo y heterodímeros de NF-κB.

Dentro de la familia de inhibidores IκB, la proteína IκB-α es la mejor caracterizada. La degradación de IκB-α es rápida tras la actuación de los estímulos adecuados como TNF-α y LPS, pero su ausencia es transitoria ya que una hora después de haber sido degradada aparecen niveles detectables de proteína. La transcripción de IκB-α está sujeta a un circuito auto-regulatorio inductor por ser uno de los primeros genes que se transcriben tras la entrada de NF-κB al núcleo. Su estructura consiste en un dominio central con seis repeticiones de anquirina y en su extremo carboxilo está contenida la región PEST (Figura 8). Las secuencias PEST son secuencias ácidas ricas en prolina, serina, ácido aspártico y treonina, que dirigen a las proteínas al proteosoma para su degradación. IκB-α regula predominantemente los heterodímeros p65:p50 de NF-κB. IκB-α regula además la actividad nuclear de NF-κB, gracias a la presencia de secuencias de exportación nuclear (NES) capaces de unirse al receptor CRM1 (Tam et al., 2000). Dentro del núcleo IκB-α puede unir formas libres de NF-κB e incluso disociar al factor del ADN. La entrada de este inhibidor al núcleo permite mantener la respuesta de NF-κB frente a estímulos persistentes.

El inhibidor I κ B- β presenta seis repeticiones de anquirina seguidas y en su extremo carboxilo terminal también se encuentran secuencias PEST (Figura 8). I κ B- β ha sido asociada con una respuesta bifásica y persistente de NF- κ B. Su degradación se inicia mucho más tarde que la de la proteína I κ B- α . La síntesis de novo de I κ B- β es más lenta que la de I κ B- α de manera que los niveles de este inhibidor en el citoplasma se normalizan más lentamente. También puede viajar al núcleo celular y capturar NF- κ B activo, sin embargo, no es capaz de romper las uniones funcionales entre este factor y la maquinaria de transcripción (Ferlito et al., 2001). El espectro de proteínas Rel con que puede interactuar I κ B- β es prácticamente el mismo que I κ B- α , aunque la afinidad es menor en el caso de la isoforma beta.

La proteína I κ B- ε tiene una estructura similar a las otras variantes referidas, con múltiples repeticiones de anquirina (Figura 8). Aunque también puede captar distintos dímeros de NF- κ B, se le ha implicado en la regulación de complejos de NF- κ B que contienen c-Rel (Whiteside et al., 1997). La degradación de I κ B- ε tiene una cinética parecida a la de I κ B- β , es decir ocurre a tiempos más largos que I κ B- α , sin embargo, la síntesis *de novo* es mucho más rápida de manera que los niveles basales de la proteína pueden restablecerse hasta tres horas después del estímulo. Se ha sugerido que este inhibidor puede regular un tipo de respuesta de NF- κ B lento y transitorio.

La proteína I κ B- γ corresponde a la región C terminal de p105 (Figura 8). Se han identificado además dos isoformas de I κ B- γ , I κ B- γ -1 e I κ B- γ -2. I κ B- γ -1 se puede encontrar en el citoplasma y en el núcleo mientras que I κ B- γ -2 se localiza preferentemente en el compartimento nuclear. Ambas formas tienen distintas afinidades relativas por p50. I κ B- γ se expresa en un número muy discreto de tipos celulares y tiene una marcada preferencia por homodímeros de p50 y c-Rel (Vallabhapurapu and Karin, 2009).

El extremo C terminal de la subunidad p100, también llamada I κ B- δ (Figura 8), se une preferentemente a RelB reteniéndolo en el citoplasma. Estudios recientes han demostrado que p100 también juega un papel importante en la regulación de los homodímeros de p65 (Vallabhapurapu and Karin, 2009).

La proteína Bcl-3, aunque estructuralmente es muy parecida al resto de la familia de inhibidores I κ B, tiene siete dominios de anquirina, su función parece ser completamente distinta (Figura 8) (Ghosh et al., 1998). De todos los miembros conocidos de I κ B, Bcl-3 es la única proteína que no es degradada por acción de las vías clásicas de NF- κ B. De hecho su localización es preferentemente nuclear y parece estimular el potencial transcripcional de NF- κ B. Bcl-3 generalmente captura dímeros p52/p50 y se cree que puede tener distintos efectos sobre la unión de estos al ADN dependiendo de su propio estado de fosforilación, de su concentración nuclear y de sus interacciones con otros factores nucleares (Heissmeyer et al., 1999). Lo que está claro es que puede translocarse al núcleo ya sea libre o unida a NF- κ B y regular la actividad transcripcional del factor.

I κ B- ζ es un miembro de la familia I κ B recientemente identificado que muestra una homología mayor con Bcl-3 que con otros miembros de la familia (Figura 8). En su extremo amino terminal se encuentra un dominio de transactivación y una señal de localización nuclear (NLS, *Nuclear Localization Signal*). En su extremo C-terminal contiene 6 repeticiones de anquirina y se expresa de una manera inducible en respuesta LPS, por ello se le conoce también como MAIL (*Molecule possessing Ankyrin repeats Induced by LPS*) (Seshadri et al., 2009). Presenta dos isoformas conocidas como MAIL-L y MAIL-S. Actúa en el núcleo, parece que potenciando la actividad transcripcional de los dímeros NF- κ B. Se asocia a homodímeros p50 y puede funcionar como coactivador de una manera similar a Bcl-3, aunque también regula negativamente los dímeros que contienen p65 (Vallabhapurapu and Karin, 2009). Esta isoforma nuclear se caracteriza por presentar un patrón de distribución en espículas nucleares, lo que puede indicar una interacción con la maquinaria de transcripción puesto que las espículas sirven de reservorio de factores de splicing.

I κ BNS es otra proteína de la familia I κ B que se induce rápidamente tras la ligación del receptor de células T (TCR) en timocitos. Contiene 6 repeticiones de anquirina (Figura 8). Correspondientemente, I κ BNS juega un papel importante en la selección negativa de las células T. Tiene preferencia por la unión de subunidades p50 *in vitro* y se induce fuertemente por la citoquina anti-inflamatoria IL-10 (Ghosh and Hayden, 2008).

4.1.3 Sistema NF- κ B/ I κ B

De manera constitutiva, los dímeros de NF- κ B se encuentran secuestrados en el citoplasma como moléculas inactivas formando un complejo con los inhibidores. La interacción con I κ B oculta una de las dos secuencias que determinan el destino nuclear de NF- κ B, de manera que para alcanzar el núcleo, debe disociarse primero de los inhibidores. La activación celular da lugar a una fosforilación rápida de los residuos de serina N-terminales de los inhibidores por parte del complejo multiproteico IKK (I κ B kinase), lo que resulta en la ubiquitinización y consecuente degradación de los mismos por el proteosoma. El factor nuclear NF- κ B se transloca al núcleo y se produce la transcripción de los genes diana, incluyendo los genes de *I κ B* (Figura 9) (Greetham et al., 2007). Este mecanismo es un ejemplo interesante dentro de la biología molecular, de cómo un factor de transcripción puede regular su propia actividad al promover la síntesis de la proteína que lo inhibe. Como se ha visto, la regulación de NF- κ B es un fenómeno complejo y multifactorial. Hay que resaltar que en una misma célula existen diferentes poblaciones del factor NF- κ B, que son sensibles a distintos estímulos y que están sujetos a complicados mecanismos de regulación.

La función de las proteínas I κ B ha evolucionado en años recientes desde ser simplemente inhibidores de la actividad de NF- κ B a ser reguladores más complejos de la expresión génica. Además de secuestrar dímeros de NF- κ B en el citoplasma, los miembros de la familia I κ B también estabilizan los dímeros nucleares. Algunas proteínas I κ B pueden regular positivamente la transcripción actuando como co-activadores transcripcionales. Por lo tanto, en lugar de considerar esta familia sólo como inhibidores de NF- κ B, sería más apropiado considerarlos como cofactores de NF- κ B (Ghosh and Hayden, 2008).

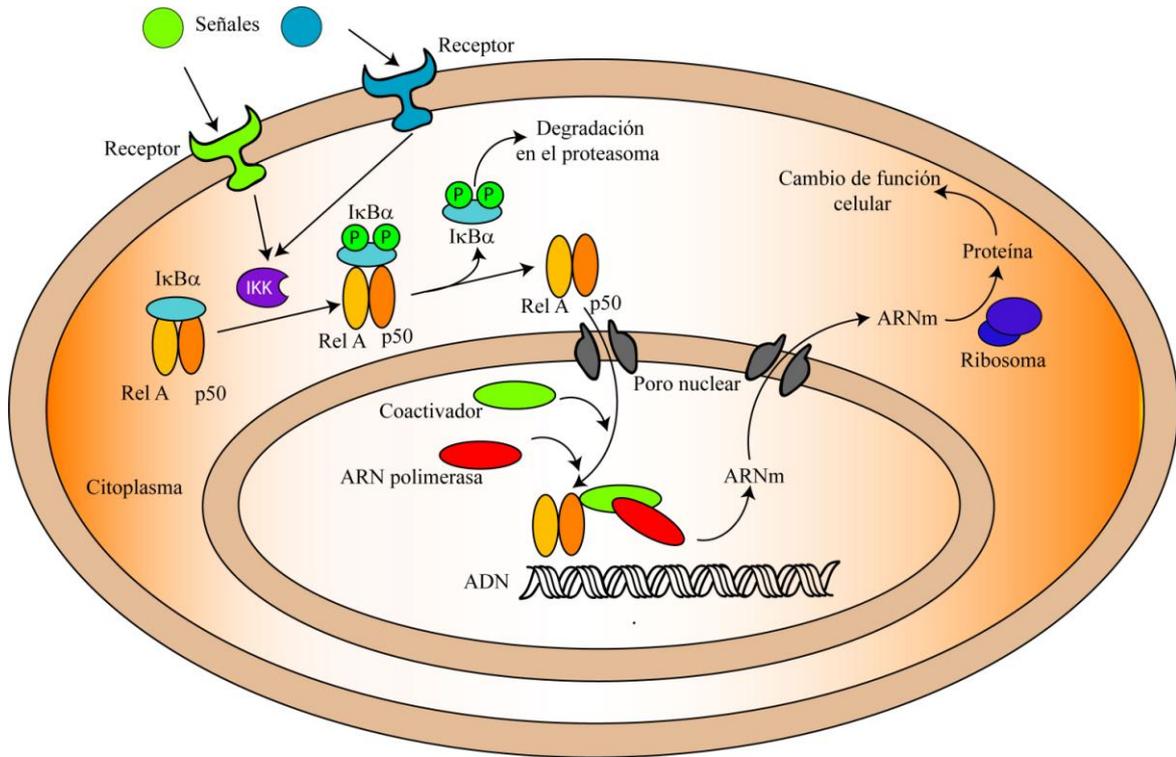


Figura 9. Mecanismo clásico de activación de NF- κ B. Tras las señales adecuadas, los inhibidores del factor NF- κ B se fosforilan por la quinasa IKK y se produce su degradación en el proteosoma. NF- κ B se transloca al núcleo y se transcriben los genes diana, entre los que se encuentra la familia de inhibidores.

4.1.4 Dominios presentes en la proteína NF κ BIL1

Los análisis de la secuencia de aminoácidos predicha de la proteína NF κ BIL1 revelan la presencia de tres repeticiones de anquirina en los exones II y III que se asemejan a las repeticiones de anquirina segunda y tercera de la subunidad p100 de NF- κ B (Albertella and Campbell, 1994). Estas secuencias de anquirina se encuentran presentes en los miembros de la familia de proteínas I κ B y actúan como dominios de interacción de proteína-proteína necesarios para la unión al factor nuclear κ B (Figura 8) (Hatada et al., 1993). Además de estas repeticiones de anquirina, la mayoría de los miembros de la familia I κ B contienen secuencias PEST en su extremo C-terminal (Lin et al., 1996). Estas secuencias PEST se encuentran en el extremo N-terminal de NF κ BIL1 y tras las repeticiones de anquirina, en el extremo amino terminal (Albertella and Campbell, 1994). En el extremo C-terminal de NF κ BIL1 se encuentran además tres

residuos de leucina y una valina cada siete aminoácidos, lo que podría constituir una reminiscencia de una cremallera de leucina (Figura 8) (Albertella and Campbell, 1994).

4.2 *NFκBIL1* y splicing

En humanos se caracterizaron dos formas de splicing de este gen llamadas *NFκBIL1-α* (NM_005007.3) y *NFκBIL1-β* (NM_001144961.1) (Semple et al., 2002), las cuales codifican proteínas de 381 y 366 aminoácidos respectivamente, que difieren en la presencia de 15 aminoácidos en la variante α que no están presentes en la variante β , concretamente en el final del exón III. El nuevo sitio de splicing alternativo en el exón III sigue la regla GT de los sitios consenso 5' de splicing y la pauta de lectura se mantiene en el exón IV. Las dos isoformas constan de cuatro exones y sólo presentan en sus secuencias la diferencia de 45 nucleótidos en el exón III (Figura 10).

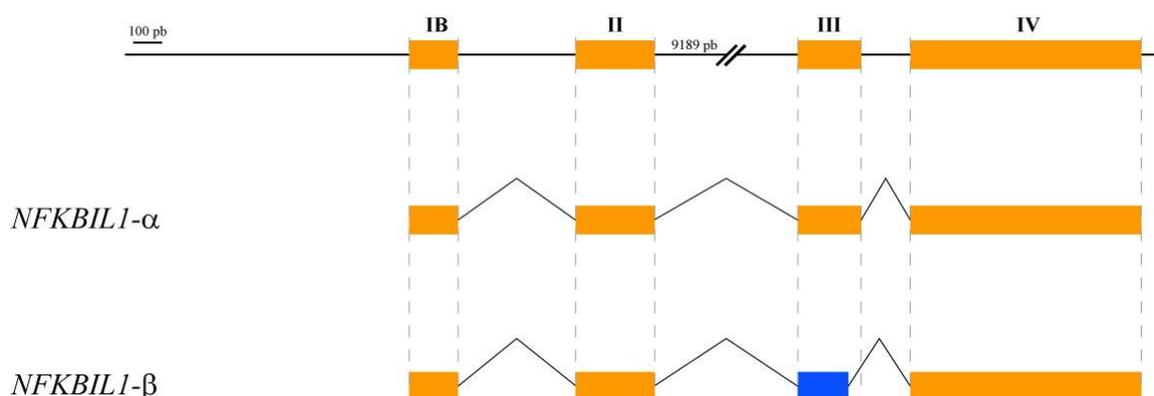


Figura 10. Isoformas de *NFκBIL1* descritas en humanos. La isoforma β consta de un exón III con una longitud 45 pb menor que la isoforma α .

Estas dos isoformas pueden asociarse con ARNm y con otras tres proteínas, identificadas por espectrofotometría de masas como son la leucofisina, el factor 1α de la elongación de la traducción y la sintasa I de CTP (Greetham et al., 2007). Se ha observado que la proteína *NFκBIL1* se acumula en espículas nucleares, lo que podría implicar un papel en la regulación del splicing (Semple et al., 2002). Estos datos apoyan el posible papel de *NFκBIL1* en el procesamiento de ARNm o en la regulación de la traducción.

4.3 Asociación del gen *NFκBIL1* con enfermedades

Polimorfismos en la región promotora de *NFκBIL1* se han asociado a enfermedades como la diabetes tipo I (Yamashita et al., 2004) y se han asociado al posible desarrollo de cardiomiopatías crónicas de Chagas entre los individuos infectados por *Trypanosoma cruzi* (Ramasawmy et al., 2008) y al posible riesgo de desarrollar lupus eritematoso sistémico y el síndrome de Sjögren (Castiblanco and Anaya, 2008). Asociaciones de polimorfismos de este gen con AR han generado mucha controversia. Okamoto y colaboradores (Okamoto et al., 2003) demostraron que *NFκBIL1* es un gen de susceptibilidad para el desarrollo de AR en pacientes japoneses mientras que un resultado similar no pudo ser encontrado en pacientes españoles (Collado et al., 2004; Martinez et al., 2003).

La AR es un desorden inflamatorio común de compleja etiología que afecta aproximadamente al 1% de la población mundial. Aunque las causas de esta enfermedad no están claras todavía, se atribuye a factores genéticos y ambientales. Durante las pasadas décadas, se han identificado muchas regiones genéticas asociadas con AR, entre ellas el MHC de clase III. Una explicación a la gran cantidad de genes identificados puede ser la existencia de interacciones entre genes, lo que hace el análisis más complejo. En un estudio reciente de posible interacciones génicas asociadas a AR, se ha visto que la interacción de *NFκBIL1* con *TNF* o *VEGF* (*Vascular Endothelial Grown Factor*) puede estar implicadas en el desarrollo de esta enfermedad (Huang et al., 2009). También se ha demostrado la presencia de proteína *NFκBIL1* en líquido sinovial reumatoide, sugiriendo un posible papel de este gen en la patogénesis de AR (Greetham et al., 2007).

Debido a la implicación de *NFκBIL1* en diversas enfermedades y a la presencia de varias isoformas, es de gran interés el estudio de este gen a un nivel más detallado. *NFκBIL1* se encuentra en una región del MHC donde mapean muchas enfermedades y en donde se ha comprobado que el fenómeno de splicing alternativo es frecuente. El splicing alternativo contribuye a la complejidad fenotípica con lo que es importante la comparación de este fenómeno entre diferentes especies para comprobar si en organismo más complejos se encuentran más variantes del mismo gen. Además, se han descrito varias isoformas de *NFκBIL1* con lo cual sería interesante estudiar la presencia

de otras posibles variantes, sus niveles de expresión e intentar ver la posible relación con enfermedades a las que se le ha asociado previamente. Todo esto constituye las bases para los estudios realizados en la presente tesis.

Por todo lo hasta ahora expuesto, nuestro interés se centrará en los siguientes objetivos:

1. **Describir todas las formas de splicing del gen *NFκB1* en humanos y en otras especies diferentes.** Desarrollando un estudio comparativo que nos ayude a entender cómo ha podido influir el splicing alternativo en la diversidad de las especies y en la complejidad del sistema inmune. Entre las especies elegidas además de humano podemos destacar vaca, cerdo, rata y ratón y un conjunto de primates, lo que nos permite el estudio del orden de los primates, los ungulados y los roedores.
2. **Cuantificar la expresión de los distintos transcritos del gen *NFκB1* en diferentes líneas celulares y tejidos humanos.** Una vez caracterizadas las formas de splicing, procederemos a la cuantificación de los niveles de expresión y determinaremos la abundancia relativa de las variantes empleando para ello RT-PCR cuantitativa en tiempo real en líneas celulares y distintos tejidos humanos.
3. **Estudiar la viabilidad a nivel proteico de las isoformas de *NFκB1* y su localización.** Se sobre-expresarán dichas variantes en células COS7 y se estudiará la localización celular de las proteínas codificadas por cada variante.
4. **Analizar el posible papel del gen *NFκB1* en enfermedades autoinmunes.** Se analizará la expresión de las variantes de splicing de este gen en muestras de pacientes con Artritis Reumatoide y en el caso de diferencias de expresión se estudiarán las posibles causas.

1. ANÁLISIS INFORMÁTICOS

El estudio del gen *NFκBIL1* se inició con la búsqueda de información en las diferentes bases de datos disponibles. Las secuencias codificantes de *NFκBIL1* (NM_005007.3) y su variante descrita (NM_001144961.1) se compararon con las bases de datos de marcadores de secuencia expresadas o ESTs en la base de datos National Centre for Biotechnology Information (NCBI) (<http://www.ncbi.nlm.nih.gov/BLAST/>). Para un estudio preciso de las formas de splicing de este gen, se consultaron otras bases de datos como Ensembl (<http://www.ensembl.org>), ECgene (<http://genome.ewha.ac.kr/ECgene/>) y SpliceInfo (<http://spliceinfo.mbc.nctu.edu.tw/>). También se realizó la búsqueda de las secuencias del gen *NFκBIL1* en otras especies en las bases de datos de NCBI y Ensembl.

Para estudiar las secuencias proteicas, se llevaron a cabo análisis independientes con cada forma de splicing utilizando programas informáticos de detección de dominios, tales como Pfam (<http://pfam.sanger.ac.uk/>), SMART (<http://smart.embl-heidelberg.de/>) e InterProScan (<http://www.ebi.ac.uk/InterProScan/>). Para el estudio de los motivos PEST en las secuencias, se utilizó el programa Epestfind (http://bioinfo.nhri.org.tw/cgi-bin/emboss/epestfind?_pref_hide_optional=1).

La región promotora de *NFκBIL1* se obtuvo de la base de datos de Ensembl. El análisis de la secuencia de la región promotora para el estudio de los sitios de unión de factores de transcripción se realizó con el programa TRANSFAC de Biobase (<http://www.biobase-international.com/pages/index.php?id=transfac>).

El alineamiento de secuencias se realizó con programas informáticos como BioEdit (<http://www.mbio.ncsu.edu/BioEdit/BioEdit.html>), MultAlin (<http://bioinfo.genotoul.fr/multalin/>) y ClustalW (<http://www.ebi.ac.uk/Tools/clustalw2/index.html>).

Los oligonucleótidos utilizados en los experimentos de este estudio se diseñaron utilizando el programa Primer3Plus (<http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>) y su calidad se evaluó en el programa Oligo Analyzer (<http://eu.idtdna.com/analyzer/applications/oligoanalyzer/default.aspx>).

2. CULTIVOS CELULARES

Se realizaron cultivos de la siguientes líneas celulares humanas: K562 (eritroleucemia), U937 (monocito), Molt4 (célula T), Raji (célula B), 143B (fibroblasto), HeLa (epitelial), HL60 (monocito) y Jurkat (célula T) y de células de mono: COS-7 (fibroblasto de mono verde). El medio de cultivo utilizado para las células en suspensión (K562, U937, Molt4, Raji, Jurkat, HL60) fue RPMI (*Roswell Park Memorial Institute Medium*) y para las adherentes (HeLa, 143B, COS-7) DMEM (*Dulbecco's Modified Eagles Medium*). Ambos medios contenían suero fetal bovino (FBS) (10% v/v), 2mM de glutamina y una mezcla de antibióticos de Estreptomicina Sulfato (100mg/ml) y Penicilina G (10^5 U/ml).

Las células se descongelaron incubando el criotubo que las contenía en un baño a 37 °C con una ligera agitación. Una vez que aproximadamente el 90% del volumen del criotubo estaba descongelado, se depositaron las células en 10 mL del medio de cultivo y se centrifugaban durante 5 minutos a 1500 r.p.m. Tras desechar el sobrenadante con el fin de eliminar el DMSO (dimetilsulfóxido) existente en el medio de congelación, se resuspendieron las células en nuevo medio de cultivo adecuado. A continuación se sembraron sobre la superficie de cultivo que se deseaba o se mantuvieron en suspensión en un incubador a una apropiada temperatura, 37°C, y mezcla de gases (5% CO₂ y 95% O₂).

3. MUESTRAS DE SANGRE DE PACIENTES Y CONTROLES

Las muestras de sangre de donantes sanos fueron suministradas por la Dra. Aurora Viejo del Hospital Universitario la Paz (Madrid). Las muestras de pacientes con Artritis Reumatoide fueron proporcionadas por el Dr. Javier González Polo del Hospital Cantoblanco (Madrid) y las muestras de pacientes con Lupus Eritematoso Sistémico las obtuvimos del Hospital Universitario Ramón y Cajal (Madrid) gracias a la colaboración con el Dr. Antonio Zea Mendoza. Se obtuvo el consentimiento informado de todos los pacientes y los proyectos fueron evaluados por los Comités Éticos correspondientes. Las muestras fueron suministradas con identificación numérica para asegurar el anonimato de los pacientes.

Cada muestra proporcionada disponía de un volumen de 2.5 ml de sangre en tubos cerrados conteniendo EDTA y se mantuvieron a 4°C durante 24 horas hasta que los resultados de los análisis de rutina fueron recibidos. Estos análisis se realizan en el centro de transfusión de la Comunidad Autónoma de Madrid para el estudio de la posible infección con los virus de la Hepatitis B, Hepatitis C y de la Inmunodeficiencia adquirida (VIH). También se realizaba la prueba de Reagina rápida en plasma (RPR) para detectar sífilis. Se utilizaron tubos Leucosep® para una separación óptima de células mononucleares de la sangre humana total. Los tubos Leucosep® se rellenaron con 3 ml de Ficoll-Paque (Sigma) y se centrifugaron durante 30 segundos a 1000 x g. Posteriormente, se añadieron los 2.5 ml de sangre y se centrifugó durante 15 minutos a 1000 x g, quitando el freno de la centrífuga. La fracción celular enriquecida en células mononucleares de sangre periférica (PBMC, *Peripheral Blood Mononuclear Cell*) se extrajo mediante una pipeta Pasteur y se centrifugó 5 minutos a 250 x g. Posteriormente se lavó con PBS (*Phosphate-Buffered Saline*) y este paso se repitió dos veces. A partir del pellet se procedió seguidamente a la extracción de ARN.

4. EXTRACCIÓN, PRECIPITACIÓN Y DETERMINACIÓN DE LA CALIDAD DEL ARN

Se realizó la extracción de ARN de las líneas celulares anteriormente citadas y de las células blancas de las muestras de sangre suministradas mediante el kit SV Total RNA Isolation System (Promega). Este sistema de extracción de ARN combina las propiedades del tiocianato de guanidina (GTC) y el β -mercaptoetanol. El GTC en asociación con SDS (dodecilsulfato sódico), rompe los complejos de nucleoproteína, permitiendo al ARN ser liberado en solución pudiendo ser aislado sin proteína. Después de la centrifugación, para limpiar el lisado de las proteínas, se precipitaba selectivamente el ARN con etanol y se unía a una columna. Se añadía DNAsa I a la membrana para digerir ADN genómico contaminante. El ARN unido se purificaba posteriormente de sales, proteínas e impurezas celulares con simple lavados. Finalmente se eluyó el ARN de la membrana en 100 μ l de agua libre de nucleasas. La concentración del ARN se midió por espectrofotometría en un NanoDrop ND-1000 (NanoDrop) del Servicio de Genómica del CBMSO.

Tras medir la concentración de ARN de las muestras, se procedió a su precipitación para poder disponer de una concentración adecuada, 200-250 $\eta\text{g}/\mu\text{l}$, para la posterior síntesis de ADNc (ADN complementario). Para precipitar el ARN total, se ajustó la concentración monovalente de cationes con Acetato Sódico 3M pH 5 (1/10 Volumen) y añadimos 3 μl de glicógeno (Applied Biosystems) a una concentración de 5 mg/ml. Posteriormente se añadieron dos volúmenes de etanol frío. Dejamos precipitando la mezcla al menos 30 minutos a -70°C o bien a -20°C durante toda la noche. A continuación se centrifugó a 10000 x g durante 15 minutos a 4°C , se removió cuidadosamente el sobrenadante y se resuspendió el pellet en el volumen de agua libre de nucleasas deseado para realizar la retrotranscripción.

La calidad del ARN comprende su pureza (la ausencia de proteínas y contaminación de ADN) y su integridad. Evaluamos la calidad del ARN mediante electroforesis usando el kit RNA 6000 Nano Chip en el Bioanalizador Agilent 2100 (ambos de Agilent) del Servicio de Genómica del CBMSO. Analizando el electroferograma (áreas de las regiones de los ARNs 18S y 28S y altura del pico del ARN 28S) se obtiene un valor conocido como RIN (*RNA Integrity Number*) que se asigna a cada muestra. Las muestras obtienen valores de RIN entre 10 (material intacto) y 0 (material degradado). Se utilizaron aquellas muestras con un valor de RIN superior a 7 para asegurarnos del que ARN que utilizamos no estaba degradado.

5. MUESTRAS DE DIFERENTES ESPECIES

Utilizamos ARN de muestras comerciales (Biochain-ams, <http://www.amsbio.com/>) de diferentes tejidos de varias especies: *Homo Sapiens*, *Macaca mulatta*, *Sus scrofa*, *Bos taurus*, *Mus musculus* y *Rattus norvegicus* (Tabla 1) para el análisis de las formas de splicing de *NFκB1*. La calidad de estos ARNs comerciales fue testada en el Bionalizador (Agilent) corroborándose que su integridad era óptima para los estudios que se iban a realizar posteriormente.

Además, gracias a una colaboración con el Zoo de Madrid mediante su director Jesús Fernández Morán y la veterinaria Eva Martínez Nevado, se nos suministraron muestras de sangre de diferentes especies de primates: *Macaca fascicularis*, *Cebus apella*, *Colobus guereza*, *Pongo pygmaeus*, *Papion sp* y *Gorilla gorilla*. Las muestras

consistían en 2.5 ml de sangre en tubos conteniendo EDTA que mantuvimos a 4 °C durante 24 horas siguiendo el mismo procedimiento que con las muestras de sangre de pacientes y voluntarios sanos. Para la extracción de linfocitos/células periféricas mononucleares, se realizó el método de Ficoll explicado en el apartado anterior.

ESPECIE	TEJIDO	ESPECIE	TEJIDO	ESPECIE	TEJIDO		
<i>Homo Sapiens</i>	NORMAL	<i>Macaca mulatta</i>	CEREBRO	<i>Mus musculus</i>	ADULTO	CEREBRO	
			CEREBRO			CORAZÓN	CORAZÓN
			RIÑÓN			HÍGADO	HÍGADO
			HÍGADO			RIÑÓN	RIÑÓN
			PULMÓN			PULMÓN	PULMÓN
			PÁNCREAS			PÁNCREAS	PÁNCREAS
	FETAL	<i>Sus Scrofa</i>	BAZO	<i>Rattus norvegicus</i>	ADULTO	11 DÍAS	
			TIMO			CEREBRO	15 DÍAS
			CEREBRO			CORAZÓN	17 DÍAS
			CORAZÓN			HÍGADO	CEREBRO
			RIÑÓN			PULMÓN	CORAZÓN
			HÍGADO			RIÑÓN	HÍGADO
TUMOR	<i>Bos taurus</i>	PULMÓN	<i>Rattus norvegicus</i>	ADULTO	RIÑÓN		
		BAZO			CEREBRO	CEREBRO	
		TIMO			CORAZÓN	CORAZÓN	
		RIÑÓN			RIÑÓN	HÍGADO	
		HÍGADO			HÍGADO	RIÑÓN	
		PULMÓN			PULMÓN	PULMÓN	
		BAZO	<i>Rattus norvegicus</i>	EMBRIÓN	19 DÍAS		
		BAZO					

Tabla 3. ARNs de los tejidos de las diferentes especies analizados en este estudio.

6. 5' RLM-RACE

Una vez hecho el análisis de ESTs del gen *NFκBIL1* se procedió a realizar un 5'RLM-RACE (*RNA Ligase Mediated Rapid Amplification of cDNA Ends*, Amplificación rápida de los extremos de ADNc mediada por una ligasa de ARN) para conocer exactamente la secuencia 5' de las distintas isoformas y comprobar la existencia de posibles variantes con distinto 5'UTR. Se realizó a partir de 6 µg de ARN total de K562 con el kit First Choice® RLM-RACE (Applied Biosystems) siguiendo las instrucciones del kit. El ARN se trató con fosfatasa alcalina de intestino de ternera (CIP, *Calf Intestine Alkaline Phosphatase*) para eliminar los fosfatos 5' de moléculas como ARNr, ARNm fragmentado, ARNt y ADN genómico contaminante. La estructura de

caperuza intacta de los extremos 5' del ARNm no se vio afectada por la fosfatasa CIP. El ARN posteriormente se trató con la pirofosfatasa ácida de tabaco (TAP, *Tobacco Acid Pyrophosphatase*) para eliminar la estructura cap de los ARNm de longitud completa, dejando un 5' monofosfato. A continuación, se ligó un adaptador de 45 nucleótidos con la ligasa T4 ARN. Primero mediante una transcripción reversa de oligos al azar, seguida de una reacción en cadena de la polimerasa o PCR (*Polymerase Chain Reaction*) anidada se consiguió amplificar el extremo 5' de un transcrito específico. La PCR anidada consiste en que el producto de una amplificación es utilizado como molde para realizar una segunda amplificación con oligos que se ubican dentro de la primera secuencia amplificada. Las reacciones de PCR contenían 1 µl de ADNc, 0,5 µM de cada oligo exterior (Tabla 4) y 10 µl de la GoTaq® Green Master Mix (Promega) en un volumen de reacción de 20 µl. Las condiciones de PCR fueron las siguientes: 95°C durante 2 minutos seguido de 35 ciclos a 95°C durante 45s, 60°C durante 45s, 72°C durante 1 minuto, y finalmente 72°C durante 5 minutos y se realizó en un termociclador MJ Mini (Bio-RAD). Para la segunda ronda de amplificación se utilizó 1 µl del producto de la primera amplificación con los mismos componentes para la reacción cambiando los oligos exteriores por otros más internos. Las condiciones de PCR fueron las mismas que en el caso anterior. Los oligos correspondientes a la secuencia del adaptador 5'RACE estaban diseñados y se suministraban con el kit (Tabla 4). Se diseñaron dos oligos antisentido anidados para la amplificación específica del gen *NFκBIL1* (Tabla 4).

NOMBRE	SECUENCIA (5'-3')	AMPLICIÓN (pb)
5'RACE exterior	GCTGATGGCGATGAATGAACACTG	670/606
3'NFκBIL1exónIII	TCAAACCTCCCCATGACTTC	
5'RACE interior	CGCGGATCCGAACACTGCGTTTTGCTGGCTTTGATG	388/324
3'NFκBIL1exónII	CAGCATGCAGTGCCGTGTCCCC	

Tabla 4. Oligos utilizados para la amplificación RACE 5'.

7. SÍNTESIS DE ADNc Y ANÁLISIS MEDIANTE PCR

Se retrotranscribió 1 µg de ARN a ADNc usando Oligo-dT con el kit Improm Reverse Transcription System (Promega) en una reacción final de 20 µl siguiendo las instrucciones del protocolo comercial (Tabla 5). Para cada reacción de

retrotranscripción se realizó un control negativo que consistía en la misma mezcla usada para el resto de muestras pero sustituyendo el ARN por agua, a esta reacción la llamábamos RT-.

1ª Desnaturalización (RT1)		
Oligo (dT)15	[Stock] 500 µg/ml	[Final] 0,5µg
ARN	0,5 µg/µl	1 µg
Agua		X
Volumen final		5 µl
Tiempo	Temp °C	
5'	70°C	
5'	4°C	
2ª RT		
Improm II React Buffer	[Stock] 5X	[Final] 1X
dNTPs	10 mM	0,5 mM
MgCl ₂	25 mM	3 mM
Retrotranscriptasa		1µL
Agua		X
Volumen final RT1+RT2		20µl
	Tiempo	T °C
Anillamiento	5'	25°C
Extensión	1h	42°C
Inactivación	15'	70°C
Final		4°C

Tabla 5. Proceso de retrotranscripción del ARN.

Se usó 1 µl de ADNc en cada PCR con los oligos para la amplificación de las isoformas de *NFκBIL1* previamente descritas (Tabla 6 para humanos y Tabla 7 para el resto de especies). En humanos, se diseñaron oligos de acuerdo a las bases de datos analizadas, basándose en las formas descritas (*NFκBIL1* y *NFκBIL1_v1*) y en los análisis de ESTs (*NFκBIL1_v2*). En las especies de *Mus musculus*, *Rattus norvegicus*, *Bos taurus* y *Sus scrofa*, sólo se encontraba descrita la isoforma canónica *NFκBIL1*, por lo que se diseñaron oligos para comprobar primero la existencia o no de un exón 3 con una longitud de 45 pb menos como ocurre en algunas de las isoformas de *NFκBIL1* en humanos y posteriormente se diseñaron oligos para amplificar posibles formas nuevas (Tabla 7). En el caso de los primates, se diseñaron unos oligos degenerados que pudieran amplificar *NFκBIL1* en todas las muestras de primates de las que disponíamos (Tabla 7).

NOMBRE	SECUENCIA (5'-3')	AMPLICIÓN (pb)
Oligos para la amplificación de las isoformas		
NFKCYVAR1 Sentido	GGTCTGATGAGTAACCCCTCCC	1143/1098
NFKVAR2Y3 Sentido	GATGAGGCAGGGGCGGAGACGG	1074/1029
NFKEXÓN4 Antisentido	TCCCTAGGGTCACTTGAGGG	

Tabla 6. Oligos diseñados para la amplificación de las isoformas de *NFκB1* en humanos.

ESPECIE	NOMBRE OLIGO	SECUENCIA (5'-3')	AMPLICIÓN (pb)	
Oligos para la comprobación del tamaño del exón 3				
<i>Mus musculus</i> y <i>Rattus norvegicus</i>	Sentido	Mmexon2-F	GACCCTGCCACCAGGAC	400
	Antisentido	Mmexon4-R	GGATCCCTCTGCCTCCAG	
<i>Sus scrofa</i> y <i>Bos taurus</i>	Sentido	Btexon2-F	GCTGACCCTGCCACCAG	342
	Antisentido	Btexon4-R	TGACCAGGCTGAGAAGGACT	
Oligos para la amplificación de las isoformas				
<i>Mus musculus</i> y <i>Rattus norvegicus</i>	Sentido	MmF	CTCCTCCCTCCCCACCTC	763
	Antisentido	MmR	GGATCCCTCTGCCTCCAG	
<i>Sus scrofa</i> y <i>Bos taurus</i>	Sentido	BtF	CAGGTCTGATGAGTGACCCCTC	688/684
	Antisentido	BtR	TGACCAGGCTGAGAAGGACT	
<i>Macaca fascicularis</i> , <i>Cebus apella</i> , <i>Colobus guereza</i> , <i>Pongo pygmaeus</i> , <i>Papion</i> , <i>Gorilla gorilla</i>	Sentido	PrimF5'	GATGAGGCAGGGGCGGAGACGG	Según especie
	Sentido	PrimF	CCAGGTTCCAGAGGAAGAAG	
	Antisentido	PrimR	TCACTTGAGGGCCTCTGC	

Tabla 7. Oligos diseñados con su secuencia 5'-3' y la longitud del amplicón en nucleótidos para el estudio de *NFKB1* en el resto de especies.

Las reacciones de PCR contenían 1µl de ADNc, 0,5 µM de cada oligo y 10 µl de la GoTaq® Green Master Mix (Promega) en un volumen de reacción de 20 µl. Las condiciones de PCR fueron las siguientes: 95°C durante 2 minutos seguido de 35 ciclos a 95°C durante 45s, 60°C durante 45s, 72°C durante 1 minuto, y finalmente 72°C durante 5 minutos. Los productos obtenidos de las isoformas descritas y de las posibles nuevas variantes, fueron analizados en geles de agarosa al 2% y las diferentes bandas fueron aisladas mediante extracción de gel con el kit Wizard® SV Gel and PCR Clean-Up System (Promega). Tras la purificación se clonaron los fragmentos aislados en el vector pGEM-T Easy (Promega). El clonaje de los fragmentos se realizó incubando la mezcla de ligación O/N a 4°C en las siguientes condiciones: 50 ng pGEM-T Easy Vector (Promega), 3 U T4-Ligasa (Promega) y 10-20 ng de inserto (fragmento aislado)

en un volumen final de 20 μ L. Una vez ligados los fragmentos se transformaron bacterias competentes *E.coli* DH5 α o XL1 Blue seleccionadas por color en placas de LB (Lysogeny Broth) agar con ampicilina (50 μ g/ml), X-Gal (0'6 mg/placa) e IPTG (0'96 mg/placa). La selección de los clones positivos se hizo mediante PCR de colonia. Las condiciones de esta PCR fueron las siguientes: 3 μ l de agua, 0,5 μ M de cada oligo y 5 μ l de la GoTaq® Green Master Mix (Promega) en un volumen de reacción de 10 μ l al que se le añadía una colonia picada con una punta. Esta misma colonia se picaba antes de añadir en la mezcla de PCR, en una nueva placa para poder ponerla de nuevo en cultivo en caso de que fuera positiva. Las condiciones de PCR fueron las siguientes: 95°C durante 2 minutos seguido de 35 ciclos a 95°C durante 45s, 60°C durante 45s, 72°C durante 1 minuto, y finalmente 72°C durante 5 minutos. Las colonias positivas se crecieron en cultivos de 2 mL de LB con ampicilina (50 μ g/mL) O/N a 37°C con agitación.

La extracción de los plásmidos se realizó con el kit Wizard® Plus SV Minipreps DNA Purification System (Promega) que se analizaron por digestión con diferentes enzimas de restricción (Promega, Roche y NEB) en función de las dianas del inserto. Se secuenciaron los plásmidos con el inserto correspondiente clonado con los oligonucleótidos T7 (5' TAATACGACTCACTATAGG3') y el reverso SP6 (5' TAGGTGACACTATAGAATAC3') en el servicio de Secuenciación del Instituto de Investigaciones Biomédicas "Alberto Sols" (Madrid).

8. RT-PCR EN TIEMPO REAL

La PCR cuantitativa en tiempo real es un método para la amplificación de ADN en el que se usan fluoróforos para detectar la cantidad de producto de PCR después de cada ciclo, al contrario de los ensayos a tiempo final donde el análisis se hace después de un número fijo de ciclos (Higuchi et al., 1993). La principal ventaja de la PCR a tiempo real es que te permite determinar el número de copias del material de partida con precisión y alta sensibilidad. Esta técnica es una herramienta muy efectiva para el análisis de la expresión génica siendo el método más sensible para la detección y cuantificación de transcritos de baja expresión en muestras biológicas. Además de la

sensibilidad, mediante la PCR a tiempo real se pueden discriminar dos secuencias de ARNm casi idénticas (Wong and Medrano, 2005).

Para realizar PCR cuantitativa se utiliza un termociclador acoplado a un detector de fluorescencia y a un ordenador. El equipo está dotado con filtros de excitación y filtros de emisión, de tal modo que un haz de luz atraviesa el filtro de excitación seleccionado, llega a la muestra y provoca la emisión de fluorescencia correspondiente a una determinada longitud de onda. La luz emitida es captada por el filtro de emisión seleccionado y es recogida por un detector. La longitud de onda de la fluorescencia emitida depende del fluoróforo que es añadido a la muestra. El fluoróforo utilizado en nuestros experimentos fue SYBR Green-I, que es capaz de intercalarse en el ADN y de emitir fluorescencia (520 nm) al excitarse con la luz adecuada (excitación a 497 nm). El SYBR Green-I emite fluorescencia muy baja en solución, pero se produce un aumento de ésta cuando este fluoróforo se intercala en el ADN de doble cadena a medida que avanzan los ciclos de PCR. La fluorescencia captada es analizada por el “software” específico como unidades relativas de fluorescencia (RFU, *Relative Fluorescent Units*). El equipo comienza a captar fluorescencia antes del primer ciclo. Esta fluorescencia es la denominada *fluorescencia base*. En cada ciclo siguiente el equipo recoge nuevamente valores de fluorescencia y los compara con los del ciclo anterior. La fluorescencia debida al SYBR Green-I que se intercala en el ADN durante la amplificación es muy baja durante los primeros ciclos. Llegado un ciclo determinado en el que dicha fluorescencia empieza a ser superior a la fluorescencia base, debido a la amplificación continua ciclo tras ciclo. Este ciclo en el cual la fluorescencia de la muestra comienza a ser mayor que la fluorescencia base se denomina Ciclo Umbral o C_t (threshold cycle). El valor de C_t es el valor clave para los experimentos de PCR en tiempo real. C_t es inversamente proporcional a la cantidad inicial de moléculas diana en nuestra muestra, de manera que cuanto mayor sea nuestro número inicial de moléculas diana más bajo será nuestro C_t es decir la cantidad de ADN amplificado necesario para superar la fluorescencia base aparecerá antes. Aquellos casos en los que nunca se supere la fluorescencia base, es decir en los que no haya amplificación, se corresponden con ausencia de moléculas diana iniciales en las muestras (Figura 11).

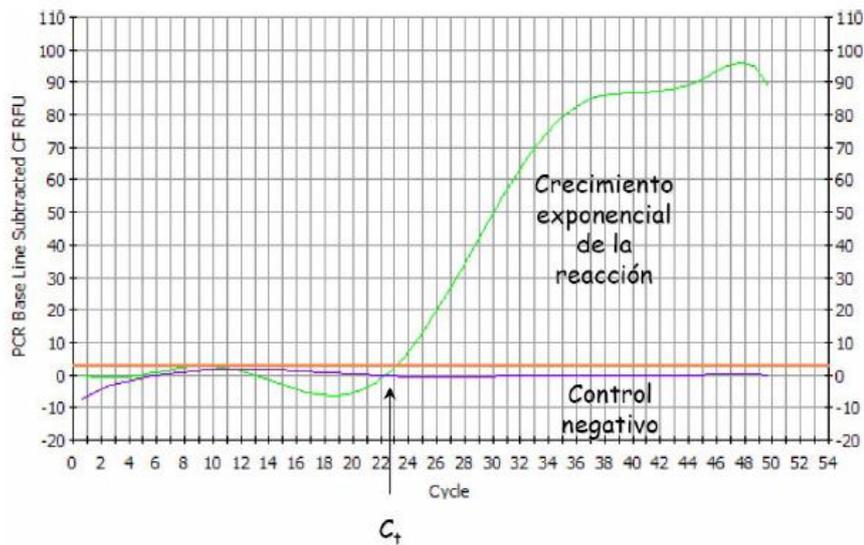


Figura 11. Diagrama en el que se representa la emisión de fluorescencia con respecto al ciclo de reacción. Esta emisión de fluorescencia irá en función de la cantidad de producto de reacción que se obtiene en cada ciclo. En verde una muestra control con ADNc y en morado un control negativo de agua. Como podemos observar, la curva del control negativo no llega a emitir fluorescencia. C_t , Ciclo Umbral.

La temperatura de melting se define como la temperatura a la cual el 50 % de una molécula de ADN de doble cadena se encuentra desnaturalizada. Este paso de desnaturalización o de melting hace que las moléculas de doble cadena de ADN, correspondientes al producto de amplificación, se abran y formen hebras simples de ADN, lo que se traduce en una pérdida de la emisión de fluorescencia, ya que el SYBR Green se separa de ellas. Esto se consigue mediante el aumento progresivo de la temperatura con incrementos de $0,5^{\circ}\text{C}$ desde 55°C a 95°C y midiendo los valores de fluorescencia tras cada incremento.

Esta pérdida de fluorescencia, al ser representada gráficamente como la derivada del valor de fluorescencia con respecto a la temperatura ($d\text{RFU}/dT$) en el eje Y, frente a la temperatura en el eje X, se traduce en una *curva o pico de "melting"* (Figura 12). Cada pico de melting es específico de cada producto de reacción, ya que depende su composición en nucleótidos (A, G, T, C) y de su tamaño, en pares de bases. La generación de este pico de melting es específico para cada producto de PCR y permite distinguir nuestro producto de PCR de posibles contaminaciones, bandas inespecíficas, o ruido de fondo provocado por la formación de dímeros de cebadores.

8.1 Curvas patrón

Para calcular los niveles de expresión de las diferentes isoformas de *NFκBIL1* en la misma muestra, es necesario realizar una cuantificación absoluta de las isoformas en cada muestra. Para conseguirlo, generamos curvas patrón estándares.

Se diseñaron oligos específicos que amplificaran cada isoforma para la realización del análisis de PCR cuantitativa (Tabla 8). Se realizaron ensayos cruzados de PCR a tiempo real con cada plásmido estándar de cada isoforma clonada en PGEM-T enfrentado a cada pareja de oligos para comprobar que sólo amplificaban la isoforma correspondiente.

El número de copias de cada plásmido se calculó usando el número de Avogadro y el tamaño de los plásmidos. Para generar las curvas patrón, se hicieron diluciones seriadas 1:10 de los plásmidos estándar desde 10^6 a 10^0 copias. Estas curvas permiten determinar las concentraciones de muestras desconocidas basados en sus valores Ct.

NOMBRE		SECUENCIA (5'-3')	AMPLICIÓN (pb)
Oligos específicos de cada isoforma (PCR cuantitativa)			
<i>NFKBIL1</i>	Sentido	AAGCCTCCACATCTGTCTG	519
	Antisentido	TCAAACCTCCCCATGACTTC	
NFKBIL1_v1	Sentido	AAGCCTCCACATCTGTCTG	488
	Antisentido	GGAGGCATCACCTGG	
<i>NFKBIL1_v2</i>	Sentido	TGTGAATGCGGTGAAAAGC	514
	Antisentido	TCAAACCTCCCCATGACTTC	
<i>NFKBIL1_v3</i>	Sentido	TGTGAATGCGGTGAAAAGC	483
	Antisentido	GGAGGCATCACCTGG	
Oligos para amplificar genes controles (PCR cuantitativa)			
GAPDH	Sentido	AGCCACATCGCTCAGACAC	67
	Antisentido	CGCCAATACGACCAAT	
β-actina	Sentido	ATTGGCAATGAGCGGTTT	76
	Antisentido	GGATGCCACAGGACTCCAT	

Tabla 8. Oligos diseñados para los ensayos de PCR a tiempo real.

8.2 Ensayos de PCR a tiempo real

Los ensayos se realizaron en el termociclador de tiempo real ABI Prism 7900HT (Applied Biosystems) del Servicio de Genómica del CBMSO y cada mezcla de reacción contenía 1x Power SYBR® Green Master Mix (Applied Biosystems) y 250 nM de cada oligo. Para la cuantificación, se utilizaron 50 ng de ADNc (el equivalente a ARN total) y las condiciones del termociclador fueron: 95°C durante 15 minutos, 40 ciclos de 95°C 15 segundos y 60°C durante 1 minuto. Se analizaron los correspondientes controles negativos de la RT (RT-), consistentes en sustituir el ARN por agua en el proceso de retrotranscripción de las muestras experimentales para detectar una posible contaminación con genómico así como muestras sin ADNc para controlar nuestros reactivos. También se incluyó una curva de melting desde los 60°C a 95°C y una incubación a 95 °C durante 15 segundos al final del programa para verificar la especificidad de la PCR. La fluorescencia se adquirió en el minuto de extensión a 60°C de cada uno de los 40 ciclos y en la curva de melting (transición 60-95 °C e incubación final de 15 segundos a 95 °C).

8.3 Análisis de datos

Para las cuantificaciones absolutas, el número de moléculas de cada transcrito se determinó interpolando los valores de Ct de las muestras desconocidas en las curvas estándar anteriormente realizadas y el análisis se realizó con el Software SDS 2.2.1 (Applied Biosystems).

Las cuantificaciones relativas fueron siempre normalizadas usando controles endógenos como *GAPDH* y *β -actina*. Los valores de Ct se exportaron en formato .txt de tabla para posteriores análisis con el software GenEx (MultiD, Suecia) que nos permitió realizar los siguientes pasos: corrección de la eficiencia de la PCR, media de los tres replicados de PCR, normalización de los genes usando los genes control en todas las muestras (como determina el algoritmo Normfinder) y la cuantificación relativa en los casos necesarios. Los análisis estadísticos, T-test, se realizaron con este Software para la comparación entre grupos y los p-valores < 0.01 se consideraron estadísticamente significativos.

9. ANÁLISIS DE LA REGIÓN PROMOTORA DE *NFκBIL1*

9.1 Extracción de ADN genómico

El ADN genómico se extrajo de las muestras de sangre a partir de un volumen inicial de 0.5 ml (FlexiGene DNA kit, Qiagen). Primero se añadió buffer de lisis a la muestra y se centrifugó para precipitar los núcleos y las mitocondrias. Posteriormente, se incubó el pellet en buffer de desnaturalización con una proteasa para remover eficientemente contaminantes como las proteínas. El ADN se precipitó añadiendo isopropanol, recuperado después de centrifugar, lavado con etanol al 70%, secado y resuspendido en agua libre de nucleasas. La concentración del ADN se midió por espectrofotometría en un NanoDrop ND-1000 (NanoDrop). La región promotora de *NFκBIL1* se analizó en muestras de pacientes con AR y en muestras de controles. Estos análisis se realizaron en colaboración con el Profesor Stephan Beck en el Sanger Institute (Hinxton, Cambridge, UK).

9.2 Análisis de mutaciones y SNPs

Se diseñaron oligos específicos para la amplificación del promotor de *NFκBIL1* (Tabla 9). El promotor fue amplificado en todas las muestras a 95°C durante dos minutos seguido de 35 ciclos de 95 °C durante 45 segundos, 60°C durante 45 segundos, 72°C durante un minuto, y finalmente 72°C durante 5 minutos. Los productos fueron purificados, secuenciados y las secuencias se alinearon con el programa MultAlin y analizadas para comprobar la existencia de mutaciones y SNPs.

9.3 Análisis de metilación en la región promotora

La metilación del ADN se caracteriza por la adición bioquímica de un grupo metilo (-CH₃) a la citosina C₅ en los dinucleótidos CpG a través de la enzima metiltransferasa. En el genoma de mamíferos existen segmentos abundantes con dinucleótidos CpG, las llamadas islas CpG, las cuales corresponden frecuentemente a las zonas promotoras y los exones adyacentes. Cambios en la metilación de estas islas

pueden ocurrir y se han asociado a cambios en la transcripción y expresión génica. Para determinar los patrones específicos de metilación en una localización única, como en una región promotora, el método de conversión con bisulfito se ha convertido en una de las herramientas favoritas. Con este método, el ADN se desnaturaliza y se trata con bisulfito sódico, causando que las citosinas no metiladas se conviertan en uracilos mientras que las citosinas metiladas permanecen sin cambios (Figura 13). El bisulfito reduce el anillo aromático de una citosina no metilada, dando lugar a un producto intermedio que se desamina espontáneamente. La adición de la base NaOH causa desulfonación del anillo, que finalmente se convierte en un anillo de uracilo. El ADN convertido puede amplificarse mediante PCR y analizarse mediante secuenciación.

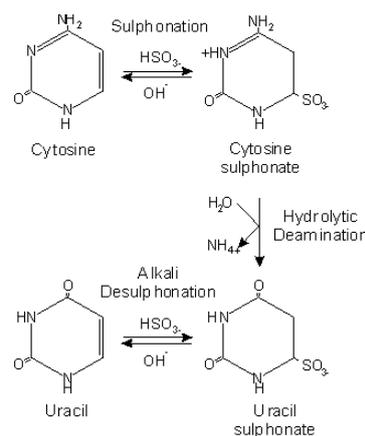


Figura 13. Tratamiento del ADN con bisulfito.

Para nuestro estudio, se diseñaron los oligos para la secuenciación con bisulfito de la región promotora de *NFκB1* (Tabla 9). En el diseño de estos oligos particulares hay que tener unas consideraciones especiales, como son que en su secuencia no existan sitios CpGs y que contengan un mínimo de citosinas que no se encuentren formando dinucleótidos CpGs para amplificar sólo el ADN tratado con bisulfito. Se recomienda además diseñar oligos que reconozcan fragmentos pequeños (300-400 pb) y que contengan un contenido en GC del 30 al 40%.

La conversión de bisulfito del ADN de todas las muestras fue realizada utilizando el kit EZ DNA Methylation KitTM (Zymo Research). Después de la conversión, la zona promotora fue amplificada en todas las muestras y secuenciada para estudiar los patrones de metilación. Las PCRs se realizaron en placas de 96 pocillos en un volumen final de 25 μl conteniendo 250 μM de dNTPs, 1x de Buffer para la PCR (Qiagen), 10 μmol de cada oligo, 1 U de Taq polimerasa (Qiagen) y 8 ηg de ADN genómico tratado con bisulfito. Las condiciones de PCR fueron las siguientes: 95°C durante 15 minutos seguido de 40 ciclos de 95°C durante 1 minuto, 55°C durante 45 segundos, 72°C durante 90 segundos y una extensión final de 10 minutos a 72 °C. Los amplicones de PCR fueron purificados usando placas de PCR MultiScreen (Millipore,

Billerica, Massachusetts, United States) y secuenciados directamente con los mismos oligos usados en la reacción de PCR.

NOMBRE		SECUENCIA (5'-3')	AMPLICIÓN (pb)
Oligos para la amplificación del promotor			
Promotor	Sentido	TGGTGGTAACAGAGGAAGCA	833
	Antisentido	TATCATGTACCCGGCAGACA	
Oligos para el análisis de la metilación del promotor			
Bis-A	Sentido	GGGAGGGGTTATTTATTAGATT	380
	Antisentido	ACCTCAAATAACTCTCTTTTCTC	
Bis-B	Sentido	GGGAGAAAAGAGAGTTATTTGGA	278
	Antisentido	CAAATCCAAATTTTACATCTCA	
Bis-C	Sentido	TTTGAATGTTGTTAAAGTATTAGATGG	427
	Antisentido	TCTTTTCCCAACACAAATTACA	

Tabla 9. Oligos diseñados para los análisis de la región promotora del gen *NFκBIL1*.

10. SOBRE-EXPRESIÓN DE LAS ISOFORMAS DE *NFκBIL1*

10.1 Construcciones para la expresión de cada isoforma

10.1.1 Epítomos en el extremo carboxilo terminal

Las diferentes isoformas humanas de *NFκBIL1* se clonaron en un vector pcDNA3.1/V5-His-TOPO (Invitrogen) para su posterior expresión en células eucariotas (Figura 14). Para clonar la secuencia codificante de las diferentes isoformas de *NFκBIL1*, exceptuando el codón stop, con el fin de continuar la fase con los epítomos V5 e Histidinas (Tabla 10), se utilizó un oligonucleótido directo, portador de una diana *HindIII* y como oligonucleótido reverso se utilizó uno que portaba en su extremo 3' una diana para *XbaI* que sustituía el codón stop por una alanina (Tabla 11). Las reacciones de PCR contenían 50 ng de ADNc de tejido cerebral adulto, 0,5 μM de cada oligo y 10 μl de la GoTaq® Green Master Mix (Promega) en un volumen de reacción de 20 μl. Las condiciones de PCR fueron las siguientes: 95°C durante 2 minutos seguido de 35 ciclos

a 95°C durante 45s, 60°C durante 45s, 72°C durante 1 minuto, y finalmente 72°C durante 5 minutos. Los productos de PCR de las isoformas de *NFκBIL1* se digirieron con *HindIII* y *XbaI* (Promega). Se ligaron los fragmentos de la digestión en el vector pcDNA3.1/V5-His-TOPO digerido con los mismos enzimas de restricción. Con todos los plásmidos construidos se transformaron bacterias competentes *E.coli* TG1 que se sembraron en placas de LB agar con ampicilina (50 µg/mL). Las colonias obtenidas se crecieron en 2 mL de LB con ampicilina (50 µg/mL) O/N a 37°C con agitación. La extracción de los plásmidos se realizó con el kit Wizard® Plus SV Minipreps DNA Purification System (Promega) que se analizaron por digestión con enzimas de restricción. Se secuenciaron los plásmidos bidireccionalmente para comprobar la fidelidad nucleotídica antes de realizar la expresión de las isoformas. Se seleccionaron 4 plásmidos, cada uno conteniendo una de las isoformas para realizar las transfecciones.

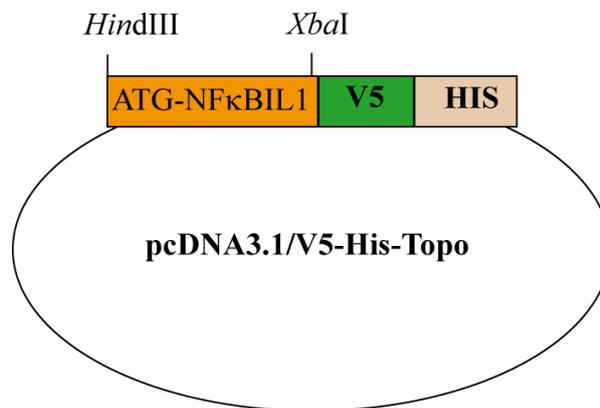


Figura 14. Construcción de cada isoforma de NFκBIL1 con epítomos V5 e Histidinas en el extremo carboxilo terminal.

10.1.2 Epítomos en el extremo amino y carboxilo terminal.

El epítomo Flag (Tabla 10) se clonó en el plásmido pcDNA3.1/Zeo (-) (Invitrogen) en el extremo amino terminal. Para ello se diseñaron unos oligos de 54 nucleótidos conteniendo la diana *NheI*, el codón de iniciación, el epítomo Flag, la diana *EcoRV*, un codón stop y la diana *XhoI* (Tabla 11). Se anillaron los oligos sentido y antisentido mediante el siguiente protocolo: 4 minutos a 100°C, un minuto a 100°C seguidos de 99 ciclos con una caída de 0,5°C en cada ciclo, 1 minuto a 50,5°C seguidos de 90 ciclos con una caída de 0,5°C en cada ciclo. La construcción obtenida se clonó en las dianas *NheI* y *XhoI* del plásmido citado anteriormente. La isoforma canónica

NFκBIL1, junto con el epítipo V5, se amplificó mediante PCR de la construcción del apartado anterior con oligos portadores de dianas *EcoRV* y *HindIII* (Tabla 11) y se insertó entre las dianas *EcoRV* y *HindIII* del plásmido clonado con el epítipo Flag. De esta manera, se obtuvo una construcción con epítipos en los extremos amino y carboxilo terminal (Figura 15). A este nuevo plásmido con un epítipo Flag en el extremo amino y un epítipo V5 en el extremo carboxilo lo llamamos Flag-pcDNA3.1-V5.

Con el plásmido construido se transformaron bacterias competentes *E.coli* TG1 que se sembraron en placas de LB agar con ampicilina (50 µg/mL). Las colonias obtenidas se crecieron en 2 mL de LB con ampicilina (50 µg/mL) O/N a 37°C con agitación. La extracción de los plásmidos se realizó con el kit Wizard® Plus SV Minipreps DNA Purification System (Promega) que se analizaron por digestión con enzimas de restricción y se secuenciaron para comprobar la secuencia.

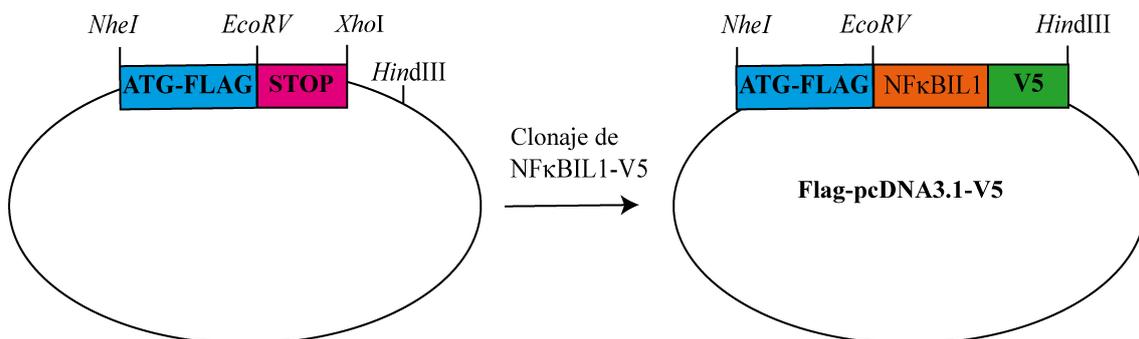


Figura 15. Construcción de la isoforma canónica NFκBIL1 con epítipos en el extremo amino y carboxilo terminal. En el extremo amino terminal se encuentra el epítipo Flag y en el extremo carboxilo terminal el epítipo V5.

Epítipo	Secuencia
HIS	HHHHHH
V5	GKPIPPELLGLDST
FLAG	DYKDDDDK

Tabla 10. Epítipos y secuencias utilizados en nuestro estudio.

NOMBRE		SECUENCIA (5'-3')	AMPLICIÓN (pb)
Oligos con dianas <i>HindIII</i> y <i>XbaI</i> para amplificación CDS			
FCyv1 <i>HindIII</i>	Sentido	CCCAAGCTTATGAGTAACCCCTCCC	1143/1098
Fv2y3 <i>HindIII</i>	Sentido	CCCAAGCTTATGGCCTCCACTTCCC	1074/1029
RCo <i>XbaI</i>	Antisentido	GGGTCTAGACTTGAGGGCCTCTGCA	
Oligos con epítipo Flag y dianas <i>NheI</i>, <i>EcoRV</i> y <i>XhoI</i>			
FlagF	Sentido	GGGGCTAGCATGGATTACAAGGATGAC GACGATAAGGATATCTGACTCGAGGGG	54
FlagR	Antisentido	CCCCTCGAGTCAGATATCCTTATCGTC GTCATCCTTGTAATCCATGCTAGCCCC	
Oligos con dianas <i>EcoRV</i> y <i>Hind III</i> para clonar <i>NFKBIL1</i> junto con V5			
<i>EcoRV</i> NFKBIL1	Sentido	GGGGATATCATGAGTAACCCCTCCCCC	1227
V5 <i>HindIII</i>	Antisentido	GGGAAGCTTTCACGTAGAATCGAGACC	

Tabla 11. Oligos utilizados para el clonaje de las isoformas del gen *NFKBIL1* en vectores de expresión.

10.2 Transfecciones y análisis por Western Blot

Se sembraron células COS-7 en placas de 24 pocillos hasta el 50-70% de confluencia antes de transfectarlas con los plásmidos construidos. En cada pocillo se sembraron 2.5×10^4 células y se transfectaron con 0.5 μg de ADN utilizando el Kit TransIT® COS Transfection (Mirus-BioNova) para la transfección. Se recogieron las células 48 horas post-transfección y para lisarlas se usó tampón de carga para proteínas (60 mM Tris-HCl pH=6.8, 2% SDS, 10% glicerol, 5% β -mercaptoetanol y 0,01% bromofenol). Las muestras fueron hervidas durante 5 minutos antes de la resolución por electroforesis en gel.

Los extractos celulares obtenidos se cargaron en geles de acrilamida al 12% siguiendo el método SDS-PAGE. Los geles obtenidos se transfirieron a membranas de nitrocelulosa (PROTRAN) mediante transferencia húmeda en el Mini Trans-Blot® Electrophoretic Transfer Cell (Bio-rad). Las membranas obtenidas se procesaron utilizando el siguiente protocolo: bloqueo toda la noche a 4°C en PBS con 0,05% Tween 20 (Sigma) y 5% de leche en polvo. Tras el bloqueo, se incubó con el anticuerpo primario (Tabla 12) diluido en PBS con 0,05% Tween 20 y 5% leche en polvo durante 1 hora tras lo que se lavó la membrana 3 veces durante 10 minutos en PBS con 0,05%

Tween 20. Posteriormente, se incubó con el anticuerpo secundario correspondiente (Tabla 12) diluido en PBS con 0,05% Tween 20 durante 1 hora. Después se lavó 2 veces durante 10 minutos en PBS con 0,05% Tween 20 y 2 veces durante 10 minutos en PBS. Finalmente revelado con el kit ECL Plus Western Blotting (Amershan).

ANTICUERPO PRIMARIO	REFERENCIA	DILUCIÓN	ANTICUERPO SECUNDARIO	REFERENCIA	DILUCIÓN
Anti V5 Antibody Mouse monoclonal	SIGMA	1:10000	Anti Mouse IgG- Peroxidase	SIGMA	1:5000
Anti Flag Antibody mouse monoclonal	SIGMA	1:1000	Anti Mouse IgG- Peroxidase	SIGMA	1:5000
Anti NFκBIL1 Antibody Rabbit polyclonal	SIGMA	1:500	Anti Rabbit IgG- Peroxidase	PROMEGA	1:2500

Tabla 12. Anticuerpos primarios y secundarios con sus diluciones y referencias utilizados para los análisis por Western Blot.

10.3 Inmunofluorescencias

Se cultivaron las células COS-7 sobre cubreobjetos circulares de vidrio depositados sobre placas de 24 pocillos hasta una confluencia del 60% y se transfectaron siguiendo el protocolo anteriormente descrito. Dos días después de la transfección, se lavaron los cristales con PBS 1x dos veces para retirar los restos de medio. A continuación, se empezó la fijación de las células con paraformaldehído (PFA) al 2% durante 10 minutos a temperatura ambiente. A continuación, añadimos 500 μL de PFA al 4% y fijamos durante 20 minutos a temperatura ambiente. Retiramos y realizamos 3 lavados con PBS 1x. Incubamos con cloruro de amonio 50mM durante 30 minutos a temperatura ambiente. Permeabilizamos con PBS 0,1% tritón X-100 durante 20 minutos a temperatura ambiente. Bloqueamos con PBS 0,1% Tritón X-100 y BSA (albúmina de suero bovino) al 5% durante 20 minutos. Incubamos con el anticuerpo primario (Tabla 13) en solución de bloqueo durante una hora seguido de lavados con PBS 0,1% tritón X-100 durante una hora aproximada. A continuación incubamos con el

anticuerpo secundario (Tabla 13) correspondiente en solución de bloqueo una hora y después los lavados con PBS 0,1% tritón X-100. Incubamos con To-Pro 3 (1:500) en PBS 0,1% tritón X-100 durante 2 minutos, lavamos y montamos los cristales con prolong (Invitrogen).

ANTICUERPO PRIMARIO	REFERENCIA	DILUCIÓN	ANTICUERPO SECUNDARIO	REFERENCIA	DILUCIÓN
Anti V5 Antibody Mouse monoclonal IgG	SIGMA	1:400	Anti Mouse IgG Alexa 488	INVITROGEN	1:500
Anti NFκBIL1 Antibody Rabbit polyclonal IgG	SIGMA	1:15	Anti Rabbit IgG Alexa 555	INVITROGEN	1:500

Tabla 13. Anticuerpos primarios y secundarios utilizados en las inmunofluorescencias.

Los experimentos de co-localización se realizaron con el mismo protocolo de inmunofluorescencia explicado anteriormente. Los cristales se incubaron con una mezcla de anticuerpo antiV5 policlonal de conejo y un anticuerpo monoclonal anti SC-35 de ratón que reconoce las espículas nucleares a las diluciones adecuadas (Tabla 14) en solución de bloqueo. Después de los lavados, se incubaron las muestras con una mezcla de los anticuerpos secundarios correspondientes (Tabla 14) en solución de bloqueo. Se utilizó To-Pro-3 para la tinción de ácidos nucleicos. Se montaron los cristales para su análisis por microscopía confocal.

ANTICUERPO PRIMARIO	REFERENCIA	DILUCIÓN	ANTICUERPO SECUNDARIO	REFERENCIA	DILUCIÓN
Anti V5 Antibody Rabbit polyclonal	ABCAM	1:1000	Anti Rabbit IgG Alexa 555Antibody	INVITROGEN	1:500
SC-35 Antibody Mouse monoclonal	SIGMA	1:500	Anti Mouse IgG Alexa 488 Antibody	INVITROGEN	1:500
To-Pro 3	INVITROGEN	1:500			

Tabla 14. Anticuerpos primarios y secundarios utilizados en los ensayos de co-localización.

10.4 Microscopía confocal y análisis de imágenes

Las muestras fueron analizadas en el servicio de Microscopía Confocal del CBMSO, utilizando un Microscopio de Barrido Láser confocal LSM510 acoplado a un microscopio vertical Axio Imager.Z1 M (Zeiss). La configuración elegida para el análisis de las muestras correspondió con las longitudes de onda de excitación/emisión que se indican en la tabla 15.

Fluoróforo	Longitud onda excitación	Longitud onda emisión
Alexa 488	495nm	519nm
Alexa 555	560nm	576nm
To-Pro-3	642nm	661nm

Tabla 15. Fluoróforos utilizados en las técnicas de inmunofluorescencia. Se representa la longitud de onda de absorción y emisión para Alexa 488 y Alexa 555, conjugados a nuestros anticuerpos secundarios, y para To-Pro-3, utilizado para el marcaje de núcleos.

Esta configuración nos permitió detectar el marcaje específico de nuestro anticuerpo secundario a una longitud de onda que no interfirió con la emitida por el marcaje nuclear To-Pro-3. El fotomultiplicador de cada canal se ajustó de forma diferencial para tomar la imagen con la mayor calidad posible y posteriormente se llevó a cabo la superposición de los diferentes canales con los programas adecuados. Se utilizaron los objetivos de 40x, 63x y 100x aumentos. Las imágenes obtenidas con el microscopio confocal fueron procesadas con el programa Zeiss LSM Image Browser.

1. CARACTERIZACIÓN DE LAS ISOFORMAS DE *NFκBIL1*

El gen *NFκBIL1* ó *IκBL* (GeneID: 4795) se localiza en el cromosoma 6p21.3 dentro de la región de clase III del MHC entre los genes *LTα* (*lymphotoxin alpha*) y *ATP6v1G2* (*ATPase, H⁺ transporting, lysosomal 13kDa, VI subunit G2*). *NFκBIL1* se transcribe en la misma orientación que *LTα* pero en orientación contraria a *ATP6v1G2*. Se encuentra a 13'2 Kb del gen *LTα* y a tan sólo 30 pb del gen *ATP6v1G2*. La proteína que codifica *LTα* es una citoquina producida por linfocitos y es miembro de la familia de TNF. La proteína es altamente inducible, secretada y forma heterotrímeros con la linfotoxina beta, que ancla la linfotoxina alfa a la superficie celular. La proteína codificada por *LTα* media una gran variedad de respuestas inflamatorias, inmunoestimuladoras y antivirales. Está implicada en el desarrollo de los órganos linfoides secundarios durante el desarrollo y se asocia a la artritis psoriática. Se han descrito varias formas de splicing de este gen. *ATP6v1G2* codifica un componente de la ATPasa vacuolar (V-ATPase), una enzima de múltiples subunidades que media la acidificación de los compartimentos intracelulares de las células eucarióticas. La acidificación dependiente de esta ATPasa es necesaria para la endocitosis mediada por receptor. Se han descrito varias formas de splicing de *ATP6v1G2*.

1.1 En *Homo Sapiens*

Nuestro primer objetivo fue determinar de qué manera *NFκBIL1* se expresaba. En particular, nos interesaba conocer las distintas variantes de splicing de este gen y de qué manera podía afectar este proceso a la expresión de sus proteínas. Para determinar las formas de splicing descritas de *NFκBIL1* e identificar nuevas variantes, se realizaron búsquedas en las base de datos anteriormente citadas.

En Ensembl se encontraron dos transcritos de *NFκBIL1* al igual que en la base de datos del NCBI. Las dos isoformas descritas previamente eran *NFκBIL1-α* (ó *IκBL*) ([NM_005007.3](#)) y *NFκBIL1-β* (ó *IκBLsv1*) ([NM_001144961.1](#)). Ambas isoformas constaban de 4 exones, I, II, III y IV. El mensajero de la isoforma *NFκBIL1-α* tenía una

longitud de 1478 pb y la secuencia codificante daba lugar a una proteína de 381 aa. El mensajero de la isoforma *NFκBIL1-β* presentaba una longitud de 1401 pb y la proteína que se codificaba era de 366 aminoácidos. Lo que las diferenciaba es que la isoforma *NFκBIL1-β* carecía de 45 nucleótidos en la parte final del exón III. El nuevo sitio de splicing en el exón III seguía la regla GT para los sitios de splicing 5' y la pauta de lectura del exón IV se mantenía. Por lo tanto, las dos variantes sólo se diferenciaban en la ausencia de la secuencia aminoacídica GELEDEWQEVMGRFE en la proteína NFκBIL1-β (Ver Figura 10).

Aparte de las isoformas *NFκBIL1-α* y *NFκBIL1-β* previamente descritas, se encontraron ESTs ([AU143330.1](#), BP195124 y BM700433) en la base de datos del NCBI y Ensembl correspondientes a otra isoforma con diferente extremo 5'UTR. Esta variante comenzaría por un exón I, al que llamaremos exón IA de ahora en adelante (Figura 16), diferente del descrito para las dos isoformas previamente caracterizadas, al que designamos como exón IB. Este exón IA, de 104 pb de longitud, se encontraba a 607 pb del exón IB y a una distancia de 1189 pb del exón II común a todas las variantes. La nueva variante constaba también de 4 exones y la longitud del mensajero era de 1411 pb.

Para comprobar experimentalmente la existencia de un mensajero con un 5'UTR diferente al descrito para las dos isoformas de *NFκBIL1* previamente caracterizadas y la posibilidad de que existieran nuevas variantes con extremos 5'UTR distintos se procedió a la realización de un experimento de 5'RLM-RACE a partir de ARN de la línea celular de K562. Siguiendo este procedimiento, los ARNm completos quedaban ligados a un adaptador en su extremo 5'. Con oligos diseñados en ese adaptador y oligos antisentido diseñados en el gen *NFκBIL1* se realizó una PCR anidada cuyos productos se secuenciaron posteriormente (Figura 16).

Tras las dos rondas de la PCR anidada se obtuvieron dos bandas que tras secuenciarlas, se demostró la existencia de dos extremos 5'UTR distintos, el que estaba descrito y caracterizado ya para las isoformas *NFκBIL1-α* y *NFκBIL1-β* y el nuevo descrito para la variante encontrada en los análisis de ESTs.

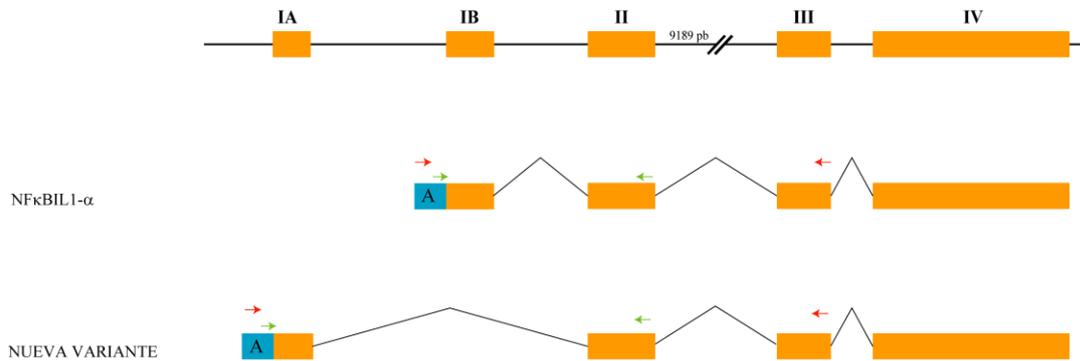


Figura 16. Oligos diseñados para la amplificación de los extremos 5'UTR de las variantes de *NFκBIL1*. Los oligos de color rojo se utilizaron en la primera ronda de la PCR anidada y los de color verde en la segunda ronda. La A simboliza el adaptador ligado en los extremos 5' de los mensajeros. Los rectángulos naranjas representan los exones. La primera variante representada es la variante α , previamente descrita y debajo se encuentra la nueva variante encontrada en las bases de datos de ESTs. Los exones están indicados con números romanos.

Se comprobó la expresión de esta nueva isoforma en líneas celulares y tejidos mediante RT-PCRs con oligos específicos en el exón IA y en el exón IV. Se consiguió amplificar esta isoforma en todas las muestras estudiadas. Analizando la pauta de lectura con el programa ORFinder se observó que el codón de inicio de la traducción se encontraba en el exón II y la secuencia codificante daría lugar a una proteína de 358 aminoácidos.

Analizando los resultados, se pensó en la posibilidad de que hubiera otra variante que tuviera el extremo 5'UTR como la nueva isoforma amplificada y el exón III más corto como ocurría en la isoforma *NFκBIL1-β* previamente descrita. Diseñamos oligos específicos para esta posible nueva isoforma en el extremo 5'UTR y el oligo antisentido solapando el exón III y IV y conseguimos identificar una nueva variante de *NFκBIL1*. Esta variante de 4 exones también, presentaba una longitud del mensajero de 1366 pb y la secuencia codificante daría lugar a una proteína de 343 aminoácidos. Para confirmar la expresión de estos transcritos, se llevaron a cabo análisis de RT-PCR en distintas líneas celulares usando los oligos descritos en la tabla I. Los productos de PCR fueron secuenciados para confirmar la especificidad nucleotídica.

Al descubrir nuevas variantes de splicing del gen *NFκBIL1* sugerimos renombrar las distintas variantes. Siguiendo las instrucciones del Comité de Nomenclatura de Genes HUGO (HGNC), la isoforma nombrada anteriormente α sería la isoforma canónica *NFκBIL1* y la primera variante de splicing descrita, llamada β previamente, sería *NFκBIL1_v1*. La nueva isoforma que se amplificó y cuyas ESTs encontramos en las bases de datos correspondería a *NFκBIL1_v2* y *NFκBIL1_v3* correspondería a la nueva isoforma que identificamos y amplificamos (Figura 17).

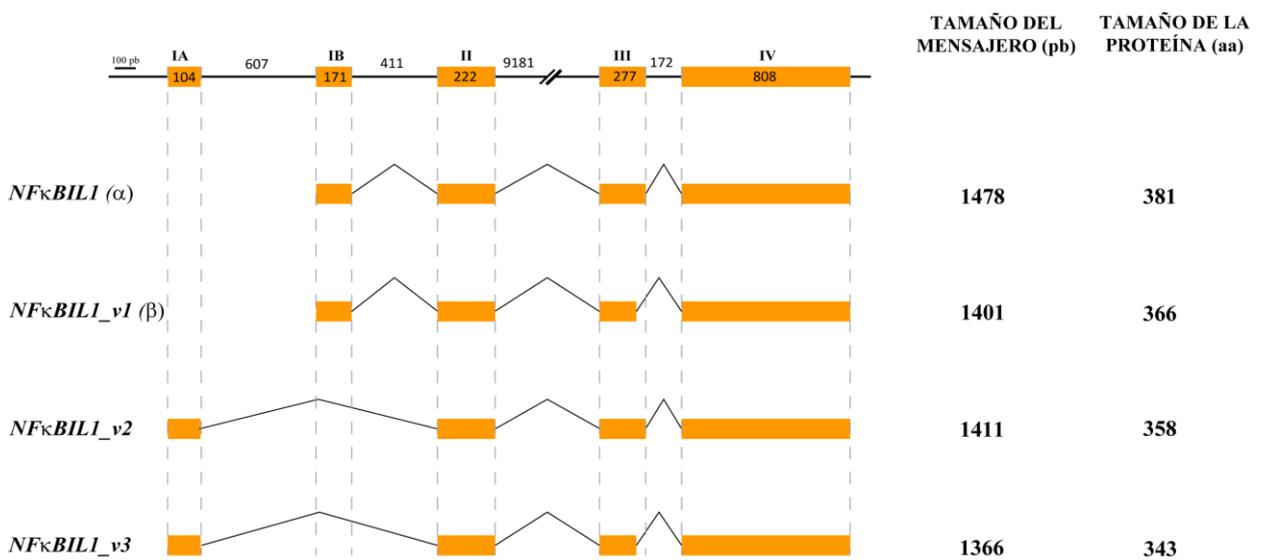


Figura 17. Representación de las distintas isoformas de *NFκBIL1*. La isoforma canónica y la variante número 1 se encontraban ya caracterizadas, mientras que las variantes 2 y 3 fueron nuevamente descritas y caracterizadas. Los rectángulos naranjas representan los exones. En la parte superior se representa la estructura genómica del gen *NFκBIL1* con los tamaños en pb de los exones y de los intrones. A la derecha de cada isoforma se muestran los tamaños de los mensajeros (pb) y de las proteínas a las que dan lugar (aa).

Analizando la secuencia de aminoácidos de las proteínas de *NFκBIL1*, se observó la presencia de dos repeticiones de anquirina correspondientes al final del exón II y otra repetición en los primeros 22 aminoácidos del exón III, comunes a todas las isoformas de *NFκBIL1* (Figura 18). Los miembros de la familia de inhibidores $I\kappa B$ se unen a $NF-\kappa B$ mediante estos dominios de anquirina sirviendo de dominios de interacción proteína-proteína. Esta similitud es la que sugirió la posibilidad de que

NFκBIL1 perteneciera a la familia de inhibidores IκB. Además en el extremo C-terminal de las proteínas encontramos tres residuos de leucina y una valina que se repetían cada siete aminoácidos, lo que podía constituir un tipo de dominio de cremallera de leucina.

También comprobamos si las isoformas de *NFκBIL1* que nosotros describimos presentaban secuencias PEST y encontramos que las nuevas isoformas identificadas, *NFκBIL1_v2* y *NFκBIL1_v3*, presentan motivos PEST a continuación de los dominios de anquirina pero no en los extremos N-terminales como ocurre en los aminoácidos 1-18 de las isoformas *NFκBIL1* y *NFκBIL1_v1* (Figura 18).

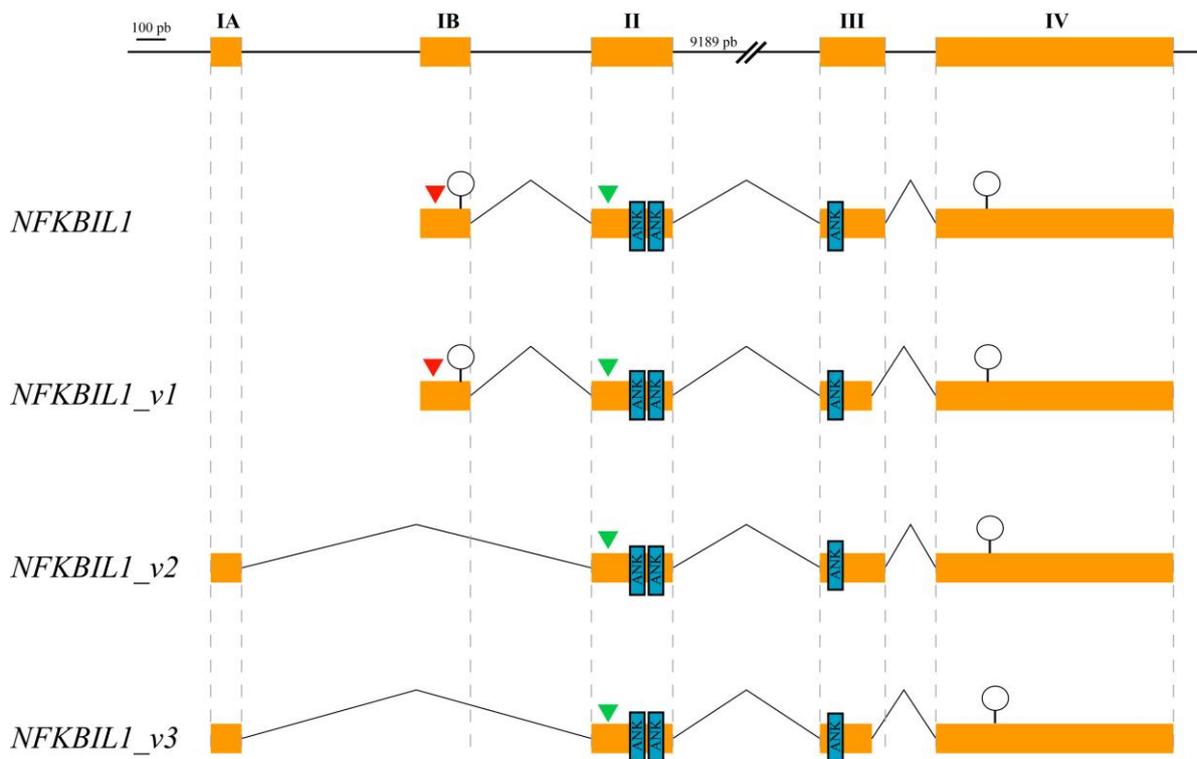


Figura 18. Isoformas de *NFκBIL1* con sus correspondientes dominios de anquirina y secuencias PEST. Los rectángulos azules representan los dominios de anquirina y los círculos las secuencias PEST. Los triángulos invertidos representan los posibles codones de inicio de la traducción.

El sitio de inicio de la traducción predicho de las isoformas *NFκBIL1* y *NFκBIL1_v1* se encuentra en el primer codón ATG del exón IB aunque la secuencia que rodea a este codón no corresponde a una secuencia Kozak adecuada para el inicio de la

traducción. La secuencia alrededor del codón de iniciación (CTGatgA) se asemeja pobremente a la secuencia Kozak consenso más adecuada (RNNatgG, donde R= G ó A, y N= G, A, T ó C) o incluso a secuencias Kozak menos adecuadas (RNNatgY ó YNNatgG, donde Y= C ó T). En cambio, en el exón II, se encuentra un segundo ATG, el primero en caso de las isoformas *NFκBIL1_v2* y *NFκBIL1_v3*, con una secuencia Kozak adecuada (TCCatgG).

1.2 En otros mamíferos

Puesto que el splicing alternativo es un mecanismo que aumenta la diversidad transcriptómica y proteómica de un organismo, decidimos caracterizar las isoformas de *NFκBIL1* en distintas especies con el fin de comprobar si encontrábamos las mismas que en humano, distintas o ninguna. Además de en *Homo sapiens*, se caracterizaron las isoformas de *NFκBIL1* en las siguientes especies: *Sus scrofa*, *Bos taurus*, *Mus musculus*, *Rattus norvegicus* y *Macaca mulatta*. De estas especies disponíamos de una batería de ARNs de tejidos comerciales (Ver materiales y métodos) donde pudimos analizar la expresión de *NFκBIL1* y caracterizar sus isoformas. Además, nos suministraron muestras de sangre de distintos primates del Zoo-Aquarium de Madrid donde pudimos analizar las posibles variantes de este gen. Los primates de los que obtuvimos muestras de sangre fueron los siguientes: *Macaca fascicularis*, *Cebus apella*, *Colobus guereza*, *Pongo pygmaeus*, *Papio sp* (híbrido) y *Gorilla gorilla*.

1.2.1 *NFκBIL1* en *Mus musculus* y *Rattus norvegicus* (Orden Rodentia)

El MHC en *Mus musculus* (ratón) se encuentra en el cromosoma 17B1 y *NFκBIL1* se encuentra en esta región. En las diferentes bases de datos (NCBI y Ensembl) sólo aparece un mensajero compuesto por los 4 exones descritos como en la isoforma canónica de humanos. El mensajero final tendría una longitud de 1419 pb y a partir de la secuencia codificante se obtendría una proteína de 381 aminoácidos. La región del MHC de *Rattus norvegicus* (rata) se encuentra en el cromosoma 20p12 donde se encuentra *NFκBIL1*. En las diferentes bases de datos (NCBI y Ensembl) también aparece sólo un mensajero compuesto por los 4 exones descritos como en la isoforma

canónica de humanos. El mensajero final tendría una longitud de 1463 pb y a partir de la secuencia codificante se obtendría una proteína de 381 aminoácidos. Mediante predicción de dominios estructurales en diversas bases de datos, encontramos las secuencias de anquirina en los exones II y III como en humanos.

Al no encontrarse más isoformas de *NFκBIL1* en las bases de datos descritas para estas especies, se procedió a continuación a la comprobación del tamaño del exón III en cada especie, puesto que en humanos dos de las isoformas caracterizadas presentaban un exón III con una falta de 45 pb en el extremo 3' del mismo. Se diseñaron oligos específicos para esa región del gen, en los exones II y IV, para estas especies y se realizaron RT-PCRs en los tejidos disponibles para cada una de ellas. En los resultados vimos que sólo se encontraba el exón III largo con lo que las variantes 1 y 3 encontradas en humanos no se expresaban en estas especies (Figura 19). Posteriormente se realizó un análisis de los extremos 5'UTR y se diseñó oligos en una región homóloga a la región 5'UTR de las isoformas *NFκBIL1_v2* y *NFκBIL1_v3* humanas y se comprobó mediante PCR que no se amplificaban estas isoformas en rata y ratón, con lo que se confirmaba lo predicho en las bases de datos consultadas y sólo se expresaba la forma canónica en ambas especies del Orden Rodentia (Figura 20).

Comparando las secuencias genómicas del gen *NFκBIL1* en rata, ratón y humano comprobamos que el intrón II en ratón y rata tiene una longitud superior al humano y de todos los exones, la secuencia del exón III es la más conservada. En los demás exones se observan pequeñas diferencias de longitud y de secuencia.

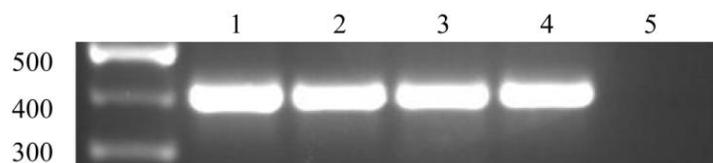


Figura 19. Comprobación del tamaño del exón III en *Mus musculus* en diferentes tejidos. 1: Cerebro, 2: Riñón, 3: Hígado, 4: Pulmón y 5: Blanco. En todos los casos se observa sólo una banda que indica que el exón III se encuentra en esta especie en su forma más larga.

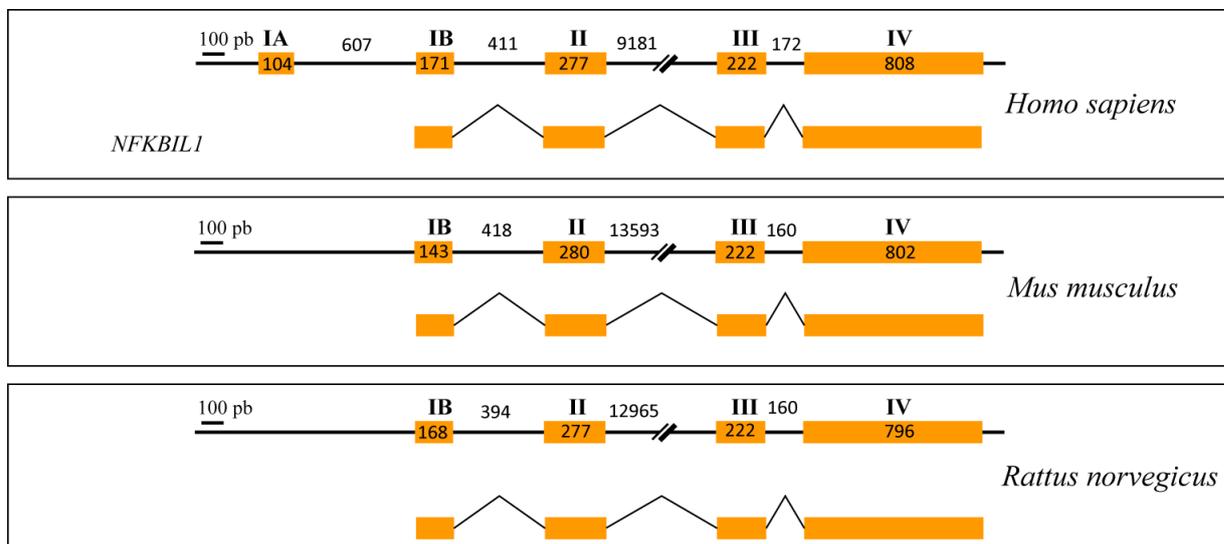


Figura 20. Comparación del ADN genómico y de las formas de splicing de *NFκB1* en *Homo sapiens*, *Mus musculus* y *Rattus norvegicus*. Se indican en pb los tamaños de los exones y de los intrones para cada especie.

1.2.2 *NFκB1* en *Sus scrofa* y *Bos Taurus* (Orden Artiodactyla)

En *Bos taurus* (vaca) la región del MHC se ubica en el cromosoma 23 y *NFκB1* se sitúa en esta región. En las bases de datos analizadas (NCBI y Ensembl) sólo aparece una variante semejante a la canónica de humanos, con 4 exones. La longitud del transcrito es de 1417 pb y la proteína a la que da lugar la secuencia codificante es de 381 aminoácidos. El genoma de cerdo, *Sus scrofa*, aún no se encuentra en el mismo estado de secuenciación y anotación que otras especies. La región del MHC en cerdo se encuentra pobremente anotada y *NFκB1* no aparece en las bases de datos. Por ello, recurrimos a la búsqueda por homología utilizando la secuencia genómica de vaca y diseñamos los oligos en los exones correspondientes.

Después de la búsqueda en las bases de datos de las isoformas de *NFκB1* descritas para estas especies, se realizó, a continuación, la comprobación del tamaño del exón III en cada especie como habíamos hecho anteriormente. Para ello diseñamos oligos que hibridaran en los exones II y IV y realizamos PCRs en las muestras de tejidos disponibles para cada especie. Como en el caso anterior, sólo se encontraba el

exón III en su longitud más larga. También realizamos un análisis de las 5'UTR para intentar amplificar isoformas con distintos extremos 5'UTR pero no se expresaban isoformas con esas características. Lo sorprendente en *Bos taurus* fue que se consiguió amplificar además de la isoforma canónica, una forma nueva en la que faltan 4 pb en el exón IV lo que cambiaba la pauta de lectura (Figura 21). De esta manera el exón IV se divide en dos nuevos exones, uno de una longitud de 122 pb y otro de una longitud de 653 pb. Las secuencias de anquirina seguían presentes en los exones II y III y la secuencia codificante daba lugar a una proteína de 328 aminoácidos.

Analizando las secuencias genómicas, observamos de nuevo que el exón III es el más conservado entre estas especies así como el exón II. Los tamaños de los intrones se asemejan y el del resto de exones, encontrando diferencias de tamaño y secuencia en comparación con humano.

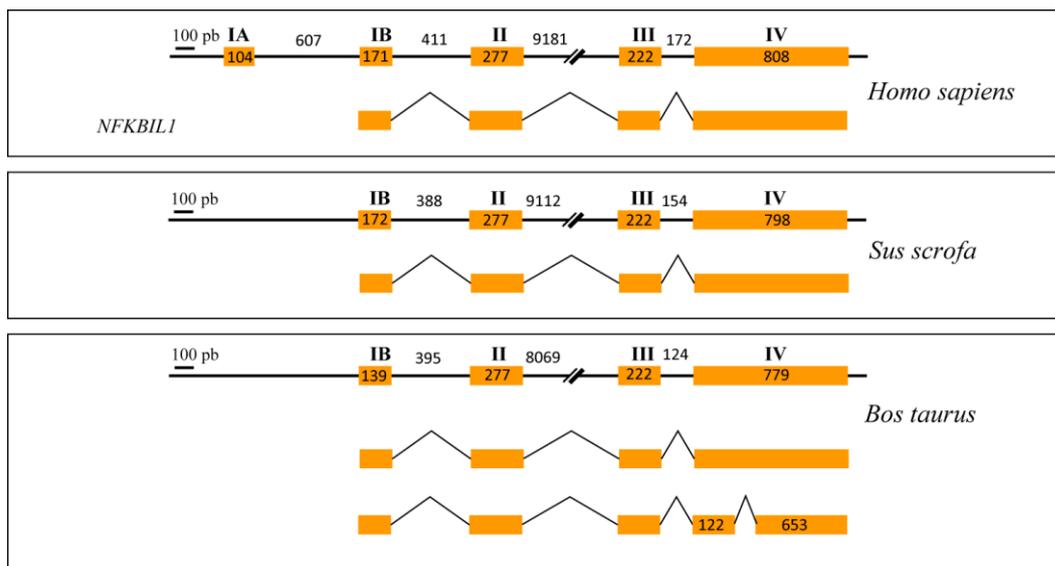


Figura 21. Comparación del ADN genómico y de las formas de splicing de *NFκBIL1* en *Homo sapiens*, *Sus scrofa* y *Bos taurus*. Se indican los tamaños de los exones e intrones en pares de bases.

Para comparar las secuencias de aminoácidos de la proteína de la isoforma canónica NFκBIL1 entre humano y el resto de especies (rata, ratón, vaca y cerdo) realizamos un alineamiento múltiple con el programa ClustaW. El mayor porcentaje de identidad fue encontrado entre *Sus scrofa* y *Bos taurus* con un valor del 97%. Entre *Mus musculus* y *Rattus norvegicus* el porcentaje de identidad era del 96%. Entre humano y

roedores el porcentaje era del 92% y entre humanos yo vaca y cedo, la secuencia proteica de NFκBIL1 era idéntica en un 95% (Figura 22).

Bos	MSDSPSQVPEGEASTSVCRPKS-SMASTRRQRRERRFRRYLSAGRLVRAQALLQRHPGL	59
Sus	MSNPSPQVPEGEASTSVCRPKS-SMASTRRQRRERRFRRYLSAGRLVRAQALLQRHPGL	59
Homo	MSNPSPQVPEEEASTSVCRPKS-SMASTRRQRRERRFRRYLSAGRLVRAQALLQRHPGL	59
Mus	MSNPSPQAPEEEASTSVCRPQSCSMASASRRHRRERRFRRYLSAGRLVRAQALLQRHPGL	60
Rattus	MSNPSPQAPEEEASTSVCRPQS-SMASVSRHRRERRFRRYLSAGRLVRAQALLQRHPGL	59
	*:****.* ** *****.* ****.**:*****	
Bos	DVDAGQPPPLHRACARHDAPALCLLLRLGADPAHQDRHGDTALHAAARQGPDAYTDFFLP	119
Sus	DVDAGQPPPLHRACARHDAPALCLLLRLGADPAHQDRHGDTALHAAARQGPDAYTDFFLP	119
Homo	DVDAGQPPPLHRACARHDAPALCLLLRLGADPAHQDRHGDTALHAAARQGPDAYTDFFLP	119
Mus	DVDAGQPPPLHRACARHDAPALCLLLRLGADPAHQDRHGDTALHAAARQGPDAYTDFFLP	120
Rattus	DVDAGQPPPLHRACARHDAPALCLLLRLGADPAHQDRHGDTALHAAARQGPDAYTDFFLP	119
	*****:*****	
Bos	LLSRCPSAMGIKKNKDGETPGQILGWGPPWDSAEDEEASKEREWQKIQGELEDEWQE	179
Sus	LLSRCPSAMGIKKNKDGETPGQILGWGPPWDSAEDEEASKEREWQKIQGELEDEWQE	179
Homo	LLSRCPSAMGIKKNKDGETPGQILGWGPPWDSAEDEEEDDASKEREWQKIQGELEDEWQE	179
Mus	LLSRCPSAMGIKKNKDGETPGQILGWGPPWDSAEDEEDEVSKEREWQKIQGELEDEWQE	180
Rattus	LLSRCPSAMGIKKNKDGETPGQILGWGPPWDSAEDEEDEVSKEREWQKIQGELEDEWQE	179
	*****:..*****	
Bos	VIGRFEDDASHETQEPESFSAWSDRMAREHAQKHQQRRREAEAGACRPPRAEGSGHSWRQQ	239
Sus	VIGRFEDDASHETQEPESFSAWSDRMAREHAQKRQQQR-ETEGACRPPRAEGSSHSWRQQ	238
Homo	VMGRFEGDASHETQEPESFSAWSDRMAREHAQKQQQREAEAGSCRPPRAEGSSQSWRQQ	239
Mus	VIGRFEDDASRDTEPESFSAWSERLAREHAQKQR-QQLEAEGSCRPPRAEGSSHSWRQH	239
Rattus	VIGRFEDDASRETQEPESFSAWSERLAREHAQKQRQQLEAEGSRPPRAEGSSHSWRQQ	239
	*:****.**:*****.*:***** : *: *:*: *****.*:****:	
Bos	EEEQRLFRERARAKEEELRESRARRAQEALRDRAPEPARAGPRAEHPRGAGRGLWRFGD	299
Sus	EEEQRLFRERARAKEEELRESRARRAQEAPRDPVPEPARAGPRAEHPRGAGRGLWRFGD	298
Homo	EEEQRLFRERARAKEEELRESRARRAQEALGDREPKPTRAGPREEHPRGAGRGLWRFGD	299
Mus	EEEQRLFRERARAKEELCESRARRAQEAQRDKGPEPPRAGPRAEHPRGAERGLWRFGD	299
Rattus	EEEQRLFRERARVKEKELCESRARRAQEAQGDRGPAPPRAEPRAEHPRGAGRGLWRFGD	299
	*****.**:** ***** * * *.* ** ***** ** *****	
Bos	VPWPCPGGDPEAMAAALVARGPPLEEQQALRRYLVRVQVRWHPDRFLQRFRSQIETWEL	359
Sus	VPWPCPGGDPEAMAAALVARGPPLEEQQALRRYLVRVQVRWHPDRFLQRFRSQIETWEL	358
Homo	VPWPCPGGDPEAMAAALVARGPPLEEQQALRRYLVRVQVRWHPDRFLQRFRSQIETWEL	359
Mus	VPWPCPGGDPEAMAAALVARGPPLEEQQALKRYLVRVQVRWHPDRFLQRFRSQIETWEL	359
Rattus	VPWPCPGGDPEAMAAALVARGPPLEEQQALKRYLVRVQVRWHPDRFLQRFRNQIETWEL	359
	*****:*****.*:*****	
Bos	GRVMGAVTALSQALNRHAEALK	381
Sus	GRVMGAVTALSQALNRHAEALK	380
Homo	GRVMGAVTALSQALNRHAEALK	381
Mus	GRVMGAVTALSQALNRHAEALK	381
Rattus	GRVMGAVTALSQALNRHAEALK	381

Figura 22. Alineamiento múltiple de la proteína NFκBIL1 en *Homos sapiens*, *Sus scrofa*, *Bos taurus*, *Mus musculus* y *Rattus norvegicus*. Con un asterisco se representan los aminoácidos idénticos, con dos puntos (:) las sustituciones conservadas y con un punto (.) las sustituciones de aminoácidos semi-conservadas.

1.2.3 *NFκBIL1* en otros mamíferos del Orden Primates

1.2.3.1 *NFκBIL1* en distintos tejidos de *Macaca mulatta*

La región del MHC en *Macaca mulatta* está localizada en el cromosoma 4 donde se encuentra *NFκBIL1*. *Macaca mulatta* pertenece a la familia Cercopithecidae, dentro del infraorden simiiformes, parvorden **Catarrhini**. En la base de datos de Ensembl sólo aparece una isoforma semejante a la canónica humana con sus cuatro exones correspondientes. Sin embargo, en el NCBI se encuentra para esta especie, tres isoformas correspondientes a las variantes *NFκBIL1*, *NFκBIL_v1* y *NFκBIL_v2*. Por ello diseñamos oligos para intentar amplificar no sólo estas tres variantes si no también la cuarta encontrada en humanos.

En *Macaca mulatta* se amplificaron en distintos tejidos comerciales las cuatro isoformas encontradas en humanos. Pero además aparecieron bandas adicionales que una vez secuenciadas y analizadas se vio que correspondían a dos isoformas nuevas que no se encontraban en la base de datos. Una de las isoformas se asemejaba a la isoforma canónica *NFκBIL1* humana pero retenía las primeras 100 pb del intrón comprendido entre los exones II y III. La otra isoforma tenía el extremo 5'UTR como la isoforma *NFκBIL_v2* humana pero también retenía las 100 pb del intrón comprendido entre los exones II y III (Figura 23).

Analizando las secuencias de las nuevas isoformas, se observó en la predicción del programa ORFinder del NCBI, que la retención de esa parte del intrón II provocaba un cambio en la pauta de lectura y se producirían proteínas de menor tamaño. Esto era debido a que al cambiar la fase de lectura, en el comienzo del exón III se introduce un codón de terminación. En concreto, en la isoforma semejante a la isoforma *NFκBIL_v2* humana, se expresaría una proteína de 142 aa, los 88 primeros aa iguales a los de la proteína que daría lugar a la variante humana a la que se asemeja y después produciéndose un cambio de lectura que además producía la inserción de un codón stop, dando lugar a una proteína más corta de lo esperado. En el caso de la isoforma que se asemejaba a la variante canónica, la retención del intrón de 100 pb también produce un cambio en la fase de lectura, que codificaría una proteína de 165 aa y terminando

también en un codón stop producido por el cambio de fase debida a la retención de esa porción del intrón.

El contexto genómico del gen *NFκBIL1* en *Macaca mulatta* es muy semejante al encontrado en *Homo sapiens*, encontrando la mayor conservación de secuencias en los exones II y III. Los tamaños de los intrones también tienen una longitud parecida a los humanos. Comparando la secuencias proteicas de la variantes de *NFκBIL1* comunes entre *Homo sapiens* y *Macaca mulatta* encontramos porcentajes de identidad del 99%.

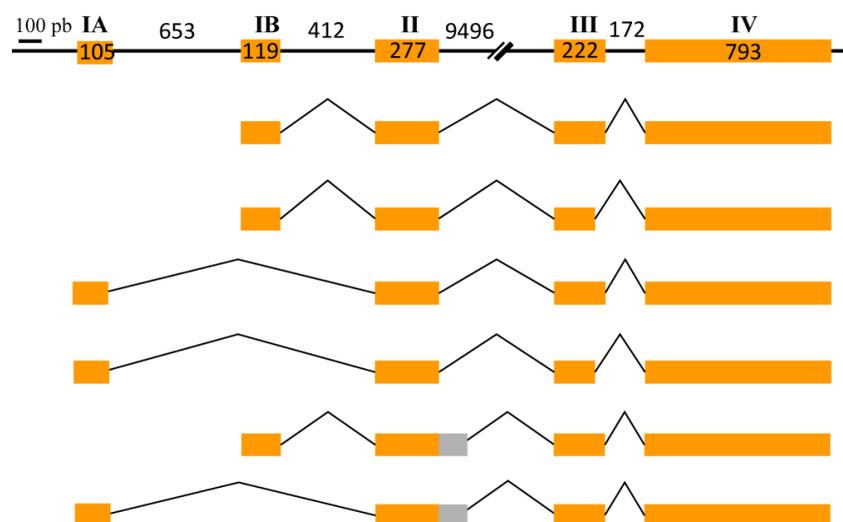


Figura 23. Isoformas de *NFκBIL1* encontradas en *Macaca mulatta*. Además las dos isoformas nuevas, *NFκBIL1* en *Macaca mulatta* también presenta las cuatro isoformas identificadas en humanos. La isoforma que se asemeja a la forma canónica retiene 100 pb del intrón situado entre los exones II y III mientras que la isoforma que se asemeja a la variante 2 humana retiene 100 pb del mismo intrón.

1.2.3.2 *NFκBIL1* en muestras de sangre de otros primates

El estudio del gen *NFκBIL1* se extendió a otros primates como son *Macaca fascicularis*, *Cebus apella*, *Colobus guereza*, *Pongo pygmaeus*, *Papio sp* y *Gorilla gorilla*. De estas especies, conseguimos muestras de sangre gracias a la colaboración con el Zoo-Aquarium de Madrid para el estudio de las variantes de *NFκBIL1* pero no disponíamos de muestras de distintos tejidos.

Cebus apella pertenece a la familia Cebidae, dentro del infraorden simiiformes, parvorden **Platyrrhini**. *Macaca fascicularis*, *Papio sp* y *Colobus guereza* pertenecen a la

familia Cercopithecidae, dentro del infraorden simiiformes, parvorden **Catarrhini**. *Pongo pygmaeus* y *Gorilla gorilla* pertenecen a la familia Hominidae, dentro del infraorden simiiformes, parvorden **Catarrhini** (Figura 24).

En la base de datos de Ensembl sólo aparece la secuencia de *NFκBIL1* en *Gorilla gorilla* de todos los primates de los que disponemos. En algunos de ellos no se encuentra ni el genoma secuenciado. En el NCBI sólo aparece la secuencia de este gen en *Pongo pygmaeus*. Puesto que no se disponía de la secuencia de *NFκBIL1* en todos los primates en los cuales se quería realizar el estudio, se decidió alinear las secuencias de *NFκBIL1* de todos los primates de los que hubiera disponible la secuencia en las bases de datos y diseñar unos oligos degenerados que pudieran hibridar en las muestras de primates de las que disponíamos.

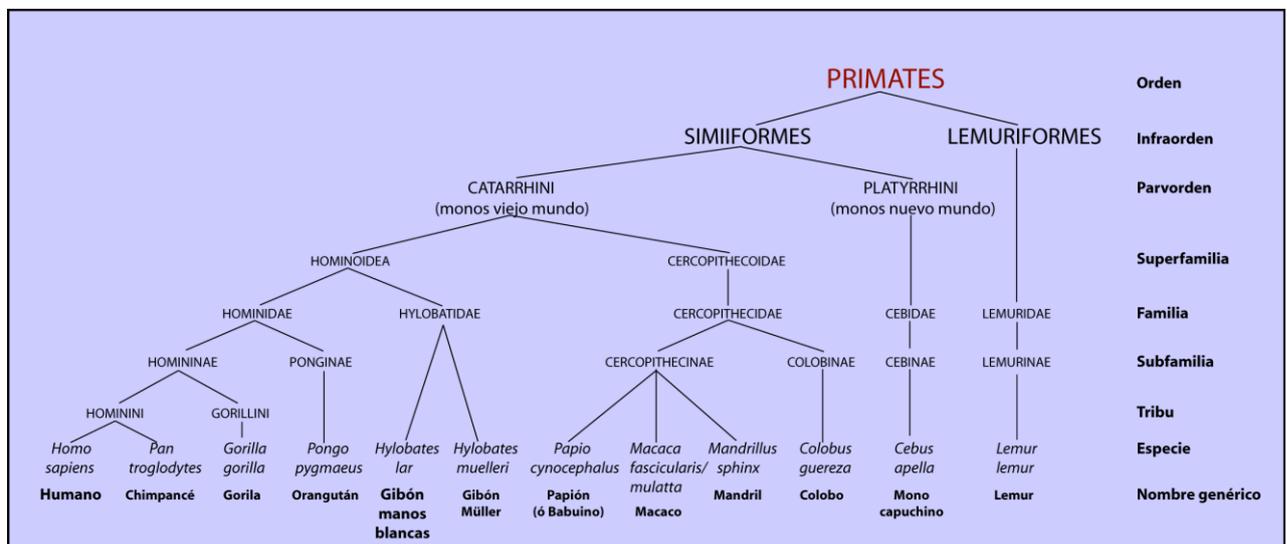


Figura 24. Representación del orden Primates.

Una vez diseñados los oligos, se realizaron PCRs sobre las muestras de ADNc obtenido a partir del ARN extraído de los linfocitos de las muestras de sangre de cada primate. A continuación, se clonaron los productos que se obtuvieron y se secuenciaron.

En *Cebus apella*, el único primate perteneciente a los monos del Nuevo Mundo del que disponíamos, sólo presentaba una variante de *NFκBIL1*, la semejante a la isoforma canónica humana. No se encontraba ni la variante con el exón III más corto ni las variantes con distintos extremos 5'UTR (Figura 25).

Entre los monos del Viejo Mundo, teníamos representantes de la superfamilia Hominoidea y de la superfamilia Cercopithecoidea. Dentro de la familia Cercopithecidae, disponíamos de muestras de *Macaca fascicularis* y *Papio sp*, pertenecientes a la subfamilia Cercopithecinae, y muestras de *Colobus guereza*, representante de la familia Colobinae. En los análisis de las isoformas de gen *NFκBIL1*, se vio que en *Colobus guereza* aparecían dos isoformas, la isoforma semejante a la canónica humana *NFκBIL1* y la variante 1 con el exón III 45 nucleótidos más corto. En el caso del *Papio sp*, se encontraron las cuatro isoformas amplificadas en humanos, *NFκBIL1*, *NFκBIL1_v1*, *NFκBIL1_v2* y *NFκBIL1_v3*. En el caso de *Macaca fascicularis* aparecían también las cuatro isoformas pero además se expresaba otra variante en la que se retenían 100 pb del intrón entre el exón II y III (Figura 25). Esta variante es la misma que se obtuvo en *Macaca mulata*. La retención de esta parte del intrón II provocaba un cambio en la pauta de lectura produciéndose un codón stop prematuro y dando lugar a una proteína de menor tamaño del esperado.

Dentro de la familia Hominidae, disponíamos de representantes de la subfamilia Ponginae, *Pongo pygmaeus* y de la subfamilia Homininae, *Gorilla gorilla*. En estos dos primates se encontraron las cuatro isoformas descritas en humanos, sin aparición de ninguna variante nueva (Figura 25)

Como resumen de la expresión de *NFκBIL1* en linfocitos de muestras de sangre de distintos primates, el gen *NFκBIL1* se expresaba en cuatro variantes de splicing en *Homo sapiens*, *Pongo pygmaeus*, *Gorilla gorilla* y *Papio sp*. En *Macaca fascicularis* además de estas cuatro isoformas, se encontraba en sangre otra isoforma nueva característica por la retención de parte de un intrón entre los exones II y III que provocaba un cambio en la fase de lectura. En *Colobus guereza* se encontraron sólo dos variantes de *NFκBIL1*, las correspondientes a las isoformas humanas canónica y variantes 1, luego no se encontraron las isoformas correspondientes a las variantes humanas con distinto 5'UTR. Y por último, en *Cebus apella*, el único primate perteneciente a los monos del Nuevo Mundo del que disponíamos, sólo se encontró una isoforma de *NFκBIL1*, la semejante a la variante canónica de humanos (Figura 25).

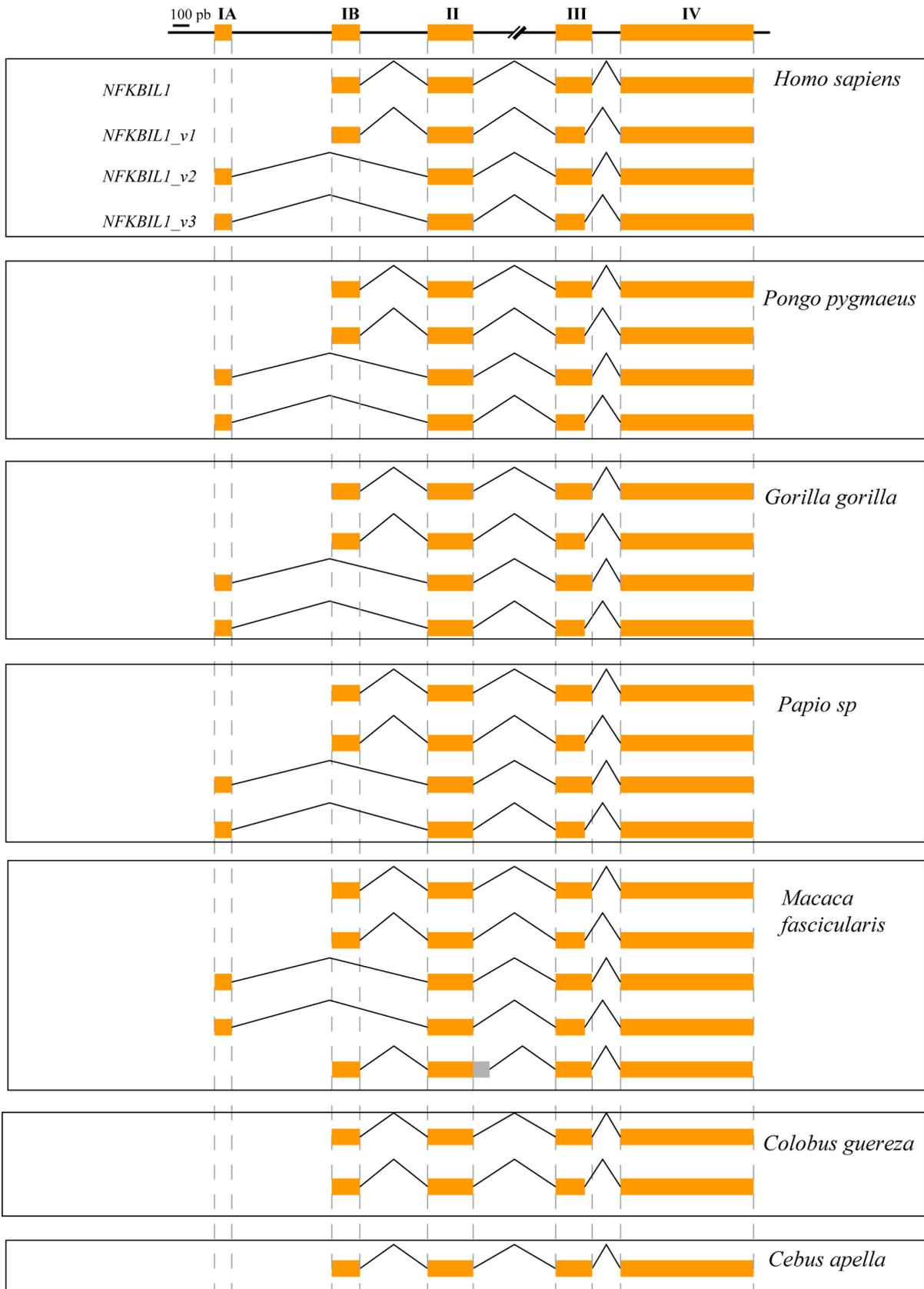


Figura 25. Isoformas de *NFκBIL1* en muestras de sangre de diferentes primates.

2. ANÁLISIS DE LA EXPRESIÓN DE LAS ISOFORMAS DE *NFκBIL1* EN LÍNEAS CELULARES Y TEJIDOS HUMANOS

Con el fin de cuantificar en humanos la expresión de las distintas isoformas de *NFκBIL1* en distintas muestras de tejidos y líneas celulares, se diseñó un ensayo de PCR a tiempo real, en el cual cada par de oligos sólo reconociera una isoforma y no las demás (Figura 26). El éxito de un análisis de PCR a tiempo real depende de la eficiencia y la amplificación específica del producto. Para que los oligos fueran específicos de cada isoforma, en el caso de la variante canónica *NFκBIL1* se diseñaron con el programa Primer3 los oligos en el exón IB y en la parte final del exón III ausente en las isoformas *NFκBIL1_v1* y *NFκBIL1_v3*. Para la isoforma *NFκBIL_v1*, el diseño de los oligos se realizó en el exón IB y el antisentido entre los exones III y IV. En el caso de las isoformas *NFκBIL1_v2* y *NFκBIL1_v3* los oligos se diseñaron en el exón IA y en el final del exón III en el caso de la variante 2 y entre el exón III y IV en el caso de la variante 3 (Figura 26). De esta manera obtuvimos un ensayo capaz de diferenciar cada isoforma. Estas isoformas fueron clonadas en plásmidos PGEM-T y de cada una de ellas se generaron curvas estándares o curvas patrón usando diluciones seriadas de cada plásmido.

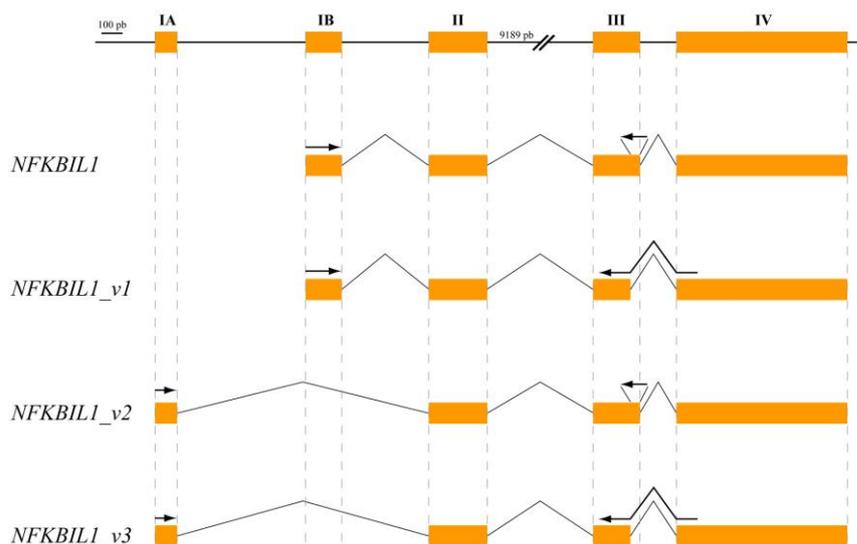


Figura 26. Diseño de ensayo de PCR a tiempo real donde se representa cada isoforma con los oligos diseñados específicamente para cada una de ellas. Las flechas indican los oligos diseñados para la amplificación de cada variante.

Para comprobar que realmente los oligos diseñados para amplificar cada isoforma eran específicos de cada una de ellas, se realizaron amplificaciones cruzadas que consistían en enfrentar cada pareja de oligos a su una isoforma y contra el resto de plásmidos conteniendo las demás isoformas que en teoría no debían amplificar. De esta manera, nos aseguramos de que la amplificación fuera específica (Figura 27).

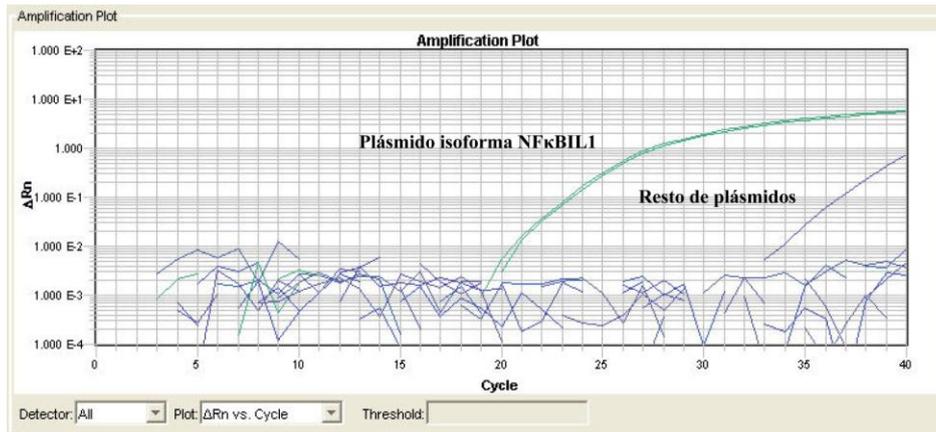


Figura 27. Gráfica representativa del análisis de la amplificación cruzada de los oligos utilizados para RT-PCR cuantitativa. Análisis de la amplificación de los oligonucleótidos específicos para la isoforma *NFκBIL1* utilizando como molde todos los plásmidos portadores de las cuatro isoformas descritas.

Una vez que se comprobó que cada par de oligos era específico de una isoforma, se procedió a medir los niveles de expresión de las isoformas de *NFκBIL1* mediante cuantificación absoluta en ADNc de líneas celulares, en tejidos adultos, fetales y tumorales. La cuantificación absoluta se consiguió comparando los valores de C_T de las muestras estudiadas con las curvas estándares generadas anteriormente. De esta manera se obtuvo el número de copias de cada isoforma dentro de una misma muestra y se calcularon los porcentajes de cada isoforma en cada línea celular o tejido.

Analizando la curva de melting de los amplicones generados en este ensayo de PCR a tiempo real, vimos que no presentaba un patrón normal sino que presentaba una curva compleja (Figura 28A). Las curvas de melting dependen de la proporción GC/AT y de la longitud total del amplicón (Ririe et al, 1997). Los amplicones de este experimento eran largos (~500 pb) puesto que era la única manera de conseguir diferenciar las cuatro isoformas de *NFκBIL1*. Esta longitud es la responsable de que las curvas de melting no fueran simples. Se realizó una electroforesis de los productos

amplificados y se obtuvo el tamaño esperado para cada isoforma (Figura 28B). La explicación de las curvas de melting complejas de debía a que los fragmentos más largos de ADN constan de dominios internos de melting que dan lugar a estas curvas mientras que productos más pequeños daban lugar a un sólo pico de temperatura (Benight et al., 1981; Li et al., 2003).

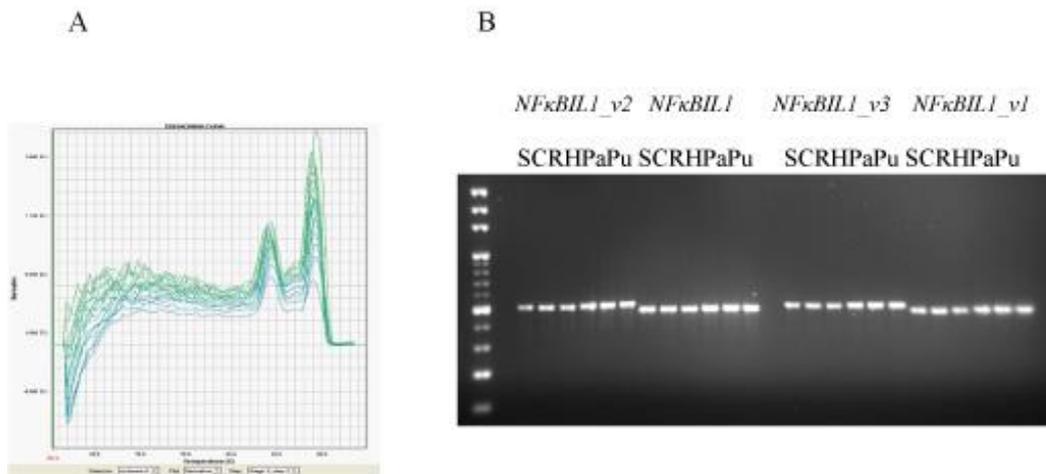


Figura 28. A Un ejemplo de curva de melting compleja obtenida de los amplicones de las isoformas de *NFκBIL1*. B. Electroforesis de los amplicones en diferentes tejidos tras la carrera de PCR en tiempo real. S: Sangre, C: Corazón, R: Riñón, H: Hígado, Pa: Páncreas y Pu: Pulmón

Analizando los resultados en los distintos tejidos y líneas celulares, se observaron diferencias significativas en la expresión de cada isoforma dentro de cada muestra. En todas las muestras analizadas, la isoforma canónica *NFκBIL1* era la que mayoritariamente se expresaba mientras que la variante *NFκBIL1_v3* la de menor expresión en todos los casos, llegando a no expresarse en algunos tejidos. Los porcentajes de expresión de la isoforma canónica rondaban entre el 60-90%, los de la variante 1, entre el 10-20% y los de las variantes 2 y 3 con porcentajes menores del 10% siendo la variante 3 la de menor expresión en casi todos los casos.

En los **tejidos adultos normales** analizados (Figura 29), observamos que aunque en todos los tejidos la variante *NFκBIL1* era la de mayor expresión, en páncreas alcanzaba una expresión del 90% donde además la variante 3 no se expresaba. En los

demás tejidos, *NFκBIL1_v1* resultaba ser la segunda isoforma con mayor expresión a nivel de mensajero exceptuando en riñón, cerebro y timo donde las variantes 1 y 2 presentan unos porcentajes de expresión similares.

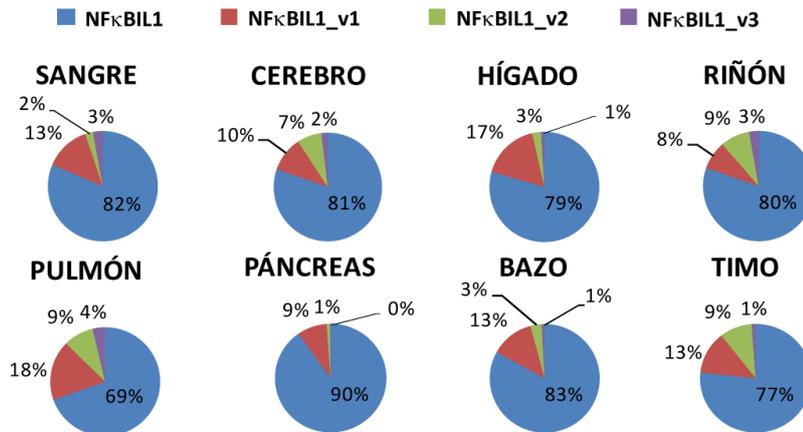
En los **tejidos fetales** estudiados (Figura 29), observamos, de nuevo, que la variante *NFκBIL1* se expresaba en mayor proporción que las demás isoformas. En el caso del tejido cerebral fetal, era donde alcanzaba su menor proporción, con un 66%, y era en este mismo tejido donde las variantes 1 y 2 alcanzaban una expresión similar así como en timo y riñón. La variante 3 se expresaba también en todos los tejidos fetales pero con los niveles más bajos de expresión.

En los **tejidos adultos tumorales** analizados (Figura 29), la isoforma *NFκBIL1* seguía siendo la de mayor expresión (80-90%) y lo sorprendente es que las variantes *NFκBIL1_v2* y *NFκBIL1_v3* prácticamente no se expresaban en estos tejidos tumorales, llegando incluso la variante 3 a no expresarse en hígado tumoral.

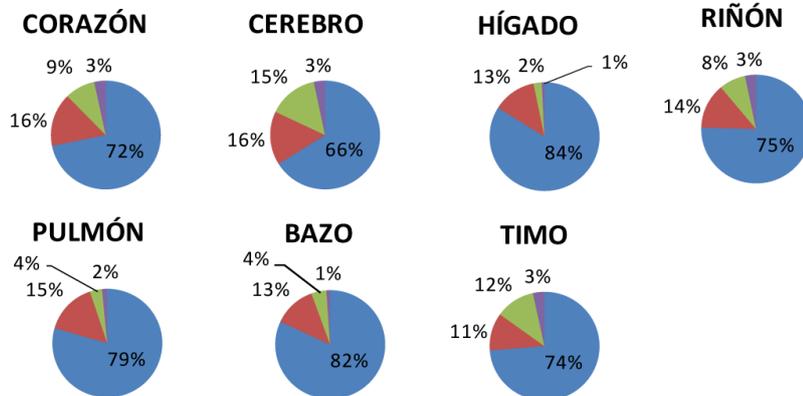
Por último en las **líneas celulares** estudiadas (Figura 29), se obtenía el mismo patrón de expresión que en la mayoría de los tejidos analizados, con unas proporciones semejantes a los tejidos tumorales, siendo la variante canónica la más expresada, seguida de la variante 1 y con una baja expresión de las variantes 2 y 3. En la línea celular K562, la variantes 3 no se expresaba.

Con estos datos, se pudo comprobar que en cada muestra analizada, curiosamente siempre era las isoformas canónica *NFκBIL1* y la de mayor expresión a nivel de mensajero, seguida por la variante 1, las cuales comparten el mismo exón IB, mientras que las variantes 2 y 3 eran las de menor expresión, las cuales comparten el mismo extremo 5'UTR.

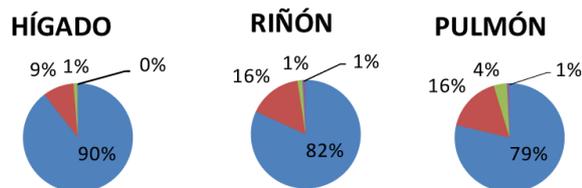
TEJIDOS ADULTOS



TEJIDOS FETALES



TEJIDOS ADULTOS TUMORALES



LÍNEAS CELULARES

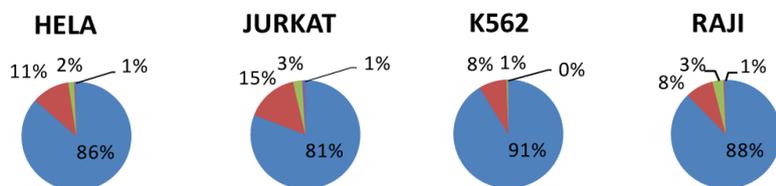


Figura 29. Cuantificación absoluta de los niveles de expresión de todas las isoformas de *NFκBIL1* en tejidos adultos normales, tumorales, fetales y líneas celulares. En rojo se representa la isoforma canónica *NFκBIL1*, en morado la isoforma *NFκBIL1_v1*, en azul la isoforma *NFκBIL1_v2* y en verde la isoforma *NFκBIL1_v3*.

3. EXPRESIÓN DE LAS PROTEÍNAS NFκBIL1

Una vez caracterizadas las formas de splicing de *NFκBIL1* a nivel de ARNm y cuantificado su expresión en distintas líneas celulares y tejidos humanos, procedimos a la expresión de las potenciales proteínas codificadas por cada isoforma con el fin de estudiar posibles diferencias en cuanto a su expresión y localización. Se clonó cada isoforma en un vector de expresión fusionada a epítomos reconocidos por distintos anticuerpos, puesto que los anticuerpos comerciales que reconocen NFκBIL1 disponibles hasta el momento, no distinguen entre las diferentes isoformas ya que están diseñados frente a una zona común de las cuatro variantes. Además, realizamos estudios de localización para analizar posibles diferencias de distribución celular de las proteínas codificadas por cada isoforma.

3.1 Clonación, transfección y análisis por Western Blot

3.1.1 Análisis de las construcciones con epítomos en el extremo carboxilo terminal.

Cada isoforma de *NFκBIL1* fue clonada en un vector pcDNA3.1/V5-His-TOPO y se realizó la transfección en células COS-7. Los pesos moleculares teóricos de cada isoforma de *NFκBIL1* contando con los aminoácidos de los epítomos V5 y la cola de Histidinas eran los siguientes: la isoforma canónica *NFκBIL1* de 46 kDa, *NFκBIL1_v1* de 44,5 kDa, *NFκBIL1_v2* de 44 kDa y *NFκBIL1_v3* de 42,5 kDa. Al realizar el análisis por Western Blot utilizando el anticuerpo contra el epítomo V5 descubrimos bandas adicionales en la expresión de las isoformas *NFκBIL1* y *NFκBIL1_v1* de 44 y 42,5 kDa respectivamente (Figura 30) que además eran más abundantes que las proteínas esperadas para estas dos isoformas. El hecho de que aparecieran dos bandas al expresar la variante canónica y la variante 1 en células transfectadas fue observado previamente por Semple y colaboradores (Semple et al., 2002).

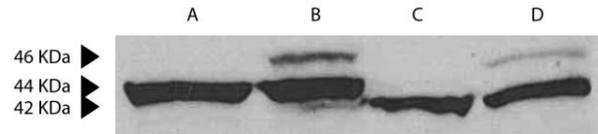


Figura 30. Análisis por Western Blot de la sobre-expresión de las isoformas de *NFκBIL1*. Utilizamos un anticuerpo que reconoce el epítipo V5 presente en el plásmido usado para clonar cada isoforma. En el carril A se encuentra la variante *NFκBIL1_v2*, en el B la isoforma canónica *NFκBIL1*, en el C la variante *NFκBIL1_v3* y en el D la variante *NFκBIL1_v1*. Los pesos moleculares se encuentran a la izquierda de la imagen.

Analizando en detalle la secuencia proteica de las isoformas, comprobamos que en las que aparecen bandas adicionales, *NFκBIL1* y *NFκBIL1_v1*, existen dos sitios de inicio de la traducción, uno en el exón IB y otro en el exón II (Figura 31) observando que la secuencia Kozak que rodea al segundo ATG es más adecuada para el inicio de la traducción. En el caso de la isoforma *NFκBIL1*, a partir del primer ATG se obtendría una proteína con un peso molecular de 46 kDa y a partir del segundo ATG se obtendría una proteína de 44 kDa. En el caso de la isoforma *NFκBIL1_v1* a partir del primer ATG, la proteína tendría un tamaño de 44,5 kDa y a partir del segundo la proteína sería de 42,5 kDa. Estos datos podrían ser una explicación de las bandas que se observaban en el análisis por Western Blot.

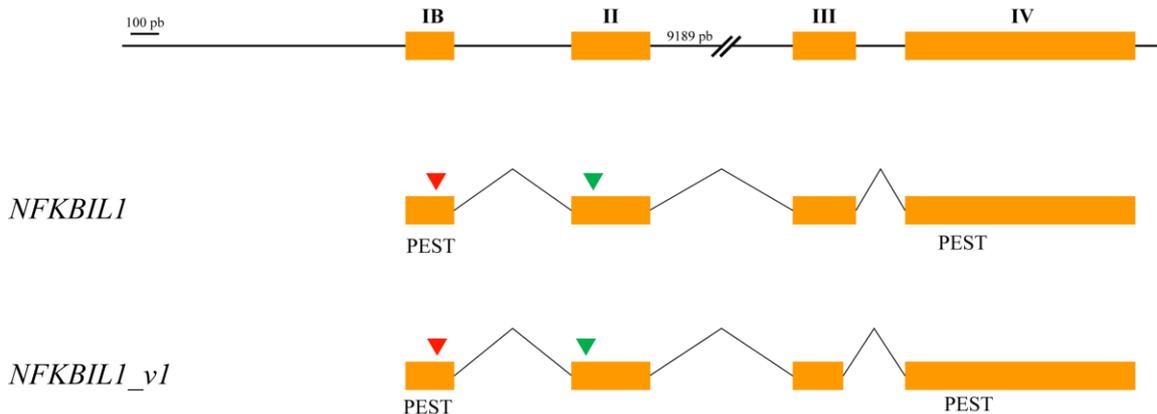


Figura 31. Descripción de las isoformas *NFκBIL1* y *NFκBIL1_v1*. En ambas se encuentran secuencias PEST en los primeros 18 aminoácidos y en el exón IV y dos ATGs (verde y rojo).

3.1.2 Análisis de las construcciones con epítomos en el extremo amino y carboxilo terminal.

Tras analizar la expresión a nivel proteico de las construcciones de cada isoforma en vectores con epítomos en el extremo carboxilo terminal, se observó la presencia de dos bandas de distinto peso molecular, al expresar la isoforma canónica y la variante 1. Además de la idea de los dos posibles inicios de traducción, se pensó en la posibilidad de que la banda de menor peso molecular que aparecía al expresar la variante canónica y la 1 podía deberse a un proceso proteolítico en el extremo amino de la proteína por lo que se procedió a realizar la expresión de la variante canónica con un epítomo también en el extremo amino terminal.

La isoforma canónica *NFκBIL1* fue clonada en un plásmido con un epítomo Flag en el extremo amino terminal y un epítomo V5 en el extremo carboxilo terminal. Las transfecciones con este plásmido y con plásmidos controles se realizaron en células COS-7. Se usó como control del análisis con el anticuerpo que reconoce el epítomo Flag, un plásmido que contenía el epítomo Flag en el extremo amino terminal fusionado a la proteína LacZ. De esta manera, si sucedía un proceso proteolítico que provocara un procesamiento de la proteína *NFκBIL1* en su extremo amino terminal, con lo que no se observaría banda en el revelado con un anticuerpo anti Flag, el control nos serviría para comprobar que el experimento funcionaba. Como control para el revelado con un anticuerpo que reconoce el epítomo V5 situado en el extremo carboxilo terminal, se transfectaron células con los plásmidos que contenían las isoformas *NFκBIL1* y *NFκBIL1_v1*. En el revelado del gel con el anticuerpo que reconoce el epítomo Flag, observamos la banda correspondiente a la proteína LacZ fusionada a este epítomo y en el caso de la isoforma canónica fusionada en su extremo amino terminal con el epítomo Flag y el carboxilo con el V5, se observó una banda con el tamaño correspondiente, 47 kDa aproximadamente (Figura 32).

En el revelado del gel con el anticuerpo V5, se observó como en los resultados del apartado anterior, que a partir de la isoforma canónica-V5 se obtenían dos bandas, mientras que de la variante 2-V5 sólo se obtenía una banda. En el caso de la isoforma canónica fusionada en el extremo N con el epítomo Flag y en el carboxilo con el epítomo

V5, al revelar con el anticuerpo V5 se obtienen dos bandas de peso molecular 47 kDa y 44 kDa aproximadamente (Figura 32). La banda de peso molecular mayor aparece en mayor cantidad que la de menor peso molecular, en la que se observa una banda de menor intensidad. La banda de 47 kDa correspondería a la proteína producida a partir del ATG situado delante del epítipo Flag mientras que la banda de 44 kDa correspondería a la proteína producida a partir del segundo ATG, lo que demostraría que este codón, junto con su secuencia Kozak, tendría la capacidad de iniciar la traducción por sí solos a pesar de tener otros codones de inicio de traducción anteriores.

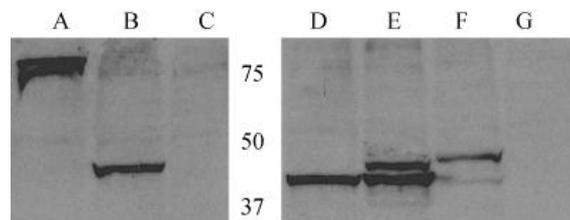


Figura 32. Análisis por Western Blot de la isoforma *NFκBIL1* clonada en un plásmido con epítipos en ambos extremos. Los carriles A, B y C del WB están revelados con un anticuerpo anti-Flag mientras que los carriles D-G están revelados con un anticuerpo anti V5. Carril A: Flag-LacZ, B: Flag-*NFκBIL1*-V5, C: Control negativo, D: *NFκBIL1_v2*-V5-His, E: *NFκBIL1*-V5-His, F: Flag-*NFκBIL1*-V5. Los pesos moleculares en kDa se encuentran en el centro de la imagen.

3.2 Inmunofluorescencias

Tras la observación de la expresión de diferentes proteínas de las isoformas de *NFκBIL1* se decidió realizar inmunofluorescencias para analizar la localización celular de los productos proteicos. Para ello, transfectamos células COS-7 con cada isoforma de *NFκBIL1* clonada en el vector pcDNA3.1/V5-His-TOPO, construcciones que habíamos usado previamente en las transfecciones y análisis por Western Blot. Utilizamos el anticuerpo que reconoce el epítipo V5 presente en el plásmido usado para clonar cada isoforma y Topro3 para teñir los núcleos celulares (Figura 33).

Una vez realizados los protocolos de inmunofluorescencias, lo que observamos fue un patrón de distribución nuclear para todas las isoformas. El punteado que aparece en los núcleos nos sugirió una distribución en espículas nucleares como se había

descrito previamente para la isoforma canónica (Semple et al., 2002). Las espículas nucleares son estructuras intracelulares enriquecidas en factores de splicing y están localizadas en regiones de intercromatina del núcleo de células de mamíferos. Aparecen en el microscopio de fluorescencia como estructuras puntuales e irregulares que varían en tamaño y forma. Para estudiar si todas las proteínas de NFκBIL1 se acumulan en estas estructuras, decidimos hacer estudios de colocalización con marcadores de espículas nucleares.

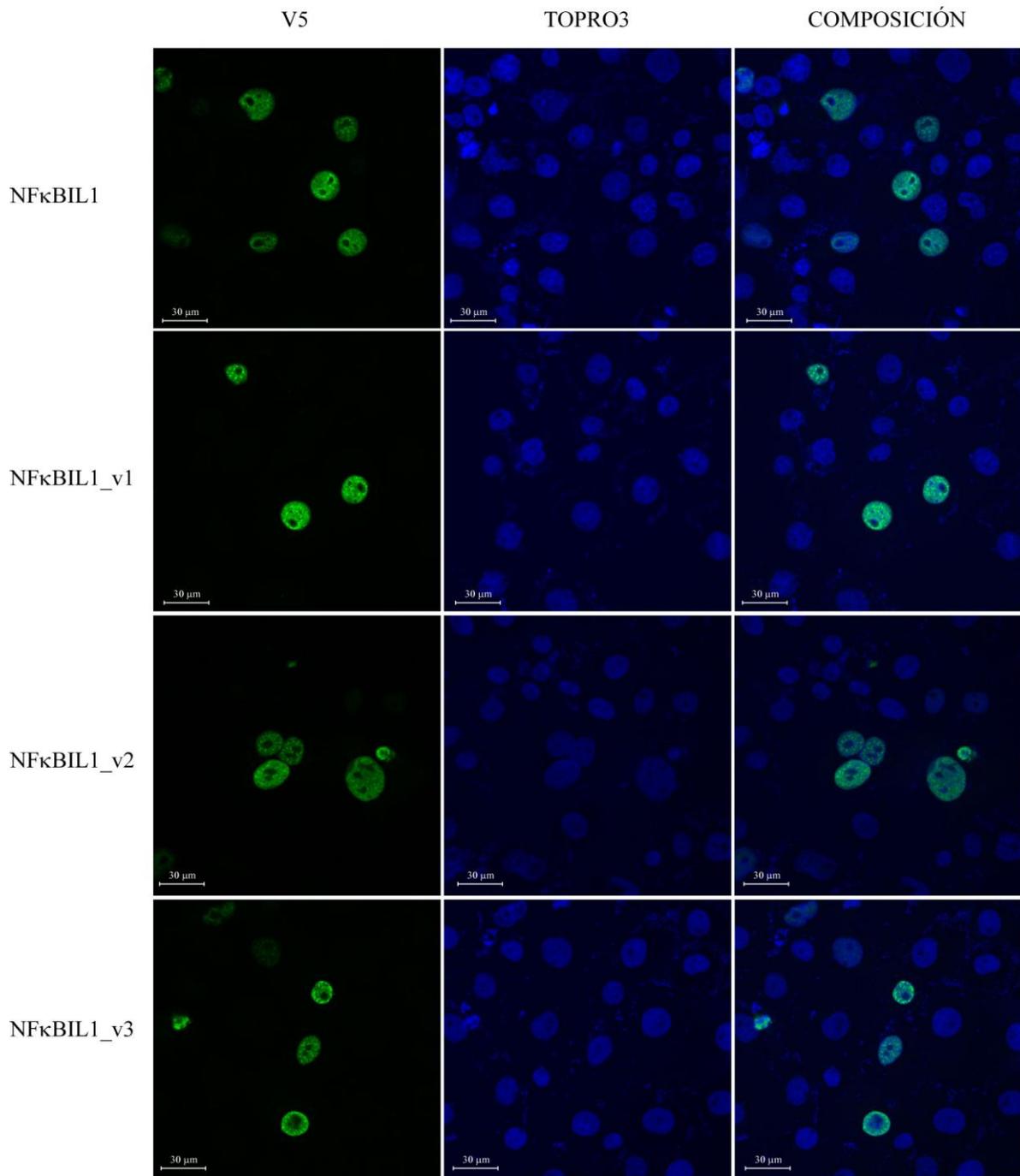


Figura 33. Inmunofluorescencias de cada isoforma de *NFκBIL1* en células COS-7. En los paneles de la izquierda se muestra cada isoforma detectada con el anticuerpo anti V5. En los paneles centrales se encuentran los núcleos teñidos con Topro3. En los paneles de la derecha se muestran las composiciones resultantes de superponer las figuras de la izquierda y centrales.

3.3 Co-localizaciones

Con el fin de estudiar la posible distribución en espículas nucleares de cada isoforma de *NFκBIL1* se realizaron inmunofluorescencias en células COS-7 transfectadas con cada isoforma. Se utilizó un anticuerpo anti SC-35, marcador de espículas nucleares, para el estudio de las co-localizaciones con las proteínas expresadas de cada isoforma. SC-35 pertenece la familia de proteínas SR (*serine/arginine rich proteins*) implicadas en el proceso de splicing alternativo. La mayoría de estas proteínas SR se localizan en las espículas nucleares aunque algunas sufren un proceso de transporte entre el núcleo y el citoplasma. SC-35 sólo se encuentra localizado dentro de las estructuras conocidas como espículas nucleares.

Lo que se observó tras estos experimentos, es que claramente todas las proteínas expresadas por cada variante de *NFκBIL1* colocalizaban con SC-35 en las mismas estructuras subnucleares, en las estructuras de espículas nucleares (Figura 34). Estos datos demuestran que las proteínas codificadas por las isoformas del gen *NFκBIL1* se encontraban en estas estructuras nucleares que están enriquecidas en proteínas relacionadas con el proceso de splicing del ARNm.

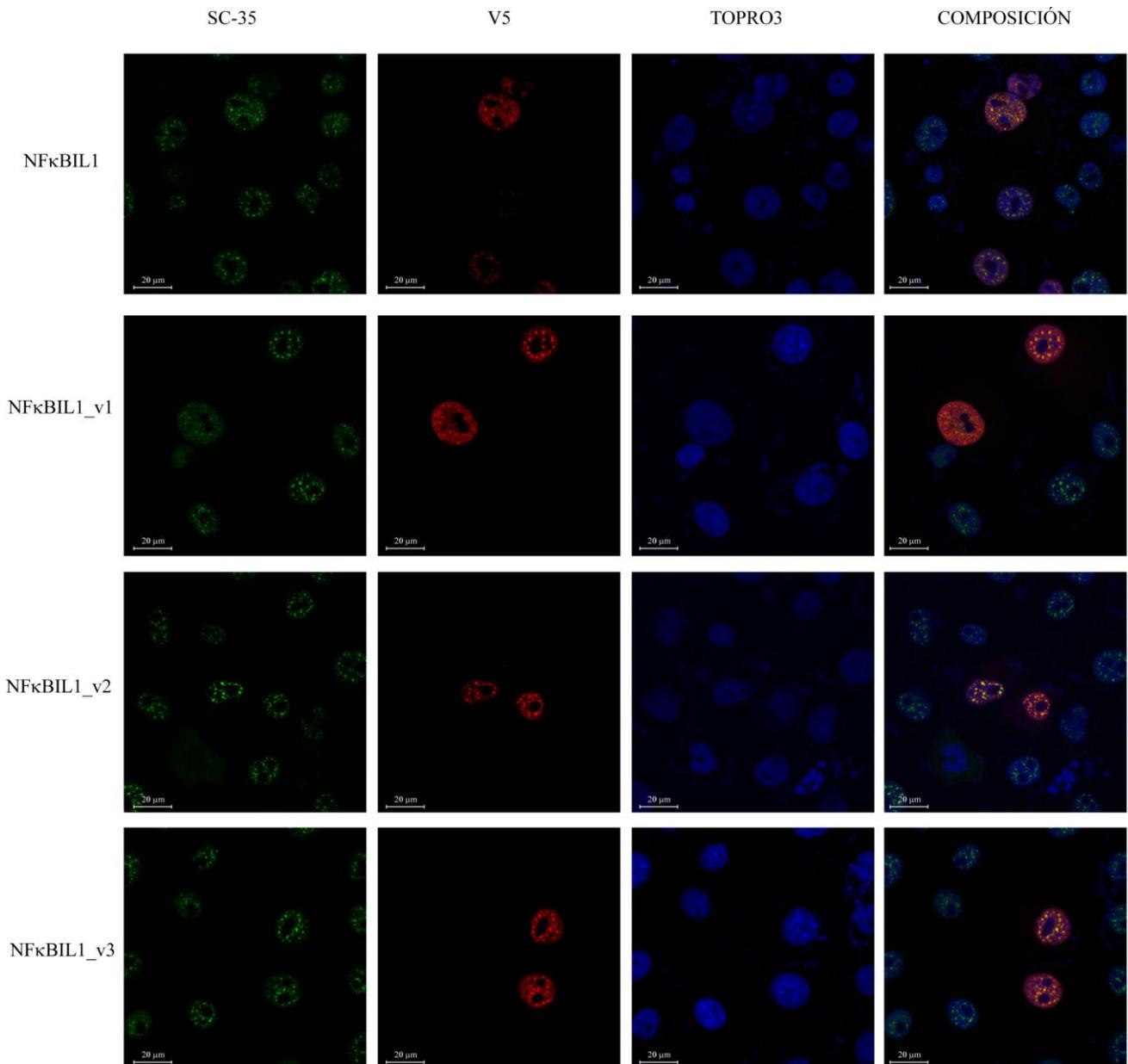


Figura 34. Distribución de las proteínas de *NFκBIL1* en espículas nucleares. Las células COS-7 fueron tratadas con una mezcla de anticuerpos anti-V5 y anti- SC35. En el panel de la izquierda se muestran las tinciones con el anticuerpo marcador de espículas, el SC35. En la segunda columna se muestra la distribución de cada isoforma, en la tercera los núcleos teñidos con ToPro3 y en la última la composición resultante de solapar los paneles.

4. ANÁLISIS DE LA EXPRESIÓN DE LAS ISOFORMAS DE *NFκBIL1* EN MUESTRAS DE PACIENTES CON ENFERMEDADES AUTOINMUNES Y CONTROLES.

El gen *NFκBIL1* se sitúa en una región del MHC de clase III donde mapean varias enfermedades relacionadas con el sistema inmune y la inflamación. Uno de nuestros objetivos consistía en analizar la expresión de las variantes de este gen en muestras de pacientes con enfermedades con las cuales se le ha relacionado, como AR, y observar posibles diferencias de expresión comparándolas con muestras de individuos sanos.

4.1 Estudio en pacientes con AR

Estudios previos demostraron una asociación del gen *NFκBIL1* con AR mediante análisis de polimorfismos o bien mediante interacciones con otros genes como *TNF* que estaban implicados en esta enfermedad. Lo que nos interesaba era analizar a nivel de cada isoforma de *NFκBIL1* qué ocurría en pacientes con AR. Se planteó la idea de, primero, ver la expresión de cada isoforma en individuos sanos y en pacientes con AR para ver si concordaban con la expresión observada en las líneas celulares y tejidos anteriormente estudiados. Y segundo, comparar los niveles de expresión de cada isoforma de *NFκBIL1* entre individuos sanos e individuos enfermos para comprobar si alguna de las isoformas estaba ausente o sobre-expresada en los pacientes con AR.

Para ver la expresión de cada isoforma en cada muestra, realizamos una cuantificación absoluta de las isoformas de *NFκBIL1* en diez muestras de sangre de pacientes con AR suministradas por el Hospital de Cantoblanco (Madrid) y en diez muestras de sangre de controles voluntarios sanos, suministradas por el Hospital Universitario la Paz (Madrid), mediante el ensayo de PCR a tiempo real descrito anteriormente. Las muestras de sangre fueron tratadas con Ficoll para quedarnos sólo con la fracción de PBMCs.

El ensayo de PCR a tiempo real nos permitía de una manera específica cuantificar cada isoforma de *NFκBIL1*. Los gráficos (Figura 35) muestran que la

isoforma canónica *NFκBIL1* era la más expresada en los individuos sanos y en pacientes, tal como ocurre en las líneas celulares y tejidos analizados para el estudio de la expresión de este gen. En los controles, las isoformas *NFκBIL1_v2* y *NFκBIL1_v3* presentaban una expresión muy baja con un porcentaje de expresión de 1.52 y 1.76 respectivamente. Lo mismo ocurría en pacientes con AR pero la isoforma *NFκBIL1_v3* parece tener una expresión superior a la variante 2. Como ocurría en el caso de todos los tejidos y líneas celulares analizadas, las isoformas canónica y variante 1 eran las de mayor expresión mientras que las variantes 2 y 3 se expresaban en menor proporción.

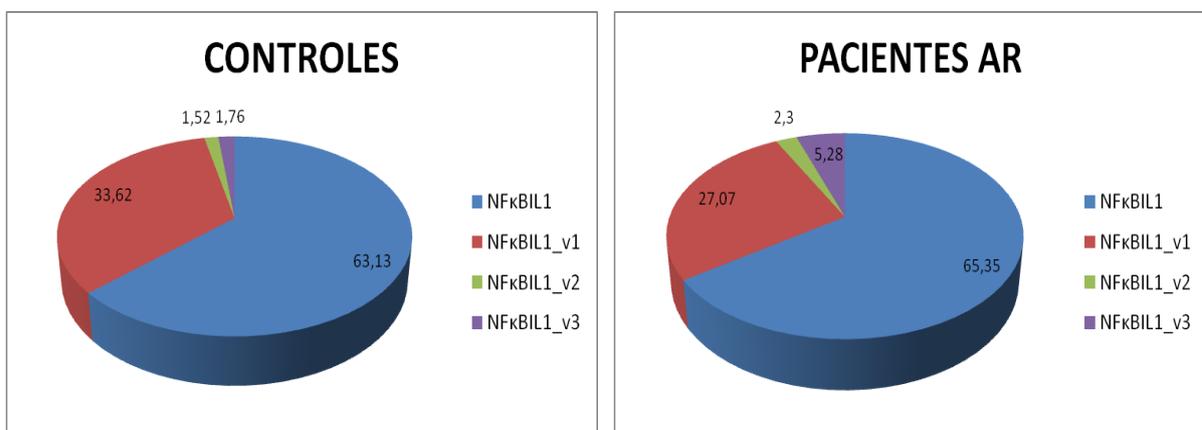


Figura 35. Cuantificación absoluta de las isoformas de *NFκBIL1* en pacientes con artritis y en individuos sanos. Los gráficos muestran los porcentajes medios de la expresión de cada isoforma en el grupo de individuos sanos y en el grupo de pacientes con AR.

Para comparar los niveles de expresión de las isoformas de *NFκBIL1* entre los controles y los pacientes con AR, realizamos una cuantificación relativa utilizando GAPDH y β -actina como genes normalizadores. Estos genes de referencia se utilizan para la cuantificación relativa puesto que su expresión permanece constante en las muestras estudiadas y no son afectados por el tratamiento o enfermedad bajo estudio. El ensayo de PCR a tiempo real que se utilizó fue el mismo que en el caso anterior y las mismas muestras de pacientes con AR y controles. La cuantificación relativa permite una comparación de la expresión de un gen entre diferentes grupos de muestras. Tras comparar la expresión de *NFκBIL1* entre individuos sanos y pacientes con AR, los resultados mostraron una sobre-expresión de las cuatro isoformas de *NFκBIL1* en pacientes con AR (Figura 36). Tras realizar los análisis estadísticos, un análisis de T-test

con el software GenEx, se encontraron diferencias significativas en la cuantificación relativa de las isoformas *NFκBIL1*, *NFκBIL1_v1* y *NFκBIL1_v3*. Estos datos reflejaban una sobre-expresión de las isoformas de *NFκBIL1* en la sangre de pacientes con AR comparándolos con los pacientes sanos.

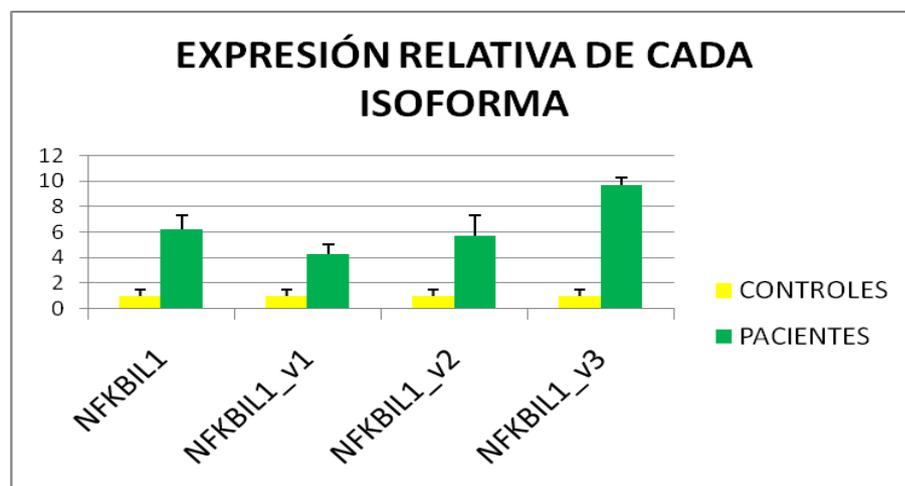


Figura 36. Expresión relativa de cada isoforma de *NFκBIL1* comparado pacientes con AR y controles. Los valores de expresión de cada isoforma en los controles se consideró como valor 1.

4.2 Estudio en pacientes con Lupus Eritematoso Sistémico

Se ha descrito que un polimorfismo en la región promotora de *NFκBIL1* está asociado con un descenso en el riesgo de adquirir Lupus eritematoso sistémico (Castiblanco and Anaya, 2008). Esta enfermedad crónica es de carácter autoinmune que afecta al tejido conjuntivo, provocando inflamación y daño de tejidos. Puesto que en pacientes con AR se observó una sobre-expresión de todas las isoformas del gen *NFκBIL1*, decidimos realizar el mismo estudio en muestras de pacientes con otra enfermedad autoinmune relacionada con este gen, para comprobar si era una respuesta generalizada a todos los procesos autoinmunes o era algo propio de la enfermedad de artritis.

Realizamos el estudio en ocho muestras de sangre de pacientes con Lupus suministradas por una colaboración con el Hospital Ramón y Cajal (Madrid) y en ocho muestras de sangre de individuos sanos, suministradas por el Hospital Universitario la

Paz (Madrid). Las muestras de sangre se trataron de la misma manera que las del apartado anterior, realizando también un Ficoll para quedarnos con la fracción de PBMCs. Cuantificamos los niveles de expresión relativa de cada isoforma con el diseño de PCR a tiempo real diseñado para el análisis de expresión de *NFκBIL1* en muestras de pacientes con AR y en individuos sanos. Al comparar los niveles de expresión de cada isoforma entre los pacientes con Lupus y los controles, se observó que no había diferencias de expresión entre ambos grupos, no había sobre-expresión ni disminución de la expresión de ninguna de las isoformas de este gen, no como ocurría en el caso de pacientes con AR. Luego la observación de la sobre-expresión de las isoformas de *NFκBIL1* sólo la encontramos en pacientes con AR.

5. ANÁLISIS DE LA REGIÓN PROMOTORA DE *NFκBIL1*

En estudios previos de polimorfismos en el promotor de *NFκBIL1* en pacientes con AR, los resultados eran contradictorios según el grupo que los presentaba, manifestando unos (Okamoto et al., 2003) que los polimorfismos en esta región estaban asociados a AR y otros (Collado et al., 2004) que no había tal asociación. En nuestro caso decidimos realizar un estudio de SNPs y posibles mutaciones en la región promotora de *NFκBIL1*. Además se realizó un análisis del patrón de metilación de esta región puesto que cambios en la metilación en la región promotora pueden dar lugar a cambios en la expresión del gen. El estudio la zona promotora del gen *NFκBIL1* se realizó en el Sanger Institute (Hinxton, Cambridge, UK) en colaboración con el Profesor Stephan Beck en una estancia de tres meses realizada de Mayo a Julio del año 2007.

5.1 Análisis de polimorfismos y mutaciones.

Se analizó el fragmento de 833 pb de longitud del gen *NFκBIL1* desde las posiciones -656 hasta la +147 del sitio de inicio de transcripción (TSS, Transcription Start Site) de la isoforma canónica. Esta amplia región comprende la zona promotora descrita en el Ensembl para *NFκBIL1*. Esta zona fue amplificada mediante PCR usando los oligos Promotor Sentido y Promotor Antisentido. Amplificamos el promotor de *NFκBIL1* en ADN extraído de ocho pacientes con AR y en ocho voluntarios sanos para

analizar los polimorfismos y posibles mutaciones en esta región que afectaran a la expresión de los transcritos. Después de la amplificación y secuenciación de la región promotora, se observó la presencia de un polimorfismo en la posición – 62A/T en las muestras analizadas. Tres de los pacientes con AR y tres de los controles sanos presentaban el alelo T. Este polimorfismo tiene unas frecuencias alélicas en Europa de A: 0.64 y T: 0.36, lo cual concordaba con nuestros resultados. Puesto que este polimorfismo se encontraba presente tanto en pacientes con AR como en individuos sanos y no se observaron evidencias significativas entre ambos grupos, no se pudo concluir que este polimorfismo estuviera asociado a AR.

5.2 Análisis de metilación.

La variabilidad epigenética es otro mecanismo de regulación de la expresión génica y puede ser relevante en la patogénesis y severidad de las enfermedades. Hay dos mecanismos epigenéticos principales, la modificación post-traducciona l de las histonas en la cromatina y la metilación del ADN. Se cuestionó si el nivel de metilación del ADN en el promotor de *NFκBIL1* podría ser relevante en la expresión de sus isoformas y podría ser un factor que contribuyera a la patogénesis inflamatoria de AR.

En un estudio publicado por Nile y colaboradores (Nile et al., 2008) se observó un patrón alterado de metilación en ADN genómico del gen *IL-6* (interleucina 6) de pacientes con AR. Examinaron el estado de metilación de los sitios CpGs en la región reguladora 5' de este gen y encontraron un sitio CpG no metilado en todos los pacientes con AR que contribuía a su elevada expresión.

Como la metilación del ADN ocurre principalmente en las islas CpGs, decidimos analizar la metilación de estos dinucleótidos. En la región promotora del gen *NFκBIL1* descrita en Ensembl, nos encontramos con 31 sitios CpG. Realizamos un análisis detallado de la metilación de los sitios CpG en la región promotora del gen *NFκBIL1* en 8 muestras de controles y 8 pacientes con AR, las mismas muestras donde se había realizado el análisis de polimorfismos. Para determinar la metilación de los sitios CpG de la región promotora del gen *NFκBIL1*, se analizó la región del promotor A conteniendo 18 CpGs, la región B conteniendo 4 sitios CpGs y la región C conteniendo 9 sitios CpGs mediante la secuenciación del ADN genómico tratado con bisulfito

(Figura 37). La región A, de 380 pb de longitud, se amplificó con los oligos Bis-A sentido y antisentido, la región B, de 278 pb de longitud, fue amplificada usando los oligos Bis-B sentido y antisentido y la región C con los oligos Bis-C sentido y antisentido.

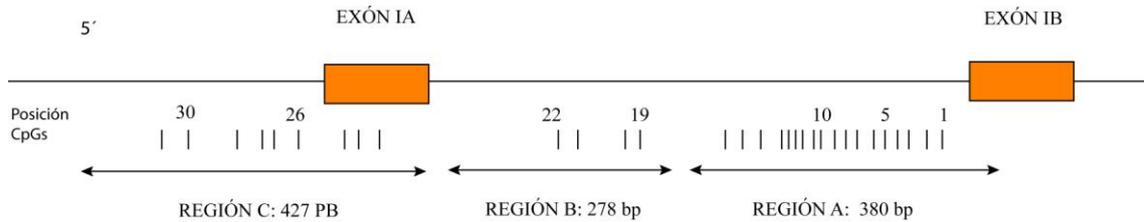


Figura 37. Mapa de los sitios CpG de la región promotora de *NFκB1*. Las líneas verticales muestran las posiciones de los sitios CpG en la región promotora. Las regiones A,B y C contienen 18, 4 y 9 sitios CpG respectivamente.

Amplificamos estas regiones en el ADN de pacientes con AR y en controles. El ADN estaba tratado con bisulfito puesto que con este procedimiento, las citosinas se transformaban en uracilo a no ser que estuvieran protegidas por la metilación. Los uracilos no protegidos se leían como timidinas durante la PCR, y la secuenciación entonces permitía la identificación de las citosinas metiladas.

Tras la secuenciación, observamos que estas regiones de la zona promotora del gen *NFκB1* se encontraban predominantemente sin metilar como ocurre en los genes transcripcionalmente activos. La única isla CpG metilada que se comprobó en el análisis se encontró en la posición 5 de la región A del DNA de uno de los pacientes con AR, dando lugar a diferencias no significativas entre pacientes con AR e individuos sanos (Figura 38). Este cambio en un sólo sitio CpG del promotor del gen *NFκB1* en un único paciente con AR no es un hecho significativo, luego otros mecanismos diferentes a los cambios de metilación están implicados en las diferencias de expresión del gen *NFκB1* en pacientes con AR y los individuos sanos.

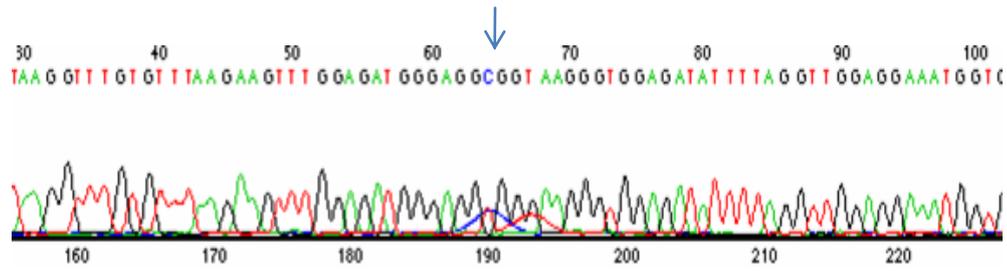


Figura 38. Secuenciación de ADN tratado con bisulfito de un paciente con AR. En la figura se muestra una Citosina metilada que permanece como tal después del tratamiento con bisulfito.

La complejidad del transcriptoma de los mamíferos excede, de lejos, lo que se esperaba hace décadas. Poco a poco, se han ido identificando mecanismos que permiten explicar cómo organismos fenotípicamente tan diferentes poseen un número similar de genes en su genoma. Uno de los mecanismos que influye en esta complejidad y diversidad fenotípica es el splicing alternativo. En humanos, se estima que el 95% de los genes sufren este proceso dando lugar a distintas variantes del mismo gen. En este estudio, hemos caracterizado las variantes de splicing del gen *NFκBIL1* en distintas especies y hemos analizado la expresión y localización celular de las distintas isoformas en humanos. Como además el gen *NFκBIL1* se encuentra en el MHC, la región más importante en el genoma de vertebrados con respecto a la infección y autoinmunidad, se ha tratado de analizar la relación de este gen con varias enfermedades autoinmunes en las que hay evidencias de su implicación.

1. CARACTERIZACIÓN DE *NFκBIL1* EN HUMANOS Y EN OTRAS ESPECIES

La región del MHC se encuentra altamente conservada en todos los mamíferos y estudios recientes muestran que existe también una alta conservación en reptiles, aves y anfibios (Hurt et al., 2004; Renard et al., 2006). El contexto genómico del gen *NFκBIL1* está altamente conservado mostrando una gran similitud en las especies estudiadas. El hecho de tener la secuencia altamente conservada en mamíferos, junto con su localización dentro del MHC y su potencial asociación con AR, sugieren que *NFκBIL1* pueda tener una función conservada en inmunidad.

En humanos se habían caracterizado previamente dos isoformas del gen *NFκBIL1* llamadas α y β , siendo la variante β 45 pb más corta que la isoforma α (Semple et al., 2002). En nuestro estudio, hemos identificado nuevas variantes de *NFκBIL1* mediante análisis informáticos y posteriores comprobaciones con ensayos de RACE y PCR. Las nuevas variantes identificadas se diferencian en los extremos 5'UTR siendo el exón I diferente a las isoformas descritas anteriormente. Por lo tanto, *NFκBIL1* expresa cuatro isoformas en humanos a las que hemos denominado *NFκBIL1*, *NFκBIL1_v1*, *NFκBIL1_v2* y *NFκBIL1_v3*.

Uno de nuestros objetivos era comprobar cómo el proceso de splicing alternativo contribuye a la diversificación transcriptómica y proteómica. Por ello decidimos comparar cómo se expresaba *NFκBIL1* en diferentes especies. Estudios previos tratando de comparar la tasa de splicing alternativo en diferentes organismos, han proporcionado resultados contradictorios. Brett y colaboradores (2002) en un análisis a gran escala de ESTs, estimaron que la tasa de splicing alternativo era similar en una variedad de organismos analizados entre los que se encontraban humanos y nematodos. Por el contrario Kim y colaboradores (2004) estimaron que la tasa de splicing en mamíferos era muy superior a la de invertebrados. Sin embargo, ambos estudios dependían de la cobertura de las ESTs, que variaba mucho entre los organismos estudiados.

En nuestro estudio no sólo hemos analizado las ESTs y formas descritas en las bases de datos para cada especie, si no que hemos intentado amplificar específicamente las variantes encontradas en humanos, u otras diferentes experimentalmente. En algunas especies que analizamos, sólo encontramos la isoforma canónica de *NFκBIL1* como en *Mus musculus*, *Rattus norvegicus*, *Sus scrofa* y *Cebus apella*. El gen *NFκBIL1* expresaba cuatro variantes de splicing en *Homo sapiens*, *Pongo pygmaeus*, *Gorilla gorilla* y *Papio sp.* En *Macaca fascicularis* además de estas cuatro isoformas, se encontraba en sangre otra isoforma nueva característica por la retención de parte de un intrón entre los exones II y III que provocaba un cambio en la fase de lectura. Esta isoforma también se encontró en *Macaca mulata*. En *Colobus guereza* se encontraron sólo dos variantes de *NFκBIL1*, las correspondientes a las isoformas humanas canónica y variantes 1.

Con estos resultados, podemos ver que el proceso de splicing de *NFκBIL1* en primates superiores da lugar a las mismas isoformas. En otros organismos como ratón, rata y el mono capuchino, sólo se expresa la isoforma canónica y en otras especies como las del género *Macaca* aparecen variantes nuevas de este gen, además de las encontradas en humanos.

2. EXPRESIÓN Y LOCALIZACIÓN DE LAS ISOFORMAS DE *NFκBIL1* EN HUMANOS

Los datos de los ensayos de PCR a tiempo real en distintos tejidos y líneas celulares humanas mostraron cómo la isoforma canónica *NFκBIL1* es la que más se expresaba de todas las isoformas, seguida de la variante 1. Las variantes 2 y 3, con distinto exón I presentaban muy baja expresión en todas las muestras estudiadas siendo la isoforma 3 la de menor expresión en todos los tejidos y líneas celulares llegando en algunos tejidos a no expresarse. En un estudio anterior (Greetham et al., 2007) en el que se realizaban ensayos de PCR a tiempo real para cuantificar las isoformas *NFκBIL1* y *NFκBIL1_v1* observaban que, en algunos tejidos, las dos isoformas se expresaban por igual, en otros la canónica predominaba y en otros la variante 1. Estudiando el diseño experimental realizado por Greetham y colaboradores (2007), encontramos varias diferencias con el nuestro. En primer lugar, ellos sólo describían dos isoformas de *NFκBIL1* mientras que en este estudio hemos caracterizado 4 isoformas de este gen. Pero lo más significativo es que el diseño de PCR a tiempo real que realizaron con una sonda Taqman supuestamente específica para cada isoforma no fue tal, puesto que la sonda diseñada para la variante 1 reconocía una zona común de las dos isoformas, canónica y variante 1, con lo que no se pueden utilizar los resultados de sus ensayos para compararlos con los obtenidos en nuestro estudio de la expresión de las cuatro isoformas de *NFκBIL1*.

El hecho de que las nuevas isoformas caracterizadas, *NFκBIL1_v2* y *NFκBIL1_v3*, comiencen con un exón I distinto al que estaba caracterizado para las otras dos variantes, puede indicar que se empleen dos promotores en la transcripción puesto que la variación en el primer exón se considera como una manifestación del uso de promotores alternativos (Kreth et al., 2008). El hecho de usar diferentes promotores no sólo controla los niveles de transcripción de las correspondientes isoformas de ARNm si no también contribuye a la eficiencia de traducción. En todas las muestras y tejidos estudiados, las variantes 2 y 3 presentan una baja expresión en comparación con las otras isoformas, lo que podría indicar que están sujetas a un promotor menos eficiente o que se activa bajo estímulos o condiciones distintas.

Otro objetivo fue estudiar las proteínas a las que daba lugar cada isoforma del gen *NFκBIL1*. Puesto que los anticuerpos comerciales disponibles reconocían a todas las isoformas sin poder diferenciarlas, se decidió transfectar cada una de las variantes unidas a epítomos V5 e Histidinas en el extremo carboxilo terminal, en células COS-7 y ver su expresión y localización. Como se observó en los resultados, las variantes 2 y 3 daban lugar a las proteínas correspondientes según los estudios del ORFinder y viendo cuál era el inicio de traducción más adecuado. Sin embargo, a partir de las variantes canónica y 1 aparecían dos bandas de distinto peso molecular que corresponderían a dos proteínas. Este hecho sorprendente, ya fue descrito previamente para la variante canónica (Semple et al., 2002). Analizando los posibles inicios de traducción, se observó que en estas dos variantes había en el exón IB un codón ATG rodeado de una secuencia Kozak poco adecuada para el inicio de la traducción, mientras que en el segundo exón se encontraba un codón ATG rodeado de una secuencia Kozak adecuada. Según Semple y colaboradores (2002), este hecho podría explicar la aparición de dos bandas al expresar estas isoformas. Además, otro hecho sorprendente era que la isoforma de mayor peso molecular, la que correspondería al inicio de traducción del ATG en el exón IB aparecía en menor cantidad que la que empezaría en el exón II. Este hecho podría explicarse debido a que en el exón I se encuentran secuencias PEST que son las responsables de dirigir a la proteína a su degradación en el proteosoma, lo que podría explicar que la proteína de mayor peso molecular estuviera en menor cantidad (Rogers et al., 1986). Estas secuencias PEST podrían afectar a la regulación post-traducciona de la expresión de *NFκBIL1*.

Semple y colaboradores (2002) demostraron que al expresar una construcción en un vector con la isoforma canónica *NFκBIL1* en la que se incluían 44 pb del extremo 5'UTR endógeno, sólo se producía una banda del tamaño de la proteína con menor peso molecular, que empezaría tras el segundo ATG. Este extremo 5'UTR que rodea al primer ATG, tiene un gran contenido en GC, y basados en el modelo de traducción de escaneado del ribosoma, este hecho podría provocar que el inicio de la traducción no fuera posible en ese primer codón ATG.

También existía la posibilidad de que se produjera un proceso proteolítico en el extremo amino terminal por lo que realizamos una construcción de la isoforma

NFκBIL1 con un epítopo Flag en el extremo N terminal y un epítopo V5 en el extremo carboxilo. Realizamos las transfecciones de la misma manera que con las otras construcciones y revelando con el anticuerpo anti Flag se observó la presencia de una sola banda que correspondía a la proteína de mayor longitud (desde el primer ATG). Cuando revelamos con un anticuerpo anti V5 se observaron 2 bandas, una mayoritaria correspondiente a la proteína que se obtendría desde el primer ATG colocado antes del epítopo Flag y otra banda de menor intensidad que correspondería a la proteína de menor peso molecular, traducida a partir del ATG del segundo exón. Esto nos confirmaría que el segundo ATG está rodeado por una secuencia Kozak que favorece el inicio de traducción aún cuando hay ATGs que le preceden.

Una vez que se analizó la expresión de proteínas, se pasó al estudio de la localización de cada una de ellas. La localización de la isoforma canónica se había realizado previamente en células transfectadas (Semple et al., 2002) y se había observado una localización nuclear en unas estructuras conocidas como espículas nucleares. Puesto que en este estudio se describían variantes nuevas del gen *NFκBIL1*, se estudió la localización de esas nuevas variantes para comprobar si diferían de la localización de la canónica. Se transfectaron células COS-7 con plásmidos conteniendo cada isoforma y se realizaron ensayos de inmunofluorescencia. Lo que se observó, fue que todas las isoformas presentaban un patrón de distribución nuclear. El patrón que se observaba era un patrón punteado que recordaba al que se observa en las espículas nucleares. Puesto que se había observado previamente que la isoforma canónica se encontraba en estas estructuras, se realizó una colocalización con el factor de splicing SC35 presente en las espículas. El resultado fue que todas las isoformas colocalizaban con el factor SC35 lo que demostró que la localización de todas las proteínas se encontraban en esas estructuras. En las espículas nucleares, se acumulan muchos factores que intervienen en el proceso de splicing, luego el hecho de que *NFκBIL1* se encontrara en las espículas podría sugerir que este gen interviniera de alguna manera en el proceso de splicing.

Tanto las repeticiones de anquirina como la señal de localización nuclear presente en el extremo amino de *NFκBIL1* son dominios necesarios para su distribución en las espículas nucleares (Semple et al., 2002). Otro gen conocido como *spergen-2*,

que consta de 5 repeticiones de anquirina y una señal de localización nuclear en su extremo amino, parece localizarse también en las espículas nucleares como se ha visto mediante estudios de microscopía confocal (Iida et al., 2003). Se han identificado otros dominios diferentes responsables de dirigir a las proteínas a las espículas nucleares: los motivos de reconocimiento de ARN del factor de splicing asociado a PTB (PSF) y de los factores de splicing ricos en serina/arginina (Caceres et al., 1997; Dye and Patton, 2001); el dominio Sm de la pequeña ribonucleoproteína nuclear U1 (Malatesta et al., 1999) y el dominio rico en treonina/prolina del factor de splicing SF3b (Davuluri et al., 2008). El hecho de que NFκBIL1 se encuentre en estas estructuras donde se acumulan factores de splicing podría indicar un papel en la regulación de este proceso.

IκB-ζ, uno de los miembros no “clásicos” de la familia de inhibidores IκB, también presenta una distribución nuclear acumulándose en las espículas nucleares (Seshadri et al., 2009). Esta proteína tiene curiosamente en su extremo N terminal una secuencia de localización nuclear y un dominio de transactivación mientras que en su extremo C terminal consta de 6 repeticiones de anquirina. Se induce por LPS y regula la producción de IL-6 en monocitos humanos. Además es capaz de unirse a las subunidades p50 y p65 del factor NF-κB. El gen *NFκBIL1* comparte con este miembro de la familia de inhibidores la posesión en su extremo amino de una señal de localización nuclear, las repeticiones de anquirina aunque en el caso de *NFκBIL1* son sólo 3 y su distribución en espículas nucleares. Estudios funcionales realizados con la isoforma canónica NFκBIL1 han mostrado que este gen no se une a las subunidades p50 ni p65 del factor nuclear NF-κB (Greetham et al., 2007). En cambio, se ha visto que puede asociarse a la proteína EF1α, un receptor de leucocitos, implicado en la traducción de ARNm y se cree que es un posible antígeno del síndrome Felty, caracterizado por ser una enfermedad sistémica de algunos pacientes con AR (Ditzel et al., 2000). El hecho de que la proteína NFκBIL1 no se una a las proteínas p50 y p65 del factor nuclear NF-κB no implica que no pueda unirse a otras subunidades de este factor, para lo cual se necesitarían realizar más ensayos de co-inmunoprecipitación con el resto de subunidades de la familia NF-κB. Otra posibilidad es que NFκBIL1 pueda unirse a NF-κB bajo determinadas condiciones o estímulos adecuados.

3. ANÁLISIS DE LA EXPRESIÓN DE *NFκBIL1* EN MUESTRAS DE PACIENTES CON AR Y LUPUS

Tras caracterizar las distintas isoformas del gen *NFκBIL1*, cuantificar su expresión en distintas líneas celulares y tejidos y describir su localización, procedimos al estudio de la expresión de las variantes en muestras de pacientes con AR y Lupus. Para ello, se utilizó el mismo diseño de PCR a tiempo real utilizado en las cuantificaciones anteriores.

Primero se analizaron las muestras de sangre de pacientes con AR y de donantes sanos. Realizando un Ficoll, nos quedamos con la fracción de PBMCs, de los cuales extrajimos el ARN y sintetizamos el ADNc para su estudio. Se realizó una cuantificación absoluta para analizar el número de copias de cada isoforma en las muestras. Los resultados mostraron que la isoforma canónica *NFκBIL1* era la de mayor expresión en cada muestra, lo que concordaba con los datos obtenidos en las células y tejidos estudiados. De la misma manera, las variantes 2 y 3 son las de menor expresión como ocurría en los análisis previos.

Posteriormente, se realizó una cuantificación relativa para poder comparar los niveles de expresión de cada isoforma entre el grupo de pacientes y el grupo de voluntarios sanos. Lo que se observó es que todas las variantes de *NFκBIL1* estaban sobre-expresadas en los pacientes con AR. En un estudio previo realizado por Greetham y colaboradores (2007) realizaron también una cuantificación de la expresión de las variantes canónica y 1 en PBMCs de tres pacientes con AR y tres individuos sanos. No observaron diferencias de expresión pero como hemos comentado anteriormente, la sonda que diseñaron no era específica de la variante 1 y reconocía las dos isoformas. Además, una vez aislados los PBMCs, los crecieron en cultivo y activaron con LPS, luego el ensayo estaba realizado en unas condiciones distintas al nuestro.

La AR es una enfermedad inflamatoria de compleja etiología, cuyas causas se atribuyen a factores genéticos y ambientales. Durante décadas se han identificado multitud de genes relacionados con esta enfermedad, hecho que puede deberse a las interacciones entre genes, con lo que los análisis se vuelven más complicados. Estudios recientes de interacciones de genes asociados con AR (Huang et al., 2009) muestran que

podría haber una relación entre NFκBIL1 y TNF-α, una citoquina proinflamatoria presente en pacientes con AR, la cual estimula el factor NF-κB provocando la expresión de genes que intervienen en la respuesta inflamatoria, con lo cual la vía de señalización de este factor está implicada en la patogénesis de muchas enfermedades inflamatorias (Pasparakis, 2009).

La presencia de las proteínas NFκBIL1 en las espículas nucleares, reservorios de factores de splicing, puede indicar un papel en el proceso de regulación de splicing. También se ha demostrado que puede unirse a ARNm y proteínas relacionadas con la transcripción, procesamiento y traducción del ARNm, entre las que se encuentra EF1α que como hemos comentado anteriormente se asocia a AR. El hecho de que las isoformas de NFκBIL1 estén sobre-expresadas en pacientes con AR podría indicar que está favoreciendo la transcripción de genes implicados en esta patología. Puesto que NFκBIL1 tiene cierta homología con miembros no “clásicos” de la familia de inhibidores IκB, podría ser un potencial regulador de NFκB actuando como supresor o como activador de este factor uniéndose a subunidades diferentes a las que se han usado anteriormente en ensayos de co-inmunoprecipitación en el núcleo y provocando la transcripción de genes relacionados con la patología de AR.

También se realizó un estudio de la expresión de *NFκBIL1* en muestras de pacientes con Lupus y en voluntarios sanos, observando que no había diferencias significativas entre ambos grupos, lo que indicaba que la sobre-expresión de los mensajeros de *NFκBIL1* en pacientes con AR no era una respuesta generalizada en pacientes con otras enfermedades autoinmunes.

4. ANÁLISIS DE LA REGIÓN PROMOTORA DE *NFκBIL1*

Polimorfismos en la región promotora de *NFκBIL1* se han asociado a enfermedades como la diabetes tipo I (Yamashita et al., 2004) y se han asociado al posible desarrollo de cardiomiopatías crónicas de Chagas entre los individuos infectados por *Trypanosoma cruzi* (Ramasawmy et al., 2008) y al posible riesgo de desarrollar lupus eritematoso sistémico y el síndrome de Sjögren (Castiblanco and

Anaya, 2008). Asociaciones de polimorfismos de este gen con AR han generado mucha controversia. Okamoto y colaboradores (2003) demostraron que *NFκBIL1* es un gen de susceptibilidad para el desarrollo de AR en pacientes japoneses mientras que un resultado similar no pudo ser encontrado en pacientes españoles (Collado et al., 2004; Martinez et al., 2003).

En este estudio realizamos el análisis de la región promotora de *NFκBIL1* en pacientes con AR y en individuos sanos con el fin de encontrar posibles polimorfismos o mutaciones. Los resultados mostraron la presencia de un polimorfismo en la posición -62A/T de *NFκBIL1* pero no se encontraron diferencias significativas entre pacientes con AR y controles. Nuestros resultados concuerdan con los obtenidos por otros grupos en análisis de polimorfismos del promotor de *NFκBIL1* en un estudio español de casos-controles (Collado et al., 2004; Martinez et al., 2003) en los que concluyen una falta de asociación de este polimorfismo con AR. Estos resultados contradicen con los encontrados por Okamoto y colaboradores en los que demostraban la asociación del polimorfismo -62A/T de *NFκBIL1* con AR (Okamoto et al., 2003). Los efectos de heterogeneidad genética y de poblaciones tienen que considerarse a la hora de detectar genes de susceptibilidad para AR en diferentes poblaciones. El hecho de que aparezcan diferencias dependiendo de los estudios poblacionales, puede ser debido a diferentes historias en la evolución de las poblaciones o por interacciones de los polimorfismos del gen *NFκBIL1* con exposiciones ambientales. El riesgo de enfermedad también puede verse afectado por otros polimorfismos o por efectos de haplotipo creados por alelos de *NFκBIL1* que difieran entre poblaciones caucásicas y japonesas.

Como muestran nuestros resultados, encontramos altos niveles de *NFκBIL1* en sangre de pacientes con AR en comparación con individuos sanos, hecho que no se ha podido explicar con mutaciones o polimorfismos en la zona promotora de este gen. Como se ha comentado previamente, la epigenética, que se define como cambios hereditarios en los patrones de expresión génica que no están causados por cambios en la secuencia de ADN primaria, se ha convertido en un área nueva de investigación en las enfermedades reumáticas. La epigenética juega cada vez un papel más importante en la respuesta inflamatoria e inmune (Lee et al., 2006; Su and Tarakhovsky, 2005). Existe un número de rasgos de AR que se han propuesto como consecuencia de cambio

epigenéticos, incluyendo la discordancia en gemelos monocigóticos (70%) y el desarrollo tardío de la enfermedad (en las décadas cuarta y quinta) (Petronis, 2001).

Dentro de los cambios epigenéticos, la metilación en los dinucleótidos CpG de las islas ricas en CpG es el mecanismo más estudiado. Las islas CpG son regiones ricas en dinucleótidos CpG que se encuentran en las regiones promotoras principalmente. La densidad de metilación es baja en genes transcripcionalmente activos y alta en los silenciados (Egger et al., 2004). En un estudio previo del gen *IL6* (Nile et al., 2008) se ha demostrado una relación entre la metilación de un único sitio CpG con AR. Lo que observaron, fue que la hipometilación de un único sitio CpG en la zona promotora del gen *IL6* en pacientes con AR afectaba su regulación provocando la unión de proteínas al ADN y produciéndose altos niveles de ARNm en pacientes comparados con controles. Cada vez hay más evidencias de la importancia del estado de metilación de las secuencias CpG en la regulación de la expresión génica.

Viendo las diferencias de expresión encontradas en el gen *NFκB1* entre pacientes y controles, nos cuestionamos si podría deberse también a cambios en la metilación del ADN en la zona promotora de este gen y podría ser un posible factor que contribuyera a la patogénesis de AR. Para ello, realizamos un análisis de los sitios CpG de la zona promotora de este gen en pacientes y controles. No encontramos diferencias significativas en la metilación de esta zona cuando comparamos los dos grupos. Esto nos sugirió que otros mecanismos diferentes a la metilación estaban implicados en las diferencias de expresión encontradas. Sería interesante estudiar los cambios epigenéticos en otros genes implicados en esta enfermedad puesto que cambios en la metilación de las islas CpG se ha visto que están asociados con AR y otras enfermedades de carácter inflamatorio (Trenkman et al., 2009).

CONCLUSIONES:

1. El gen *NFκBIL1* presenta en humanos cuatro variantes debidas al proceso de splicing alternativo usando un sitio 5' de splicing alternativo en el exón III o bien mediante un uso del exón I alternativo. Estas cuatro isoformas son *NFκBIL1*, *NFκBIL1_v1*, *NFκBIL1_v2* y *NFκBIL1_v3*.
2. La forma de splicing *NFκBIL1* está muy conservada en las especies analizadas mientras que la variante 1, con la que sólo se diferencia en 45pb en el exón III, y las variantes 2 y 3, las cuales presentan un exón IA alternativo, sólo aparecen en los primates más cercanos en la evolución a *Homo Sapiens*.
3. La isoforma canónica *NFκBIL1* es la de mayor expresión a nivel de mensajero en todos los tejidos y líneas celulares analizadas. La variante 1 es la siguiente en expresión mientras que las variantes 2 y 3 son las de menor expresión, llegando la variante 3 en algunos tejidos a no expresarse.
4. El segundo ATG presente en las proteínas *NFκBIL1* y *NFκBIL1_v1*, es el más adecuado para el inicio de la traducción.
5. Las proteínas codificadas por todas las isoformas de *NFκBIL1* presentan un patrón de distribución nuclear, concretamente en unas estructuras conocidas como espículas nucleares, donde se encuentran proteínas implicadas en el proceso de splicing.
6. Las isoformas de *NFκBIL1* están sobreexpresadas en sangre de pacientes con AR en comparación con la sangre de individuos sanos. En el caso de enfermos con Lupus Eritematoso Sistémico, no se observa esta sobreexpresión.
7. El polimorfismo -62A/T encontrado en la región promotora del gen *NFκBIL1* no está asociado a un desarrollo de AR como se ha visto en otros estudios de este polimorfismo en familias españolas.
8. No existen diferencias significativas en cuanto al patrón de metilación en la región promotora del gen *NFκBIL1* en pacientes con AR e individuos sanos.

CONCLUSIONS

1. The *NFκBIL1* gene undergoes alternative splicing and expresses four mRNA variants that differ on the 5'UTR or in the use of an alternative 5'splice site in exón III. These variants are *NFκBIL1*, *NFκBIL1_v1*, *NFκBIL1_v2* y *NFκBIL1_v3*.
2. *NFκBIL1* isoform is conserved among the species analysed in this study, whilst the rest of the variants of this gene only appear in highly evolved primates.
3. At mRNA level, *NFκBIL1* canonical isoform presents the highest level of expression in all the cell lines and tissues analysed. *NFκBIL1_v2* y *NFκBIL1_v3* variants are expressed in a very low manner, being variant 3 not expressed in some cases.
4. Translation can initiate from either of the first two ATGs in *NFκBIL1* and *NFκBIL1_v1* proteins, being the second ATG preferentially used for it.
5. All of the *NFκBIL1* proteins analysed were localised to nuclear speckles, which are subnuclear structures that are enriched in splicing factors.
6. There is an overexpression of the four *NFκBIL1* mRNA isoforms in patients with RA in comparison to healthy controls.
7. There is no association between -62 A/T polymorphism and RA as it has been demonstrated in other studies in Spanish populations.
8. Methylation analyses for the CpG sites in *NFκBIL1* promoter region do not show a clear correlation with the expression status, suggesting that mechanisms other than CpG methylation are involved in the regulation of *NFκBIL1* expression in RA patients.

BIBLIOGRAFÍA

- Aguado, B., and R.D. Campbell. 1995. The novel gene G17, located in the human major histocompatibility complex, encodes PBX2, a homeodomain-containing protein. *Genomics*. 25:650-9.
- Aguado, B., and R.D. Campbell. 1998. Characterization of a human lysophosphatidic acid acyltransferase that is encoded by a gene located in the class III region of the human major histocompatibility complex. *J Biol Chem*. 273:4096-105.
- Aguado, B., and R.D. Campbell. 1999. Characterization of a human MHC class III region gene product with S-thioesterase activity. *Biochem J*. 341 (Pt 3):679-89.
- Albertella, M.R., and R.D. Campbell. 1994. Characterization of a novel gene in the human major histocompatibility complex that encodes a potential new member of the I kappa B family of proteins. *Hum Mol Genet*. 3:793-9.
- Antequera, F. 2003. Structure, function and evolution of CpG island promoters. *Cell Mol Life Sci*. 60:1647-58.
- Bali, D., S. Gourley, D.D. Kostyu, N. Goel, I. Bruce, A. Bell, D.J. Walker, K. Tran, D.K. Zhu, T.J. Costello, C.I. Amos, and M.F. Seldin. 1999. Genetic analysis of multiplex rheumatoid arthritis families. *Genes Immun*. 1:28-36.
- Barak, M., E.Y. Levanon, E. Eisenberg, N. Paz, G. Rechavi, G.M. Church, and R. Mehr. 2009. Evidence for large diversity in the human transcriptome created by Alu RNA editing. *Nucleic Acids Res*.
- Bass, B.L. 2002. RNA editing by adenosine deaminases that act on RNA. *Annu Rev Biochem*. 71:817-46.
- Beck, S., A. Olek, and J. Walter. 1999. From genomics to epigenomics: a loftier view of life. *Nat Biotechnol*. 17:1144.
- Belicha-Villanueva, A., J. Blickwedehl, S. McEvoy, M. Golding, S.O. Gollnick, and N. Bangia. 2009. What is the role of alternate splicing in antigen presentation by major histocompatibility complex class I molecules? *Immunol Res*.
- Benacerraf, B. 1981. Role of MHC gene products in immune regulation. *Science*. 212:1229-38.
- Benight, A.S., R.M. Wartell, and D.K. Howell. 1981. Theory agrees with experimental thermal denaturation of short DNA restriction fragments. *Nature*. 289:203-5.
- Berget, S.M., C. Moore, and P.A. Sharp. 1977. Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proc Natl Acad Sci U S A*. 74:3171-5.
- Bird, A. 2002. DNA methylation patterns and epigenetic memory. *Genes Dev*. 16:6-21.
- Blencowe, B.J. 2006. Alternative splicing: new insights from global analyses. *Cell*. 126:37-47.
- Bodmer, W.F. 1972. Evolutionary significance of the HL-A system. *Nature*. 237:139-45 passim.
- Butcher, L.M., and S. Beck. 2008. Future impact of integrated high-throughput methylome analyses on human health and disease. *J Genet Genomics*. 35:391-401.
- Calvanese, V., M. Mallya, R.D. Campbell, and B. Aguado. 2008. Regulation of expression of two LY-6 family genes by intron retention and transcription induced chimerism. *BMC Mol Biol*. 9:81.
- Carrington, M., G.W. Nelson, M.P. Martin, T. Kissner, D. Vlahov, J.J. Goedert, R. Kaslow, S. Buchbinder, K. Hoots, and S.J. O'Brien. 1999. HLA and HIV-1: heterozygote advantage and B*35-Cw*04 disadvantage. *Science*. 283:1748-52.
- Caceres, J.F., T. Misteli, G.R. Screaton, D.L. Spector, and A.R. Krainer. 1997. Role of the modular domains of SR proteins in subnuclear localization and alternative splicing specificity. *J Cell Biol*. 138:225-38.
- Castiblanco, J., and J.M. Anaya. 2008. The IkappaBL gene polymorphism influences risk of acquiring systemic lupus erythematosus and Sjogren's syndrome. *Hum Immunol*. 69:45-51.

- Collado, L., B. Rueda, R. Caliz, B. Torres, A. Garcia, A. Nunez-Roldan, M.F. Gonzalez-Escribano, and J. Martin. 2004. Lack of association between the I kappa BL promoter polymorphism and rheumatoid arthritis. *Arthritis Rheum.* 50:2032-3.
- Cucca, F., Z.B. Zhu, A. Khanna, F. Cossu, M. Congia, M. Badiali, R. Lampis, F. Frau, S. De Virgiliis, A. Cao, M. Arnone, P. Piras, R.D. Campbell, M.D. Cooper, J.E. Volanakis, and S.H. Powis. 1998. Evaluation of IgA deficiency in Sardinians indicates a susceptibility gene is encoded within the HLA class III region. *Clin Exp Immunol.* 111:76-80.
- Cudworth, A.G., and J.C. Woodrow. 1975. Evidence for HL-A-linked genes in "juvenile" diabetes mellitus. *Br Med J.* 3:133-5.
- Chacko, E., and S. Ranganathan. 2009. Comprehensive splicing graph analysis of alternative splicing patterns in chicken, compared to human and mouse. *BMC Genomics.* 10 Suppl 1:S5.
- Chow, L.T., R.E. Gelinas, T.R. Broker, and R.J. Roberts. 1977. An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell.* 12:1-8.
- Dausset, J. 1981. The major histocompatibility complex in man. *Science.* 213:1469-74.
- Davuluri, R.V., Y. Suzuki, S. Sugano, C. Plass, and T.H. Huang. 2008. The functional consequences of alternative promoter use in mammalian genomes. *Trends Genet.* 24:167-77
- de Vet, E.C., B. Aguado, and R.D. Campbell. 2001. G6b, a novel immunoglobulin superfamily member encoded in the human major histocompatibility complex, interacts with SHP-1 and SHP-2. *J Biol Chem.* 276:42070-6.
- Del Gatto-Konczak, F., C.F. Bourgeois, C. Le Guiner, L. Kister, M.C. Gesnel, J. Stevenin, and R. Breathnach. 2000. The RNA-binding protein TIA-1 is a novel mammalian splicing regulator acting through intron sequences adjacent to a 5' splice site. *Mol Cell Biol.* 20:6287-99.
- Del Gatto-Konczak, F., M. Olive, M.C. Gesnel, and R. Breathnach. 1999. hnRNP A1 recruited to an exon in vivo can function as an exon splicing silencer. *Mol Cell Biol.* 19:251-60.
- Ditzel, H.J., Y. Masaki, H. Nielsen, L. Farnaes, and D.R. Burton. 2000. Cloning and expression of a novel human antibody-antigen pair associated with Felty's syndrome. *Proc Natl Acad Sci U S A.* 97:9234-9.
- Dobrzanski, P., R.P. Ryseck, and R. Bravo. 1995. Specific inhibition of RelB/p52 transcriptional activity by the C-terminal domain of p100. *Oncogene.* 10:1003-7.
- Dye, B.T., and J.G. Patton. 2001. An RNA recognition motif (RRM) is required for the localization of PTB-associated splicing factor (PSF) to subnuclear speckles. *Exp Cell Res.* 263:131-44.
- Egger, G., G. Liang, A. Aparicio, and P.A. Jones. 2004. Epigenetics in human disease and prospects for epigenetic therapy. *Nature.* 429:457-63.
- Faustino, N.A., and T.A. Cooper. 2003. Pre-mRNA splicing and human disease. *Genes Dev.* 17:419-37.
- Feder, J.N., A. Gnirke, W. Thomas, Z. Tsuchihashi, D.A. Ruddy, A. Basava, F. Dormishian, R. Domingo, Jr., M.C. Ellis, A. Fullan, L.M. Hinton, N.L. Jones, B.E. Kimmel, G.S. Kronmal, P. Lauer, V.K. Lee, D.B. Loeb, F.A. Mapa, E. McClelland, N.C. Meyer, G.A. Mintier, N. Moeller, T. Moore, E. Morikang, C.E. Prass, L. Quintana, S.M. Starnes, R.C. Schatzman, K.J. Brunke, D.T. Drayna, N.J. Risch, B.R. Bacon, and R.K. Wolff. 1996. A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis. *Nat Genet.* 13:399-408.
- Ferlito, M., O.G. Romanenko, S. Ashton, F. Squadrito, P.V. Halushka, and J.A. Cook. 2001. Effect of cross-tolerance between endotoxin and TNF-alpha or IL-1beta on cellular signaling and mediator production. *J Leukoc Biol.* 70:821-9.

- Filipowicz, W., S.N. Bhattacharyya, and N. Sonenberg. 2008. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat Rev Genet.* 9:102-14.
- Fogdell, A., J. Hillert, C. Sachs, and O. Olerup. 1995. The multiple sclerosis- and narcolepsy-associated HLA class II haplotype includes the DRB5*0101 allele. *Tissue Antigens.* 46:333-6.
- Garcia-Blanco, M.A., A.P. Baraniak, and E.L. Lasda. 2004. Alternative splicing in disease and therapy. *Nat Biotechnol.* 22:535-46.
- Ghosh, S., and M.S. Hayden. 2008. New regulators of NF-kappaB in inflammation. *Nat Rev Immunol.* 8:837-48.
- Ghosh, S., and M. Karin. 2002. Missing pieces in the NF-kappaB puzzle. *Cell.* 109 Suppl:S81-96.
- Ghosh, S., M.J. May, and E.B. Kopp. 1998. NF-kappa B and Rel proteins: evolutionarily conserved mediators of immune responses. *Annu Rev Immunol.* 16:225-60.
- Greetham, D., C.D. Ellis, D. Mewar, U. Fearon, S.N. an Ultaigh, D.J. Veale, F. Guesdon, and A.G. Wilson. 2007. Functional characterization of NF-kappaB inhibitor-like protein 1 (NFKappaBIL1), a candidate susceptibility gene for rheumatoid arthritis. *Hum Mol Genet.* 16:3027-36.
- Gruss, H.J., and S.K. Dower. 1995. Tumor necrosis factor ligand superfamily: involvement in the pathology of malignant lymphomas. *Blood.* 85:3378-404.
- Gustincich, S., A. Sandelin, C. Plessy, S. Katayama, R. Simone, D. Lazarevic, Y. Hayashizaki, and P. Carninci. 2006. The complexity of the mammalian transcriptome. *J Physiol.* 575:321-32.
- Hacker, H., and M. Karin. 2002. Is NF-kappaB2/p100 a direct activator of programmed cell death? *Cancer Cell.* 2:431-3.
- Hagiwara, M. 2005. Alternative splicing: a new drug target of the post-genome era. *Biochim Biophys Acta.* 1754:324-31.
- Hatada, E.N., M. Naumann, and C. Scheidereit. 1993. Common structural constituents confer I kappa B activity to NF-kappa B p105 and I kappa B/MAD-3. *Embo J.* 12:2781-8.
- Hayden, M.S., and S. Ghosh. 2008. Shared principles in NF-kappaB signaling. *Cell.* 132:344-62.
- Heider, K.H., H. Kuthan, G. Stehle, and G. Munzert. 2004. CD44v6: a target for antibody-based cancer therapy. *Cancer Immunol Immunother.* 53:567-79.
- Heissmeyer, V., D. Krappmann, F.G. Wulczyn, and C. Scheidereit. 1999. NF-kappaB p105 is a target of IkappaB kinases and controls signal induction of Bcl-3-p50 complexes. *Embo J.* 18:4766-78.
- Higuchi, R., C. Fockler, G. Dollinger, and R. Watson. 1993. Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. *Biotechnology (N Y).* 11:1026-30.
- Hill, A.V. 2006. Aspects of genetic susceptibility to human infectious diseases. *Annu Rev Genet.* 40:469-86.
- Horton, R., L. Wilming, V. Rand, R.C. Lovering, E.A. Bruford, V.K. Khodiyar, M.J. Lush, S. Povey, C.C. Talbot, Jr., M.W. Wright, H.M. Wain, J. Trowsdale, A. Ziegler, and S. Beck. 2004. Gene map of the extended human MHC. *Nat Rev Genet.* 5:889-99.
- Houseley, J., and D. Tollervey. 2009. The many pathways of RNA degradation. *Cell.* 136:763-76.
- Howell, W.M., P.C. Calder, and R.F. Grimble. 2002. Gene polymorphisms, inflammatory diseases and cancer. *Proc Nutr Soc.* 61:447-56.
- Hsieh, S.L., and R.D. Campbell. 1991. Evidence that gene G7a in the human major histocompatibility complex encodes valyl-tRNA synthetase. *Biochem J.* 278 (Pt 3):809-16.

- Huang, C.H., L. Cong, J. Xie, B. Qiao, S.H. Lo, and T. Zheng. 2009. Rheumatoid arthritis-associated gene-gene interaction network for rheumatoid arthritis candidate genes. *BMC Proc.* 3 Suppl 7:S75.
- Hui, J. 2009. Regulation of mammalian pre-mRNA splicing. *Sci China C Life Sci.* 52:253-60.
- Jacob, S., M.K. McClintock, B. Zelano, and C. Ober. 2002. Paternally inherited HLA alleles are associated with women's choice of male odor. *Nat Genet.* 30:175-9.
- Hurt, P., L. Walter, R. Sudbrak, S. Klages, I. Muller, T. Shiina, H. Inoko, H. Lehrach, E. Gunther, R. Reinhardt, and H. Himmelbauer. 2004. The genomic sequence and comparative analysis of the rat major histocompatibility complex. *Genome Res.* 14:631-9.
- Iida, H., A. Urasoko, M. Doiguchi, T. Mori, K. Toshimori, and Y. Shibata. 2003. Complementary DNA cloning and characterization of rat spergen-2, a spermatogenic cell-specific gene 2 encoding a 56-kilodalton nuclear protein bearing ankyrin repeat motifs. *Biol Reprod.* 69:421-9.
- Jaenisch, R., and A. Bird. 2003. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet.* 33 Suppl:245-54.
- Jenkins, S.C., R.E. March, R.D. Campbell, and C.M. Milner. 2000. A novel variant of the MHC-linked hsp70, hsp70-hom, is associated with rheumatoid arthritis. *Tissue Antigens.* 56:38-44.
- Jones, P.A., and S.B. Baylin. 2002. The fundamental role of epigenetic events in cancer. *Nat Rev Genet.* 3:415-28.
- Karin, M., and Y. Ben-Neriah. 2000. Phosphorylation meets ubiquitination: the control of NF- κ B activity. *Annu Rev Immunol.* 18:621-63.
- Kidd, J.M., G.M. Cooper, W.F. Donahue, H.S. Hayden, N. Sampas, T. Graves, N. Hansen, B. Teague, C. Alkan, F. Antonacci, E. Haugen, T. Zerr, N.A. Yamada, P. Tsang, T.L. Newman, E. Tuzun, Z. Cheng, H.M. Ebling, N. Tusneem, R. David, W. Gillett, K.A. Phelps, M. Weaver, D. Saranga, A. Brand, W. Tao, E. Gustafson, K. McKernan, L. Chen, M. Malig, J.D. Smith, J.M. Korn, S.A. McCarroll, D.A. Altshuler, D.A. Peiffer, M. Dorschner, J. Stamatoyannopoulos, D. Schwartz, D.A. Nickerson, J.C. Mullikin, R.K. Wilson, L. Bruhn, M.V. Olson, R. Kaul, D.R. Smith, and E.E. Eichler. 2008. Mapping and sequencing of structural variation from eight human genomes. *Nature.* 453:56-64.
- Kilding, R., and A.G. Wilson. 2005. Mapping of a novel susceptibility gene for rheumatoid arthritis in the telomeric MHC region. *Cytokine.* 32:71-5.
- Kim, E., A. Magen, and G. Ast. 2007. Different levels of alternative splicing among eukaryotes. *Nucleic Acids Res.* 35:125-31.
- Kreth, S., C. Ledderose, I. Kaufmann, G. Groeger, and M. Thiel. 2008. Differential expression of 5'-UTR splice variants of the adenosine A2A receptor gene in human granulocytes: identification, characterization, and functional impact on activation. *Faseb J.* 22:3276-86.
- Kwan, T., D. Benovoy, C. Dias, S. Gurd, C. Provencher, P. Beaulieu, T.J. Hudson, R. Sladek, and J. Majewski. 2008. Genome-wide analysis of transcript isoform variation in humans. *Nat Genet.* 40:225-31.
- Kwan, T., D. Benovoy, C. Dias, S. Gurd, D. Serre, H. Zuzan, T.A. Clark, A. Schweitzer, M.K. Staples, H. Wang, J.E. Blume, T.J. Hudson, R. Sladek, and J. Majewski. 2007. Heritability of alternative splicing in the human genome. *Genome Res.* 17:1210-8.
- Lander, E.S., L.M. Linton, B. Birren, C. Nusbaum, M.C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, R. Funke, D. Gage, K. Harris, A. Heaford, J. Howland, L. Kann, J. Lehoczky, R. LeVine, P. McEwan, K. McKernan, J. Meldrim, J.P. Mesirov, C. Miranda, W. Morris, J. Naylor, C. Raymond, M. Rosetti, R. Santos, A. Sheridan, C. Sougnez, N. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson,

- R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J.C. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Shownkeen, S. Sims, R.H. Waterston, R.K. Wilson, L.W. Hillier, J.D. McPherson, M.A. Marra, E.R. Mardis, L.A. Fulton, A.T. Chinwalla, K.H. Pepin, W.R. Gish, S.L. Chissoe, M.C. Wendl, K.D. Delehaunty, T.L. Miner, A. Delehaunty, J.B. Kramer, L.L. Cook, R.S. Fulton, D.L. Johnson, P.J. Minx, S.W. Clifton, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T. Slezak, N. Doggett, J.F. Cheng, A. Olsen, S. Lucas, C. Elkin, E. Uberbacher, M. Frazier, et al. 2001. Initial sequencing and analysis of the human genome. *Nature*. 409:860-921.
- Lee, G.R., S.T. Kim, C.G. Spilianakis, P.E. Fields, and R.A. Flavell. 2006. T helper cell differentiation: regulation by cis elements and epigenetics. *Immunity*. 24:369-79.
- Li, W., B. Xi, W. Yang, M. Hawkins, and U.K. Schubart. 2003. Complex DNA melting profiles of small PCR products revealed using SYBR Green I. *Biotechniques*. 35:702-4, 706.
- Licatalosi, D.D., and R.B. Darnell. 2010. RNA processing and its regulation: global insights into biological networks. *Nat Rev Genet*. 11:75-87.
- Lin, R., P. Beauparlant, C. Makris, S. Meloche, and J. Hiscott. 1996. Phosphorylation of IkappaBalpha in the C-terminal PEST domain by casein kinase II affects intrinsic protein stability. *Mol Cell Biol*. 16:1401-9.
- Lindberg, J., and J. Lundeberg. 2009. The plasticity of the mammalian transcriptome. *Genomics*.
- Lunn, M.R., and C.H. Wang. 2008. Spinal muscular atrophy. *Lancet*. 371:2120-33.
- Luo, M.L., Z. Zhou, K. Magni, C. Christoforides, J. Rappsilber, M. Mann, and R. Reed. 2001. Pre-mRNA splicing and mRNA export linked by direct interactions between UAP56 and Aly. *Nature*. 413:644-7.
- Malatesta, M., S. Fakan, and U. Fischer. 1999. The Sm core domain mediates targeting of U1 snRNP to subnuclear compartments involved in transcription and splicing. *Exp Cell Res*. 249:189-98.
- Mallya, M., R.D. Campbell, and B. Aguado. 2002. Transcriptional analysis of a novel cluster of LY-6 family members in the human and mouse major histocompatibility complex: five genes with many splice forms. *Genomics*. 80:113-23.
- Martinez, A., M. Pascual, D. Pascual-Salcedo, A. Balsa, J. Martin, and E.G. de la Concha. 2003. Genetic polymorphisms in Spanish rheumatoid arthritis patients: an association and linkage study. *Genes Immun*. 4:117-21.
- Mendes Soares, L.M., and J. Valcarcel. 2006. The expanding transcriptome: the genome as the 'Book of Sand'. *Embo J*. 25:923-31.
- Milner, C.M., S.V. Smith, M.B. Carrillo, G.L. Taylor, M. Hollinshead, and R.D. Campbell. 1997. Identification of a sialidase encoded in the human major histocompatibility complex. *J Biol Chem*. 272:4549-58.
- Miller, W.L., and Y. Morel. 1989. The molecular genetics of 21-hydroxylase deficiency. *Annu Rev Genet*. 23:371-93.
- Modrek, B., and C. Lee. 2002. A genomic view of alternative splicing. *Nat Genet*. 30:13-9.
- Mulcahy, H., K.P. O'Rourke, C. Adams, M.G. Molloy, and F. O'Gara. 2006. LST1 and NCR3 expression in autoimmune inflammation and in response to IFN-gamma, LPS and microbial infection. *Immunogenetics*. 57:893-903.
- Muller, C.W., F.A. Rey, M. Sodeoka, G.L. Verdine, and S.C. Harrison. 1995. Structure of the NF-kappa B p50 homodimer bound to DNA. *Nature*. 373:311-7.
- Mungall, A.J., S.A. Palmer, S.K. Sims, C.A. Edwards, J.L. Ashurst, L. Wilming, M.C. Jones, R. Horton, S.E. Hunt, C.E. Scott, J.G. Gilbert, M.E. Clamp, G. Bethel, S. Milne, R. Ainscough, J.P. Almeida, K.D. Ambrose, T.D. Andrews, R.I. Ashwell, A.K. Babbage, C.L. Bagguley, J. Bailey, R. Banerjee, D.J. Barker, K.F. Barlow, K. Bates, D.M. Beare, H. Beasley, O. Beasley, C.P. Bird, S. Blakey, S. Bray-Allen, J. Brook, A.J. Brown, J.Y.

- Brown, D.C. Burford, W. Burrill, J. Burton, C. Carder, N.P. Carter, J.C. Chapman, S.Y. Clark, G. Clark, C.M. Clee, S. Clegg, V. Cobley, R.E. Collier, J.E. Collins, L.K. Colman, N.R. Corby, G.J. Coville, K.M. Culley, P. Dharni, J. Davies, M. Dunn, M.E. Earthrowl, A.E. Ellington, K.A. Evans, L. Faulkner, M.D. Francis, A. Frankish, J. Frankland, L. French, P. Garner, J. Garnett, M.J. Ghorri, L.M. Gilby, C.J. Gillson, R.J. Glithero, D.V. Grafham, M. Grant, S. Gribble, C. Griffiths, M. Griffiths, R. Hall, K.S. Halls, S. Hammond, J.L. Harley, E.A. Hart, P.D. Heath, R. Heathcote, S.J. Holmes, P.J. Howden, K.L. Howe, G.R. Howell, E. Huckle, S.J. Humphray, M.D. Humphries, A.R. Hunt, C.M. Johnson, A.A. Joy, M. Kay, S.J. Keenan, A.M. Kimberley, A. King, G.K. Laird, C. Langford, S. Lawlor, D.A. Leongamornlert, M. Leversha, et al. 2003. The DNA sequence and analysis of human chromosome 6. *Nature*. 425:805-11.
- Neeper, M., A.M. Schmidt, J. Brett, S.D. Yan, F. Wang, Y.C. Pan, K. Elliston, D. Stern, and A. Shaw. 1992. Cloning and expression of a cell surface receptor for advanced glycosylation end products of proteins. *J Biol Chem*. 267:14998-5004.
- Neville, M.J., and R.D. Campbell. 1999. A new member of the Ig superfamily and a V-ATPase G subunit are among the predicted products of novel genes close to the TNF locus in the human MHC. *J Immunol*. 162:4745-54.
- Nile, C.J., R.C. Read, M. Akil, G.W. Duff, and A.G. Wilson. 2008. Methylation status of a single CpG site in the IL6 promoter is related to IL6 messenger RNA levels and rheumatoid arthritis. *Arthritis Rheum*. 58:2686-93.
- Okamoto, K., S. Makino, Y. Yoshikawa, A. Takaki, Y. Nagatsuka, M. Ota, G. Tamiya, A. Kimura, S. Bahram, and H. Inoko. 2003. Identification of I kappa BL as the second major histocompatibility complex-linked susceptibility locus for rheumatoid arthritis. *Am J Hum Genet*. 72:303-12.
- Pan, Q., O. Shai, L.J. Lee, B.J. Frey, and B.J. Blencowe. 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet*. 40:1413-5.
- Pasparakis, M. 2009. Regulation of tissue homeostasis by NF-kappaB signalling: implications for inflammatory diseases. *Nat Rev Immunol*. 9:778-88.
- Petronis, A. 2001. Human morbid genetics revisited: relevance of epigenetics. *Trends Genet*. 17:142-6.
- Rakyan, V.K., T. Hildmann, K.L. Novik, J. Lewin, J. Tost, A.V. Cox, T.D. Andrews, K.L. Howe, T. Otto, A. Olek, J. Fischer, I.G. Gut, K. Berlin, and S. Beck. 2004. DNA methylation profiling of the human major histocompatibility complex: a pilot study for the human epigenome project. *PLoS Biol*. 2:e405.
- Ramasawmy, R., K.C. Fae, E. Cunha-Neto, S.C. Borba, B. Ianni, C. Mady, A.C. Goldberg, and J. Kalil. 2008. Variants in the promoter region of IKBL/NFKBIL1 gene may mark susceptibility to the development of chronic Chagas' cardiomyopathy among Trypanosoma cruzi-infected individuals. *Mol Immunol*. 45:283-8.
- Renard, C., E. Hart, H. Sehra, H. Beasley, P. Coggill, K. Howe, J. Harrow, J. Gilbert, S. Sims, J. Rogers, A. Ando, A. Shigenari, T. Shiina, H. Inoko, P. Chardon, and S. Beck. 2006. The genomic sequence and analysis of the swine major histocompatibility complex. *Genomics*. 88:96-110.
- Rodriguez, A.J., K. Czapinski, J.S. Condeelis, and R.H. Singer. 2008. Mechanisms and cellular roles of local protein synthesis in mammalian cells. *Curr Opin Cell Biol*. 20:144-9.
- Rogers, S., R. Wells, and M. Rechsteiner. 1986. Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science*. 234:364-8.
- Rottman, F.M., J.A. Bokar, P. Narayan, M.E. Shambaugh, and R. Ludwiczak. 1994. N6-adenosine methylation in mRNA: substrate specificity and enzyme complexity. *Biochimie*. 76:1109-14.

- Scherer, S.W., C. Lee, E. Birney, D.M. Altshuler, E.E. Eichler, N.P. Carter, M.E. Hurles, and L. Feuk. 2007. Challenges and standards in integrating surveys of structural variation. *Nat Genet.* 39:S7-15.
- Schroeder, H.W., Jr., Z.B. Zhu, R.E. March, R.D. Campbell, S.M. Berney, S.A. Nedospasov, R.L. Turetskaya, T.P. Atkinson, R.C. Go, M.D. Cooper, and J.E. Volanakis. 1998. Susceptibility locus for IgA deficiency and common variable immunodeficiency in the HLA-DR3, -B8, -A1 haplotypes. *Mol Med.* 4:72-86.
- Semple, J.I., S.E. Brown, C.M. Sanderson, and R.D. Campbell. 2002. A distinct bipartite motif is required for the localization of inhibitory kappaB-like (IkappaBL) protein to nuclear speckles. *Biochem J.* 361:489-96.
- Sen, R., and D. Baltimore. 1986. Multiple nuclear factors interact with the immunoglobulin enhancer sequences. *Cell.* 46:705-16.
- Seshadri, S., Y. Kannan, S. Mitra, J. Parker-Barnes, and M.D. Wewers. 2009. MAIL regulates human monocyte IL-6 production. *J Immunol.* 183:5358-68.
- Sharp, P.A. 1994. Split genes and RNA splicing. *Cell.* 77:805-15.
- Smith, C.W., and J. Valcarcel. 2000. Alternative pre-mRNA splicing: the logic of combinatorial control. *Trends Biochem Sci.* 25:381-8.
- Snell, G.D. 1948. Methods for the study of histocompatibility genes. *J Genet.* 49:87-108.
- Snoek, M., C. Teuscher, and H. van Vugt. 1998. Molecular analysis of the major MHC recombinational hot spot located within the G7c gene of the murine class III region that is involved in disease susceptibility. *J Immunol.* 160:266-72.
- Staley, J.P., and C. Guthrie. 1998. Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell.* 92:315-26.
- Stephens, R., R. Horton, S. Humphray, L. Rowen, J. Trowsdale, and S. Beck. 1999. Gene organisation, sequence variation and isochore structure at the centromeric boundary of the human MHC. *J Mol Biol.* 291:789-99.
- Sterenczak, K.A., S. Willenbrock, M. Barann, M. Klemke, J.T. Soller, N. Eberle, I. Nolte, J. Bullerdiek, and H. Murua Escobar. 2009. Cloning, characterisation, and comparative quantitative expression analyses of receptor for advanced glycation end products (RAGE) transcript forms. *Gene.* 434:35-42.
- Su, I.H., and A. Tarakhovsky. 2005. Epigenetic control of B cell differentiation. *Semin Immunol.* 17:167-72.
- Tam, W.F., L.H. Lee, L. Davis, and R. Sen. 2000. Cytoplasmic sequestration of rel proteins by IkappaBalpha requires CRM1-dependent nuclear export. *Mol Cell Biol.* 20:2269-84.
- Terns, M., and R. Terns. 2006. Noncoding RNAs of the H/ACA family. *Cold Spring Harb Symp Quant Biol.* 71:395-405.
- The MHC sequencing consortium 1999. Complete sequence and gene map of a human major histocompatibility complex. *Nature.* 401:921-3.
- Todd, J.A., J.I. Bell, and H.O. McDevitt. 1987. HLA-DQ beta gene contributes to susceptibility and resistance to insulin-dependent diabetes mellitus. *Nature.* 329:599-604.
- Traherne, J.A. 2008. Human MHC architecture and evolution: implications for disease association studies. *Int J Immunogenet.* 35:179-92.
- Trenkmann, M., M. Brock, C. Ospelt, and S. Gay. 2009. Epigenetics in Rheumatoid Arthritis. *Clin Rev Allergy Immunol.*
- Trowsdale, J. 1993. Genomic structure and function in the MHC. *Trends Genet.* 9:117-22.
- Trowsdale, J., and P. Parham. 2004. Mini-review: defense strategies and immunity-related genes. *Eur J Immunol.* 34:7-17.
- Utans, U., R.J. Arceci, Y. Yamashita, and M.E. Russell. 1995. Cloning and characterization of allograft inflammatory factor-1: a novel macrophage factor identified in rat cardiac allografts with chronic rejection. *J Clin Invest.* 95:2954-62.
- Vallabhapurapu, S., and M. Karin. 2009. Regulation and function of NF-kappaB transcription factors in the immune system. *Annu Rev Immunol.* 27:693-733.

- Vandiedonck, C., and J.C. Knight. 2009. The human Major Histocompatibility Complex as a paradigm in genomics research. *Brief Funct Genomic Proteomic*.
- Venter, J.C., M.D. Adams, E.W. Myers, P.W. Li, R.J. Mural, G.G. Sutton, H.O. Smith, M. Yandell, C.A. Evans, R.A. Holt, J.D. Gocayne, P. Amanatides, R.M. Ballew, D.H. Huson, J.R. Wortman, Q. Zhang, C.D. Kodira, X.H. Zheng, L. Chen, M. Skupski, G. Subramanian, P.D. Thomas, J. Zhang, G.L. Gabor Miklos, C. Nelson, S. Broder, A.G. Clark, J. Nadeau, V.A. McKusick, N. Zinder, A.J. Levine, R.J. Roberts, M. Simon, C. Slayman, M. Hunkapiller, R. Bolanos, A. Delcher, I. Dew, D. Fasulo, M. Flanigan, L. Florea, A. Halpern, S. Hannenhalli, S. Kravitz, S. Levy, C. Mobarry, K. Reinert, K. Remington, J. Abu-Threideh, E. Beasley, K. Biddick, V. Bonazzi, R. Brandon, M. Cargill, I. Chandramouliswaran, R. Charlab, K. Chaturvedi, Z. Deng, V. Di Francesco, P. Dunn, K. Eilbeck, C. Evangelista, A.E. Gabrielian, W. Gan, W. Ge, F. Gong, Z. Gu, P. Guan, T.J. Heiman, M.E. Higgins, R.R. Ji, Z. Ke, K.A. Ketchum, Z. Lai, Y. Lei, Z. Li, J. Li, Y. Liang, X. Lin, F. Lu, G.V. Merkulov, N. Milshina, H.M. Moore, A.K. Naik, V.A. Narayan, B. Neelam, D. Nusskern, D.B. Rusch, S. Salzberg, W. Shao, B. Shue, J. Sun, Z. Wang, A. Wang, X. Wang, J. Wang, M. Wei, R. Wides, C. Xiao, C. Yan, et al. 2001. The sequence of the human genome. *Science*. 291:1304-51.
- Villate, O., A. Rastrojo, R. Lopez-Diez, F. Hernandez-Torres, and B. Aguado. 2008. Differential splicing, disease and drug targets. *Infect Disord Drug Targets*. 8:241-51.
- Wang, E.T., R. Sandberg, S. Luo, I. Khrebtkova, L. Zhang, C. Mayr, S.F. Kingsmore, G.P. Schroth, and C.B. Burge. 2008. Alternative isoform regulation in human tissue transcriptomes. *Nature*. 456:470-6.
- Wang, G.S., and T.A. Cooper. 2007. Splicing in disease: disruption of the splicing code and the decoding machinery. *Nat Rev Genet*. 8:749-61.
- Weber, M., J.J. Davies, D. Wittig, E.J. Oakeley, M. Haase, W.L. Lam, and D. Schubeler. 2005. Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat Genet*. 37:853-62.
- Whiteside, S.T., J.C. Epinat, N.R. Rice, and A. Israel. 1997. I kappa B epsilon, a novel member of the I kappa B family, controls RelA and cRel NF-kappa B activity. *Embo J*. 16:1413-26.
- Wong, M.L., and J.F. Medrano. 2005. Real-time PCR for mRNA quantitation. *Biotechniques*. 39:75-85.
- Wood, M., H. Yin, and G. McClorey. 2007. Modulating the expression of disease genes with RNA-based therapy. *PLoS Genet*. 3:e109.
- Xie, T., L. Rowen, B. Aguado, M.E. Ahearn, A. Madan, S. Qin, R.D. Campbell, and L. Hood. 2003. Analysis of the gene-dense major histocompatibility complex class III region and its comparison to mouse. *Genome Res*. 13:2621-36.
- Yamashita, T., K. Hamaguchi, Y. Kusuda, A. Kimura, T. Sakata, and H. Yoshimatsu. 2004. IKBL promoter polymorphism is strongly associated with resistance to type 1 diabetes in Japanese. *Tissue Antigens*. 63:223-30.
- Yilmaz, Z.B., D.S. Weih, V. Sivakumar, and F. Weih. 2003. RelB is required for Peyer's patch development: differential regulation of p52-RelB by lymphotoxin and TNF. *EMBO J*. 22:121-30.