UNIVERSIDAD DE SALAMANCA

DEPARTAMENTO DE MEDICINA
HEMATOLOGÍA

VNiVERSiDAD
D SALAMANCA

CAMPUS DE EXCELENCIA INTERNACIONAL

TESIS DOCTORAL

Polimorfismos genéticos, alteraciones moleculares y reordenamientos clonotípicos en el linfoma B difuso de célula grande.

Correlaciones clínico-biológicas

Elena Sebastián Pérez

Salamanca, 2016

UNIVERSIDAD DE SALAMANCA

DEPARTAMENTO DE MEDICINA
HEMATOLOGÍA

VNiVERSiDAD
D SALAMANCA

CAMPUS DE EXCELENCIA INTERNACIONAL

TESIS DOCTORAL

Polimorfismos genéticos, alteraciones moleculares y reordenamientos clonotípicos en el linfoma B difuso de célula grande.

Correlaciones clínico-biológicas

Directores:

Prof. Marcos González Díaz

Dr. Ramón García Sanz

Dr. Miguel Alcoceba

Elena Sebastián Pérez

D. MARCOS GONZÁLEZ DÍAZ, Doctor en Medicina, Profesor Titular de Universidad, Jefe de Sección del Servicio de Hematología del Hospital Clínico Universitario de Salamanca.

D. RAMÓN GARCÍA SANZ, Doctor en Medicina, Profesor Asociado de Medicina y Adjunto del Servicio de Hematología del Hospital Clínico Universitario de Salamanca.

D. MIGUEL ALCOCEBA, Doctor en Biología, Investigador del Hospital Clínico Universitario de Salamanca.

CERTIFICAN:

Que el presente trabajado titulado «Polimorfismos genéticos, alteraciones moleculares y reordenamientos clonotípicos en el linfoma B difuso de célula grande. Correlaciones clínico-biológicas» ha sido realizado por D.ª Elena Sebastián Pérez bajo su dirección en el Departamento de Medicina, reúne, a su juicio, las condiciones de originalidad y calidad científica requeridas para su presentación y defensa para optar al grado de Doctor en Medicina con mención "Doctor Europeus" por la Universidad de Salamanca.

Y para que así conste a los efectos oportunos, firmamos el presente certificado en Salamanca, a 30 de noviembre de 2015.

Fdo. Prof. D. Marcos González Díaz

Fdo. Dr. D. Ramón García Sanz

Fdo. Dr. D. Miguel Alcoceba

La presente tesis doctoral corresponde a un compendio de 3 trabajos previamente publicados, que se detallan a continuación:

1. **Molecular characterization of immunoglobulin gene rearrangements in diffuse large B-cell lymphoma: antigen-driven origin and IGHV4-34 as a particular subgroup of the non-GCB subtype.**

   Sebastián E, Alcoceba M, Balanzategui A, Marín L, Montes-Moreno S, Flores T, González D, Sarasquete ME, Chillón MC, Puig N, Corral R, Pardal E, Martín A, González-Barca E, Caballero MD, San Miguel JF, García-Sanz R, González M.
   American Journal of Pathology 2012; 181(5): 1879-88; DOI: 10.1016/j.ajpath.2012.07.028.

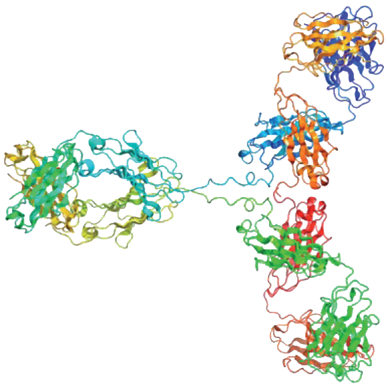
2. **High-resolution copy number analysis of paired normal-tumor samples from diffuse large B cell lymphoma.**

   Sebastián E*, Alcoceba M*, Martín-García D*, Blanco Ó, Sanchez-Barba M, Balanzategui A, Marín L, Montes-Moreno S, González-Barca E, Pardal E, Jiménez C, García-Álvarez M, Clot G, Carracedo Á, Gutiérrez NC, Sarasquete ME, Chillón C, Corral R, Prieto-Conde MI, Caballero MD, Salaverria I*, García-Sanz R*, González M*.
   Ann Hematol. 2015 Nov 14. [Epub ahead of print]

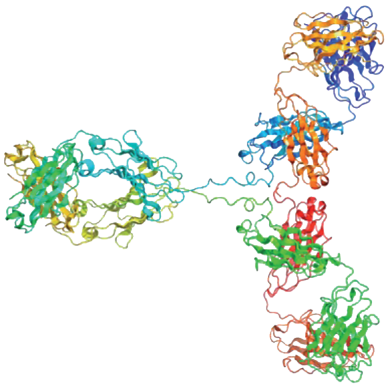
3. **HLA specificities are related to development and prognosis of diffuse large B-cell lymphoma.**

   Alcoceba M*, Sebastián E*, Marín L, Balanzategui A, Sarasquete ME, Chillón MC, Jiménez C, Puig N, Corral R, Pardal E, Grande C, Bello JL, Albo C, de la Cruz F, Panizo C, Martín A, González-Barca E, Caballero MD, San Miguel JF, García-Sanz R, González M.
   Blood 2013; 122(8): 1448-54; DOI: 10.1182/blood-2013-02-483420.
   *equal contribution

Índice



Abreviaturas



aa	Aminoácido
aaIPI	(Age-adjusted IPI) - IPI ajustado a la edad
ABC	(Actived-B cell) - célula B activada
ADN	Ácido desoxirribonucleico
AID	Activation-induced cytidine deaminase
ALK	Anaplastic lymphoma receptor tyrosine kinase
ARF	ADP Ribosylation Factors
ARHGEF1	RHO, guanine nucleotide Exchange factor 1
ARNm	Ácido ribonucleico mensajero
ASHMS	(Aberrants somatic hypermutation) - SHM aberrantes
β2M	Beta-2 microglobulina
BA	Break apart
BCL10	B-cell CLL/lymphoma 10
BCL11A	B-cell CLL/lymphoma 11A
BCL2	B-cell CLL/lymphoma 2
BCL-6	B-cell CLL/lymphoma 6
BCR	(B-cell receptor) - Receptor de célula B
BET	Bromodomain and extra terminal family
BLIMP1	B lymphocite-induced maduration protein
BM	Biología Molecular
BTK	Bruton agammaglobulinemia tyrosine kinase
C	Constante
CAR	Chimeric antigen receptor
CARD11	Caspase recruitment domain-containing protein 11
CCND1	Cyclin D1
CD40 L	CD40 ligand
CDKN1A	Cyclin-dependent kinase inhibitor 1A (p21, Cip1)
CDKN2A	Cyclin-dependent kinase inhibitor 2A
CDKN2B	Cyclin-dependent kinase inhibitor 2B
CDKN2C	Cyclin-dependent kinase inhibitor 2C
CDKN2D	Cyclin-dependent kinase inhibitor 2D
CDR	(Complementarity Determining Region) - Regiones hipervariables
CG	Centro germinal

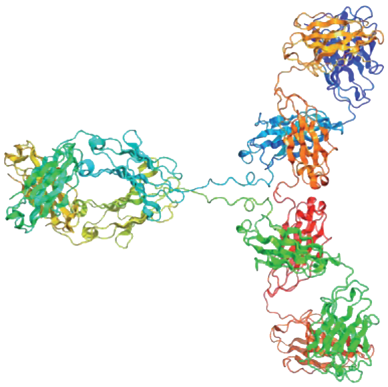
CIITA	Class II, major histocompatibility complex, transactivator
CMF	Citometría de flujo
CN arrays	Copy number arrays
CNN-LOH	Copy neutral number-loss of heterozycosity) - Pérdidas de heterocigosidad con número de copias normal
COO	(Cell of origin) – célula de origen
CREBBP	CREB-binding protein
CSR	Class-switch recombination
CTL	(Cytotoxic T lymphocite) - Linfocitos T citotóxicos
CXCR4	CXC-chemokine receptor 4
D	Diversity
DZ	Dark zone B terminal
EBER	EBV-encoded RNA-1
EBV	(Eptein-Barr virus) - virus de Epstein Barr
ECOG	"Eastern Cooperative Oncology Group"
EMR	Enfermedad mínima residual
EP300	E1A-binding protein p300
EPOCH-R	Etoposide, prednisone, vincristine, cyclophosphamide and doxorrubicin with rituximab
ERK	Extracellular signal-regulated kinase
ETS	E26 transformation-specific o E-twenty-six
EZH2	Enhancer of zeste homologue 2
Fab	Región variable
FAS/TNFRSF6/CD95	Fas cell surface death receptor
Fc	Región constante
FISH	(Fluorescent in situ hybridization) – Hibridación in situ fluorescente
FR	Framework
GEP	(Gene expression profile) – perfil de expresión génica
GNA13	Guanine nucleotide binding protein (G protein), alpha 13
H3K27me3	Trimetilación de la lisina 27 de la histona 3
HHV-8	(Human herpes virus 8) – virus herpes humano 8
HLA	Human Leukocyte Antigen
Hsp70	Heat shock protein 70

Ig	Inmunoglobulina
IGH	(Immunoglobulin heavy locus) - Gen de la cadena pesada de las inmunoglobulinas
IGHV	Immunoglobulin heavy variable
IGHD	Immunoglobulin heavy diversity
IGHJ	Immunoglobulin heavy joining
IgL	Cadenas polipeptídicas ligeras de la Inmunoglobulina
IHQ	Inmunohistoquímica
IPI	Índice Pronóstico Internacional
IRAK4	Interleukin-1 receptor-associated kinase 4
IRF4	Interferon-regulatory factor 4
IRF8	Interferon-regulatory factor 8
J	Joining
JAK	Janus Kinase
STAT3	Signal transducer and activator of transcription 3
κ	Kappa
λ	Lambda
LB	Linfoma de Burkitt
LBDCG	Linfoma B difuso de célula grande
LBDCG, NOS	LBDCG not otherwise specified
LCM	Linfoma de células del manto
LDH	Lactado deshidrogenasa
LEZM	Linfoma esplénico de la zona marginal
LF	Linfoma folicular
LH	Linfoma de Hodgkin
LLC	Leucemia linfática crónica
LMO2	LIM domain only 2
LNH-B	Linfomas no hodgkin B
LPM	Linfoma primario mediastínico
LZ	Light zone
MALT1	Mucosa-associated lymphoid tissue lymphoma translocation protein 1
MAPK	Mitogen-activated protein kinase
MDM2	MDM2 proto-oncogene, E3 ubiquitin protein ligase

MEF2B	Myocyte enhancer factor 2B
MHC	Major histocompatibility complex
MICA	MHC class I polypeptide-related sequence A
MICB	MHC class I polypeptide-related sequence B
MIP	Molecular inversion probe
MIRHG1	miR-17-92 cluster host gene
MLL3	Myeloid/lymphoid or mixed-lineage leukemia 3
MM	Mieloma múltiple
MUM-1/IRF4	Melanoma associated antigen (mutated) 1/interferon regulatory factor 4
MW	Macroglobulinemia de Waldenstrom
MYC	v-myc avian myelocytomatosis viral oncogene homolog
MYD88	Myeloid differenciation primary response protein 88
NF-κB	Nuclear factor of kappa
NK	Natural killer
P2RY8	P2Y purinoceptor 8
PAAF	Punción aspiración con aguja fina
PAX5	Paired box 5
PB	pares de bases
PCR	(Polymerase Chain Reaction) - Reacción en cadena de la polimerasa
PD1	Programmed Death-1
PDL1/CD274	Programmed Death Ligand 1 - ligando del receptor de la molécula de muerte celular programada 1
PDL2/CD273	Programmed Death Ligand 2 - ligando del receptor de la molécula de muerte celular programada 2
PET	Tomografía de emisión de positrones
PI3K	Phosphoinositide 3-kinase
PI3K/AKT	Phosphoinositide 3-kinase / protein kinase B (PKB)
PIM1	Pim-1 proto-oncogene, serine/threonine kinase
PKCß	Protein kinase Cß
PRMD1	PR domain containing 1, with ZNF domain
PSMB8	Proteasome subunit beta 8
PSMB9	Proteasome subunit beta 9
RB1	Retinoblastoma 1

R-CHOP	Rituximab, ciclofosfamida, doxorrubicina, vincristina y prednisona
RSS	(Recombination signal sequences) - secuencia señal de recombinación
S	Switch
S1PR2	Sphingosine-1-phosphate receptor 2
SG	Supervivencia global
Skp2	S-phase kinase-associated protein 2, E3 ubiquitin protein ligase
SLP-B	Síndromes linfoproliferativos B
SNC	Sistema nervioso central
SNP	(Single Nucleotide Polymorphism) - polimorfismo de nucleótido único
STAT	Signal transducer and activator of transcription
SYK	Spleen tyrosine kinase
TAP1	Transporter associated with Antigen Processing 1
TAP2	Transporter associated with Antigen Processing 2
TCR	(T cell receptor) - Receptor de célula T
TdT	Transferasa terminal de desoxinucleótidos
TLR	Toll-like receptors
TNF	(Tumor necrosis factor) - Factor de necrosis tumoral
TNFAIP3	Tumor necrosis factor, alpha-induced protein 3
TNFRSF14	Tumor necrosis factor receptor superfamily, member 14
TNFSF9	Tumor necrosis factor (ligand) superfamily, member 9
TP53	Tumor protein p53
TRAIL/TNFSF10	Tumor necrosis factor (ligand) superfamily, member 10
V(D)J	Variable (Diversity) Joining
VEGFR2	Vascular endothelial growth factor receptor 2
VHB	Virus de la Hepatitis B
VHC	Virus de la Hepatitis C
VIH	Virus de la inmunodeficiencia
XBP1	X-box binding protein

Introducción



1.
SÍNDROMES LINFOPROLIFERATIVOS B

Los Síndromes Linfoproliferativos B (SLP-B) crónicos tienen en común la proliferación de células linfoides maduras B. Aunque este término englobaría a toda proliferación linfoide B, generalmente se restringe a las formas primariamente leucémicas como la leucemia linfática crónica (LLC), leucemia prolinfocítica de origen B o tricoleucemia. El resto de SLP-B que presentan un predominio de infiltración nodal se agrupan dentro del término de linfoma B, si bien algunos de los subtipos de linfomas como el linfoma de células del manto presentan con frecuencia leucemización. Dentro de los linfoma B se engloban: a) linfomas no hodgkin B (LNH-B), entre cuyos subtipos más representativos se encuentran, linfoma folicular (LF), linfoma B difuso de célula grande (LBDCG), linfoma de burkitt (LB), linfoma de células del manto (LCM), linfoma esplénico de la zona marginal (LEZM), y el b) linfoma de hodgkin (LH). Desde el punto de vista clínico hay un grupo que se considera indolente y otros de comportamiento agresivo, entre los que se encuentra el LBDCG que es el subtipo histológico objeto del presente trabajo de tesis doctoral.

1.1. Linfoma B difuso de célula grande. Generalidades

1.1.1. Definición y clasificación

El LBDCG es una neoplasia de células B grandes con un patrón de crecimiento difuso. Es el linfoma más frecuente en el adulto y constituye el paradigma del linfoma agresivo. El LBDCG representa un grupo biológicamente heterogéneo. En la última década estudios moleculares han permitido aumentar el conocimiento sobre la biología de los diferentes tipos de LBDCG. Esto queda reflejado en la actual clasificación de la Organización Mundial de la Salud (OMS) del LBDCG[1], ver Tabla I, en la que se definen las diferentes variantes, subtipos y subgrupos, según su morfología, inmunofenotipo y características moleculares. Algunos casos están reconocidos como entidades propias por sus características clínicas o patológicas. Sin embargo, la mayoría de los LBDCG son biológicamente heterogéneos y no tienen claros criterios de subdivisión; son los denominados LBDCG, NOS (*not otherwise specified*), que incluyen varios subgrupos moleculares e inmunohistoquímicos. En el LBDCG, NOS es en el que nos centraremos a partir de ahora.

Tabla 1.
Clasificación del Linfoma de células grandes según la Organización Mundial de la salud (OMS)

Linfoma B difuso de célula grande (LBDCG), *not otherwise specified*
Variantes morfológicas comunes
Centroblástica
Inmunoblástica
Anaplásica
Variantes morfológicas raras
Subgrupos moleculares
Tipo célula B de centro germinal (CGB)
Tipo célula B activada (ABC, *activated B-cell like*)
Subgrupos inmunohistoquímicos
LBDCG CD5-positivo
Tipo célula B de centro germinal (CGB)
Tipo célula B no centro germinal (no-CGB)
Subtipos de linfoma B difuso de célula grande
Linfoma B de célula grande rico en histiocitos/células T
LBDCG primario del sistema nervioso central (SNC)
Linfoma primario cutaneo difuso de celulas B grandes,tipo pierna
LBDCG EBV-positivo del anciano
Otros linfomas de células grandes B
Linfoma primario mediastínico (tímico) de células grandes B
Linfoma de célula B grande intravascular
LBDCG asociado con inflamación crónica
Granulomatosis linfomatoide
LBDCG ALK positivo
Linfoma plasmablástico
Linfoma B de células grandes en enfermedad de Castleman multicéntrico
Linfoma primario de cavidades
Casos “borderline”
Linfoma de célula B, inclasificable, con características intermedias entre linfoma B difuso de célula grande y linfoma de Burkitt
Linfoma de célula B, inclasificable, con características intermedias entre linfoma B difuso de célula grande y linfoma de Hodgkin

El LBDCG NOS constituye el 25-30% de los linfomas B no Hodgkin en el mundo occidental. Es más común en edad avanzada, pero también puede aparecer en niños y jóvenes. Algunas formas pediátricas[2] y la asociada a edad avanzada[3] e infección por virus de Epstein Barr (E, *Eptein-Barr virus*)[1,4,5] tienen rasgos clínico patológicos y moleculares característicos.

La etiología del LBDCG no es bien conocida. Normalmente es un linfoma de nuevo diagnóstico o primario, pero a veces se puede presentar como una transformación de un síndrome linfoproliferativo B indolente, como la LLC[6,7] o el LF[8,9].

El LBDCG puede también ocurrir en el contexto de una inmunodeficiencia primaria o adquirida, así como en receptores de trasplante de órgano sólido o en pacientes infectados con el virus de la inmunodeficiencia humana (VIH)[10–12], aunque en este subgrupo de pacientes suele ser más frecuente el LBDCG asociado a EBV. En pacientes con LBDCG la infección por EBV, estudiada mediante hibridación in situ (FISH, *fluorescent in situ hybridization*) de EBER, *EBV-encoded RNA-1*, es aproximadamente del 10%[13]. También se ha visto asociación con el virus de la Hepatitis B (VHB)[14,15] y el virus de la Hepatitis C (VHC)[16].

1.1.2. Características clínicas

Es una neoplasia agresiva, ligeramente más frecuente en varones que en mujeres. La mediana de edad de aparición se sitúa en la séptima década de la vida. Al diagnóstico la mayoría de los pacientes presentan adenopatías rápidamente progresivas. Para el estadiaje se emplea el sistema "Ann Arbor", por el que casi la mitad de los pacientes tienen un estadio I o II al diagnóstico. Puede debutar como enfermedad nodal o extranodal. El sitio extranodal más frecuentemente afectado es el tracto gastrointestinal. Otras infiltraciones extranodales frecuentes son la médula ósea, los testículos, el bazo, el tiroides y el pulmón. La infiltración de médula ósea está descrita en el 11-27%[17,18], pero el porcentaje de casos con infiltración medular puede aumentar cuando se añaden técnicas de citometría de flujo (CMF) y biología molecular (BM)[19,20]. La infiltración de la médula ósea en el LBDCG puede ser por la presencia de linfocitos B grandes, por linfocitos B pequeños, o por la presencia de ambas poblaciones. En los casos con presencia de linfocitos B pequeños en la médula ósea, el estudio molecular puede revelar el mismo clon que el del

LBDCG, lo que indica que se trataría de un SLP-B indolente que no ha sido diagnosticado y que ha sufrido una trasformación histológica a LBDCG a nivel nodal, o alternativamente la presencia de clones diferentes, indicando la presencia simultánea de dos neoplasias B en diferentes localizaciones[18,21].

Es una enfermedad curable en un 60 – 70% de los casos con los tratamientos actuales con inmunoquimioterapia, pero un tercio de los pacientes presentan enfermedad refractaria o recaen tras el tratamiento. Una de las razones es la heterogeneidad de esta neoplasia que engloba múltiples subgrupos, que reflejan el origen de la célula B en diferentes estadios de maduración. Estos subtipos difieren, no sólo en la expresión génica, sino también en las vías oncogénicas que llevan al desarrollo del tumor.

1.1.3. Caracterización del LBDCG

ANATOMÍA PATOLÓGICA, MORFOLOGÍA E INMUNOFENOTIPO

En la biopsia ganglionar del LBDCG se observa una invasión difusa por linfocitos B de tamaño grande, núcleo vesicular con nucléolos muy visibles y citoplasma basófilo. El fenotipo más común es CD19+, CD20+, CD79a+, CD45+, y ocasionalmente pueden expresar CD10 o CD5. En el LBDCG, NOS se han descrito tres variantes morfológicas comunes denominadas como variante centroblástica, inmunoblástica y anaplásica[1].

PERFILES DE EXPRESIÓN GÉNICA

En los últimos años, las tecnologías de alto rendimiento o *"high-thoughput"*, han aumentado nuestro conocimiento sobre la complejidad molecular del LBDCG. Con el uso del perfil de expresión génica (GEP), Alizadeh y colaboradores[22] identificaron diferentes subtipos de LBDCG que eran morfológicamente superponibles pero que tenían patrones de expresión que indicaban diferentes estadios de la diferenciación de la célula B, similar a las células B del centro germinal (CGB) o a las células B activadas (ABC, *activated-B cell*). El linfoma primario mediastínico (LPM) sería una tercera variante, que constituye actualmente un grupo aparte. Estas variantes moleculares del LBDCG se definen en función del perfil de expresión génica determinada mediante análisis de expresión de ARNm. De forma consistente con su célula de origen (COO, *cell of origin*), el LBDCG

CGB, muestra una alta expresión del regulador principal BCL6 y tiene el gen de las inmunoglobulinas mutado, presentando mutaciones "*ongoing*", mientras que el LBDCG ABC pierde la expresión de los genes restringidos al centro germinal, muestra activación de las vías de NF-κB y del receptor de célula B (BCR, *cell receptor*) y regulación positiva de los genes requeridos para la diferenciación plasmática, estando a su vez bloqueada dicha diferenciación.

Esta clasificación de COO por GEP identifica particularidades clínicas, biológicas y genéticas claras en los dos subtipos[22–26]. La mayoría de las series publicadas asignan un porcentaje de casos al tipo ABC entre el 46 y 48%[2,27,28]. Los pacientes con LBDCG ABC tienen un peor pronóstico cuando se compara con el subtipo CGB[23,29]. A pesar de que la introducción del rituximab ha disminuido el impacto pronóstico de esta clasificación, el subtipo ABC sigue presentando una peor respuesta a la terapia que el subtipo CGB[25]. Además, estudios más recientes han demostrado que los subgrupos de LBDCG tienen diferentes lesiones genéticas y/o vías de señalización dañadas, lo que nos proporciona una justificación genética para el desarrollo de potenciales dianas terapéuticas[30–32], aunque no siempre se traduce en un beneficio clínico o los resultados son controvertidos[30,33].

ESTUDIO INMUNOHISTOQUÍMICO

Paneles de inmunohistoquímica (IHQ) e hibridación in situ (FISH, fluorescent in situ hybridization)

Como se ha comentado previamente, se reconocen al menos dos variantes moleculares de LBDCG NOS, CGB y ABC. Debido a las dificultades técnicas para aplicar el GEP en la práctica clínica diaria, se han desarrollado varios algoritmos IHQ como sustitutos del GEP. El clásico algoritmo de Hans utiliza anticuerpos contra CD10, BCL6, y MUM-1/IRF4[34]. Los casos con expresión de CD10 en ≥30% de las células, son considerados tipo CGB, así como los casos que siendo CD10 negativos, expresan BCL6, y son negativos para MUM-1/IRF4. Con el desarrollo de nuevos anticuerpos, el panel se ha extendido y se han elaborado nuevos algoritmos, como "Choi", Tally" y "Visco-Youngs"[27,28,35]. El grado con el que la clasificación inmunohistoquímica, CGB *versus* no-CGB/ABC, se correlaciona con la COO definida por el GEP es variable[27,28,34–36], y si la clasificación inmunohistoquímica tiene valor pronóstico es controvertido[37,38]. Varios es-

tudios que examinan la reproductibilidad entre diferentes laboratorios indican que esta falta de concordancia podría deberse en parte a los aspectos técnicos, tipo de tejido, calidad de la muestra, así como, las distintas interpretaciones[39–41].

A pesar de que estos dos subtipos de LBDCG clasificados por inmunohistoquímica no están todavía reconocidos como entidades diferentes o consistentemente identificadas en los informes diagnósticos, la IHQ es un método factible para la subclasificación de estos casos y hay datos que sugieren que el fenotipo CGB *versus* no-CGB/ABC pudiera ser útil cuando existan tratamientos más efectivos, por ello se recomienda incluir al menos el fenotipado según el algoritmo diagnóstico de Hans para el LBDCG[34]. Adicionalmente hay evidencia de que la coexpresión mediante inmunohistoquímica de C-MYC y BCL2 en casos de LBDCG (C-MYC ≥40%, BCL2 ≥70%) identifica un subgrupo de LBDCG, denominados LBDCG con "*double hit* inmunohistoquímico", de conducta clínica especialmente agresiva[42–44]. Se recomienda incluir la determinación de la expresión de C-MYC y BCL2, identificando el porcentaje estimado de células positivas. Asimismo, la expresión de CD30 en casos de LBDCG se asocia con rasgos clínicopatológicos y moleculares específicos[45]. La potencial disponibilidad de terapias específicas frente a esta molécula hace aconsejable identificar la positividad y el porcentaje estimado de células positivas para CD30 en cada caso de LBDCG. Estas subdivisiones pronósticas, independientes del índice IPI, no tienen hoy mucha aplicabilidad clínica fuera de los ensayos clínicos, ya que no se utilizan estrategias terapéuticas diferentes para los distintos subgrupos de pacientes.

Existe un conjunto de subtipos de linfomas B de células grandes con morfología inmunoblástica y plasmablástica que se asocian con rasgos clinicopatológicos distintivos. Para su identificación se requiere un panel que permita la identificación de un fenotipo específico (pérdida de marcadores de línea B y adquisición de marcadores de diferenciación terminal) asociado con marcadores propios de las diferentes entidades (EBV-EBER y C-MYC en el caso de linfomas plasmablásticos, HHV-8, *human herpes virus 8,* en el caso de linfomas de cavidades, por ejemplo). Se recomienda el uso de un panel específico de segunda línea en estos casos para subclasificar adecuadamente la neoplasia, así como en los casos de LBDCG asociado a EBV y edad avanzada, granulomatosis linfomatoide y síndrome linfoproliferativo post-trasplante de tipo LBDCG, en los que la demostración de EBV-EBER o una inmunotinción positiva para EBV-LMP1 son un requisito para el diagnóstico. Así, en un contexto diagnóstico en la práctica clínica se recomienda[46]:

Panel de primera línea:

– CD45 (ALC), CD20, CD3. En caso de terapia previa con antiCD20 es útil el uso de otros marcadores de línea B como PAX5, OCT2, CD79.

Paneles de segunda línea:

– Necesario en casos con el probable diagnóstico de LBDCG NOS: CD10, BCL6, MUM1, BCL2, C-MYC, Ki67, ciclinaD1.
– Aconsejable en casos con el probable diagnóstico de LBDCG NOS: CD30.
– Necesario en casos con morfología inmunoblástica/plasmablástica: CD138, CD38, Ki67, C-MYC, HHV-8, ALK, EBV-EBER (en su defecto EBV-LMP1).
– Necesario en casos con edad avanzada (>50 años), granulomatosis linfomatoide y desorden linfoproliferativo post-trasplante de tipo LBDCG: EBV-EBER (en su defecto EBV-LMP1).

Estudio de inmunoglobulinas como marcador de clonalidad

Una de las aplicaciones más útiles de las técnicas de inmunohistoquímica es la demostración de expresión de cadenas ligeras citoplasmáticas como marcador de clonalidad. Las células B linfoides neoplásicas, muestran salvo escasas excepciones la expresión de un receptor, que es una inmunoglobulina como veremos posteriormente, convenientemente modificada. La inmunoglobulina como receptor consta de dos cadenas pesadas y dos ligeras. En condiciones fisiológicas el proceso de exclusión alélica selecciona en cada célula una única forma de cadena ligera, kappa o lambda. De forma fisiológica se encuentran en proporción 2:1. La detección en un tejido de una población exclusiva o predominante con expresión de una de estas cadenas es altamente sugestivo de clonalidad.

BIOLOGÍA MOLECULAR Y CITOGENÉTICA/FISH

El LBDCG está constituido por células B grandes que albergan un reordenamiento clonal de los genes de las inmunoglobulinas (Igs).

Análisis de clonalidad linfoide (reordenamientos de Ig) mediante PCR y electroforesis capilar

En contadas ocasiones (abundancia de necrosis, mala calidad de la muestra, aberraciones inmunofenotípicas inesperadas) puede ser necesaria la ayuda de un estudio molecular que confirme la clonalidad B de la muestra. Para ello, la técnica recomendada es la amplificación de la región variable de los genes de las inmunoglobulinas y su estudio por análisis del tamaño de los fragmentos o su secuenciación[47,48]. Este estudio es siempre complementario al estudio morfológico e inmunohistoquímico de la muestra, y nunca debe utilizarse en solitario para asignar un diagnóstico.

Citogenética

La detección de reordenamientos de *MYC*, *BCL2* y *BCL6* utilizando sondas FISH, tipo BA, *break apart*, es de utilidad para identificar casos con alteraciones citogenéticas múltiples asociadas con mal pronóstico clínico y potenciales candidatos a terapias no estándar[49]. La incidencia de la t(14;18), *IGH/BCL2* es del 15-30%, y pueden derivar de un LF previo. La incidencia de t(3;V)(q27;V), *BCL6* es del 20-30% y puede tener gran variedad de parejas de reordenamiento. La t(8;14)(q24;q32) *IGH/MYC* se presenta en aproximadamente 10% de los LBDCG. La presencia de translocaciones en estos genes, particularmente en *MYC* y *BCL2* se asocia estrechamente con la sobreexpresión de la proteína[42,44,50]. Se sugiere por ello analizar mediante FISH *MYC* y *BCL2* aquellos casos con sobreexpresión de MYC en al menos el 50% de la población neoplásica. Si *MYC* se encuentra reordenado con o sin reordenamiento de *BCL2* mediante FISH, es de utilidad identificar el posible reordenamiento de BCL6[51]. Los LBDCG que presentan simultáneamente reordenamientos en *MYC* y *BCL2,* se conocen como *double hit*, y suelen presentar un fenotipo CGB. Los LBDCG *triple hit* se presentan asociando reordenamiento de *MYC*, *BCL2* y *BCL6*. Los LBDCG *double y triple hit* tienen un pronóstico desfavorable cuando *MYC* está implicado[49,52].

Cerca de un 2% de los casos de LBDCG sobreexpresan ciclina D1 mediante IHQ[53]. Si se observa la sobreexpresión homomogénea e intensa de ciclina D1 mediante IHQ es necesario realizar FISH de *CCND1* para descartar una forma pleomórfica de linfoma de células del manto. La presencia de traslocaciones de *MYC* y *ALK* está descrita en algunos subtipos de linfoma B de células grandes de morfología inmunoblástica/plasmablás-

tica[54,55]. La búsqueda de traslocaciones de *MUM1/IRF4* puede identificar un subgrupo de Linfomas B de fenotipo centrogerminal (LBDCG y LF), que se presenta en jóvenes, con una conducta clínica relativamente favorable[2]. El estudio de cariotipo puede ser de utilidad para identificar nuevos reordenamientos no conocidos, anomalías secundarias, potenciales dianas terapéuticas y en el diagnóstico diferencial de linfomas de características intermedias, en los que los cariotipos son muy complejos a diferencia de los cariotipos del linfoma de Burkitt.

ALTERACIONES CROMOSÓMICAS EN EL LBDCG: RELACIÓN ENTRE LOS DOS SUBTIPOS PRINCIPALES

En estudios recientes se ha demostrado que los dos subgrupos de LBDCG tienen diferentes alteraciones cromosómicas y/o vías de señalización, siendo algunas de ellas características de un subtipo, pero también hay alteraciones que se pueden dar en ambos grupos.

A continuación describimos algunas de las alteraciones genéticas que se encuentran más frecuentemente en cada uno de los dos subtipos CGB vs no-CGB, y aquellas que son compartidas por ambos[56,57]. El linfoma mediastínico está considerado como entidad propia. En la Figura 1 se pueden ver estas vías oncogénicas.

LINFOMA B DIFUSO DE CÉLULA GRANDE CGB

Hasta hace poco, solo se habían encontrado algunas lesiones preferencialmente asociadas con el LBDCG CGB, como las translocaciones de *BCL2* y de *MYC*. El oncogén antiapoptótico *BCL2* (18q21) está alterado muy frecuentemente en el LBDCG. La translocación cromosómica t(14;18)(q32;q21) es la más común en el LBDCG CGB, detectándose en 30%-40% de los casos[58], pero no es específica del LBDCG. Está presente también en más del 90% de los linfomas foliculares, y algunos de los LBDCG CGB podrían representar una transformación de un linfoma folicular clínicamente silente. Sin embargo, es muy infrecuente que el gen *BCL2* se encuentre translocado en los LBDCG ABC, encontrándose en este subgrupo con más frecuencia ganado o amplificado (30-40% de los casos frente al 15% en el LBDCG CGB[58,59]). Las mutaciones del gen *BCL2* también son frecuentes en el LBDCG CGB[60,61]. Las traslocaciones de *MYC* se dan en el 10 a 14% de los casos de LBDCG CGB. Ambas translocaciones, conducen a una sobreexpresión

de la proteína implicada, en parte por permitir el escape de la represión transcripcional mediada por BCL6[62,63]. Otras de las alteraciones ya conocidas en el LBDCG CGB incluye la inactivación de *PTEN*: aproximadamente el 10% de los LBDCG CGB tienen deleción del locus 10q23 que contiene el gen supresor *PTEN* [59], y hasta el 55% de los LBDCG CGB y el 14% de los LBDCG ABC presentan inactivación de *PTEN*. La inactivación de PTEN parece correlacionarse inversamente con la activación de la vía oncogénica (PI3K/AKT)[64].

Otras lesiones en el LBDCG CGB:

El locus en el cromosoma 2p (2p12-26) está frecuentemente amplificado en el LBDCG CGB y en el linfoma B difuso primario mediastínico, y muy raramente en el LBDCG ABC, pero su significado funcional no se conoce todavía. No hay datos claros que vinculen la presencia de la lesión genómica con los niveles de expresión o actividad de los dos supuestos genes diana, *REL*, que codifica una subunidad del factor de transcripción NF-κB, y *BCL11A*, que codifica una proteína dedo de zinc que puede interactuar con BCL6[29,65–67].

Se han identificado amplificaciones en el locus *MIRHG1* (13q31.3) en aproximadamente 12% de los LBDCG CGB, produciendo una sobreexpresión del micro RNA miR-17-92 que se relaciona con reordenamientos o amplificaciones de *MYC*,[59,68–74] siendo este cluster de miRNA también un regulador negativo de *PTEN*.

Las ganancias en el cromosoma 12, que normalmente provocan aumento del gen *MDM2*, que codifica un regulador negativo para el supresor de tumores *TP53*, son mucho más frecuentes en el LBDCG CGB que en el ABC. De forma similar, las ganancias del cromosoma 7 (especialmente las ganancias que afectan a su integridad) son más frecuentes en el subtipo de CGB. El efecto patogénico de las ganancias del cromosoma 7 no está claro, pero podría implicar la sobreexpresión de una serie de miRNAs en el cromosoma, especialmente miR-96, miR-182, miR589 y miR-25. La presencia de ganancias de 7q parece predecir una mejor supervivencia y una menor probabilidad de infiltración de médula ósea[75,76].

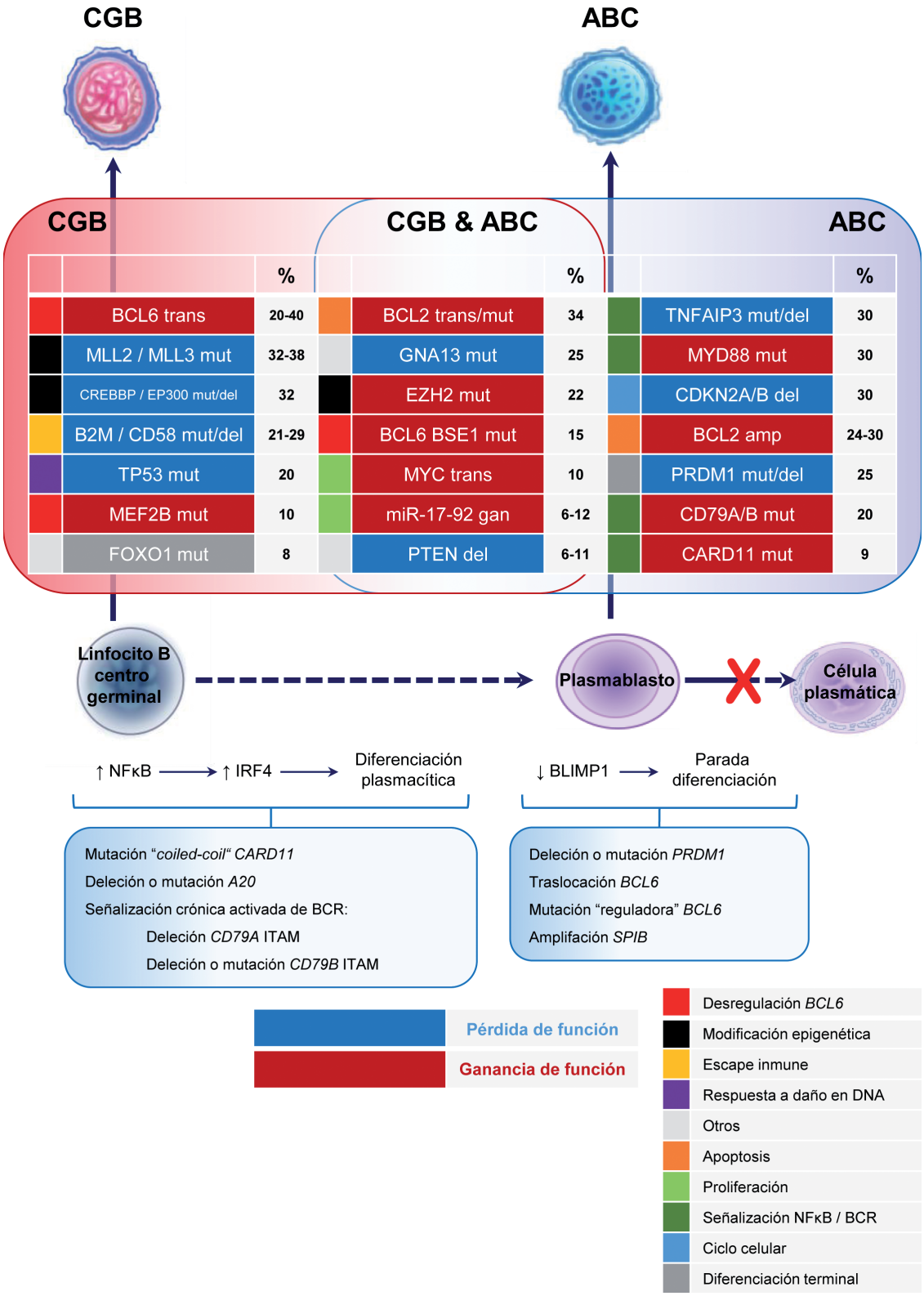


Figura 1. Patogénesis de los dos subtipos de LBDCG: de células B del centro germinal, CGB, y célula B activada, ABC (modificado de Lenz M & Staudt LM NEJM 2008, y Pasqualucci L & Dalla Favera R Sem Hematol 2015). Se ilustra esquemáticamente la célula de origen de cada subtipo. Se muestras las alteraciones más comunes específicas y compartidas de los dos subtipos, y su frecuencia. El código de colores indica la vía biológica a la que pertenece cada gen. La pérdida o ganancia de función de estas alteraciones se ilustra en azul o rojo respectivamente.

Mutación de *EZH2* y otros modificadores de la cromatina: uno de los modificadores epigenéticos más frecuentemente mutado es EZH2, cuyo gen presenta una ganancia de función en el 6-14% de los LBDCG[9,60]. Su alteración se encuentra de forma exclusiva en el LBDCG de CGB[77,78]. En la mayoría de los casos, las mutaciones de *EZH2* consisten en el cambio de un residuo conservado en la evolución (Tyr641) en el dominio SET de la proteína, lo que lleva un aumento de la capacidad catalítica y de los niveles de H3K27me3[77]. La mutación de EZH2 se produce concretamente en los LBDCG CGB que llevan reordenamiento de *BCL2*[79]. Así, la asociación de mutaciones en *EZH2* con la presencia de t(14;18) se da en 20% de los casos de LBDCG CGB y no ocurren en el LBDCG ABC[77,79]. De acuerdo con esto, la expresión de alelos mutados de EZH2 en ratones promueve hiperplasia de los CGs y coopera con BCL2 en la inducción de LBDCG[80]. Especialmente, moléculas inhibidores de EZH2 han sido testados como diana molecular prometedora[81,82]. Otros genes de este grupo mutados en el LBDCG son: *MLL2* (22-32% de los LBCDG), *CREBBP* (18-20% de los LBDCG), *EP300* (5-10%) y *MLL3* (15%) [9,60,77,78,83,84]. Debido a la baja frecuencia de estas mutaciones y a la heterogeneidad de las series estudiadas, es difícil establecer una asociación con algunos de los subtipos de LBDCG, como se describe posteriormente.

Mutaciones en la vía Ga13: aproximadamente el 30% de los LBDCG CGB llevan mutaciones estructuralmente nocivas en componentes del circuito inhibitorio acoplado a la proteína G, que regula el crecimiento y la reclusión local de las células B del CG (así como *GNA13*, *S1PR2*, entre otros)[85]. La pérdida de estos genes en ratones se ha asociado con el aumento de supervivencia de las células B del centro germinal y la diseminación linfática y a la médula ósea, conduciendo finalmente al desarrollo del linfoma[85].

LINFOMA B DIFUSO DE CÉLULA GRANDE ABC

El escenario genético del subtipo ABC es el más conocido, y está asociado con alteraciones genéticas que convergen en dos vías principales: activación de NF-κB y bloqueo de la diferenciación terminal B. Además, lesiones recurrentes en este subtipo son la amplificación de *BCL2* y deleciones o pérdidas de expresión de los genes supresores *CDKN2A/2B*[59].

a. Lesiones genéticas que impiden la diferenciación terminal de la célula B

La desregulación del proceso normal del desarrollo del centro germinal representa uno de los mecanismos principales de la patogénesis del LBDCG, habitualmente por lesiones genéticas de *BCL6*, clave en la diferenciación del centro germinal, y *BLIMP1*, el principal regulador de la diferenciación de las células plasmáticas[26,86]. Así, la diferenciación de las células del CG en células plasmáticas requiere BLIMP1, un represor transcripcional específico que está aumentado en un subgrupo de células B de la zona clara del CG preparadas para la diferenciación plasmática y también en todas las células plasmáticas[87]. Sin embargo, aproximadamente el 25% de los LBDCG ABC tienen inactivación de BLIMP1, por mutaciones o por deleciones de la región 6q21 que contiene el gen que codifica para BLIMP1[86,88]. Además, BCL6 es uno de los principales represores directos de BLIMP1. Las alteraciones de BCL6, generalmente su activación constitutiva por translocaciones o mutaciones, representan un mecanismo alternativo de la desregulación de BLIMP1[86]. Las lesiones genéticas que afectan a *BCL6* (3q27) son más frecuentes en el LBDCG ABC y son mutuamente excluyentes con las alteraciones de *BLIMP1*[9,88,89], lo que apoya su papel complementario en la linfomagénesis bloqueando la diferenciación terminal. Se cree que las translocaciones o mutaciones en *BCL6* contribuyen al desarrollo de LBDCG ABC bloqueando la diferenciación post-CG, y estos eventos probablemente cooperan con la activación constitutiva de NF-κB[86,90].

SPIB se encuentra en el cromosoma 19 (19q13.3-q13.4) y es diana de translocaciones y mutaciones en el LBDCG ABC[59,91]. Es un factor de transcripción de la familia ETS necesario para la completa señalización del receptor de célula B y para la respuesta a anticuerpos dependiente de célula T. SPIB es también una diana directa de BLIMP1 y su expresión es mayor en el LBDCG ABC que en el CGB[29]. Su silenciamiento contribuye por tanto al desarrollo de LBDCG ABC[59].

b. Lesiones genéticas que llevan a la activación constitutiva de NF-κB

La activación constitutiva de la vía NF-κB es la alteración más importante en el LBDCG ABC. Se evidenció por el aumento de expresión de los genes diana de NF-κB y el requerimiento de NF-κB para la proliferación de las líneas celulares del LBDCG ABC, pero no de las del subtipo CGB[92]. Diversos estudios han proporcionado posteriormente

evidencia de alteraciones genéticas, en moléculas cuya función común es inducir activación de NF-κB. Aunque algunos genes son inactivados por mutaciones somáticas y deleciones como *TNFAIP3* (30% de los casos), otros como *CARD11* (10% de los casos), *CD79B* y *CD79A* (20 % de los casos) y *MYD88* (30% de los casos), son activados mayoritariamente por mutaciones somáticas[26,31,61,78,93–95].

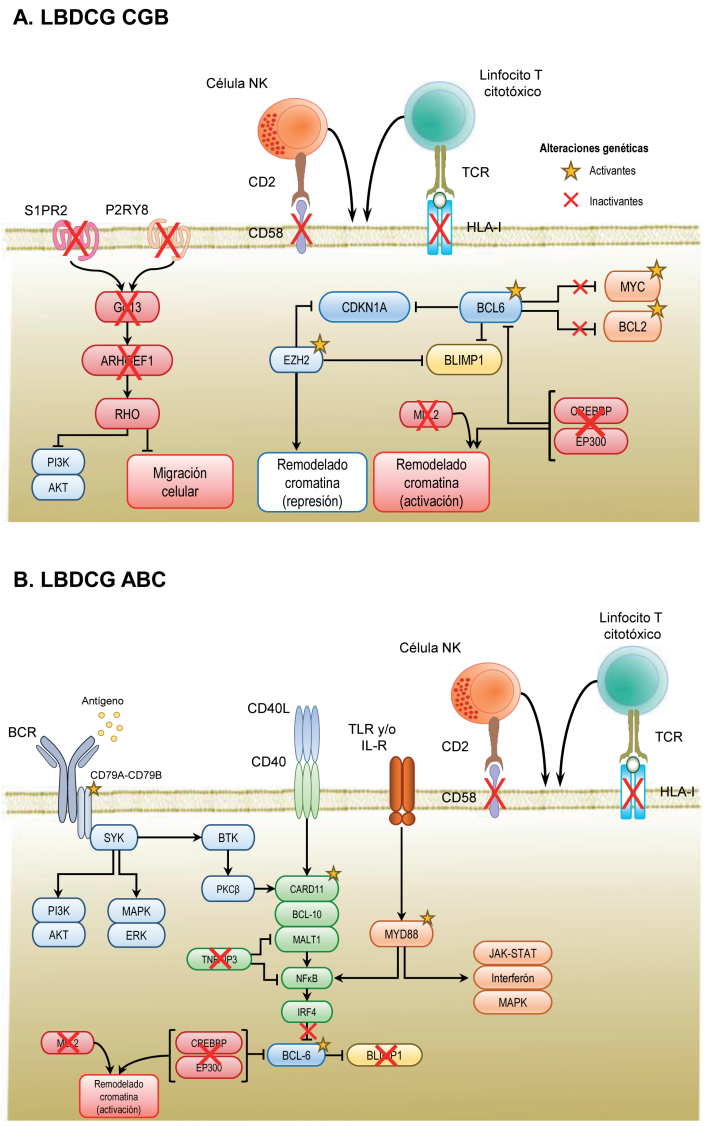


Figura 2. Representación esquemática de las vías afectadas durante la linfomagénesis en el LBDCG (modificado de Basso K & Dalla-Favera R. Nat Rev Immunol 2015). a. Vías afectadas en el LBDCG CGB. El LBDCG CGB se asocia con inactivación de la vía Gα13, que regula negativamente la migración de las células del CG y la señalización de PI3K. Además, las mutaciones activantes de EZH2, que promueve cambios epigenéticos que cooperan con el programa transcripcional de BCL6. Un subgrupo de LBDCG CGB es caracterizado por la translocación que afecta a MYC y/o BCL2. b. Se muestran las vías que afectan a los LBDCG ABC. Este subtipo se caracteriza por la activación de NF-κB por múltiples alteraciones genéticas- incluyendo las mutaciones activadoras de CD79A o CD79B, CARD11, y MYD88, e inactivación de TNFAPI3, que codifica A20 y un bloqueo en la diferenciación, que es mediado por la inactivación de BLIMP1. Ambos subtipos LBDCG CGB y ABC comparten lesiones genéticas que llevan al escape inmunológico, debido a la inactivación de β2M, CD58 y los genes que codifican el HLA-A, HLA-B y HLA-C, y la inactivación de modificadores de la cromatina, debido a las mutaciones de MLL2, CREBBP, y EP300. Los diferentes colores indican las moléculas que pertenecen a una vía específica.

c. Mutaciones activadoras de la vía de señalización del BCR.

Las células B maduras requieren el estímulo del BCR para sobrevivir, sin embargo, las células del LBDCG ABC muestran una activación crónica de la vía del BCR, que se mantiene por alteraciones genéticas en genes proximales de esta vía.

Más del 20% de pacientes con LBDCG ABC tiene mutaciones en los componentes del BCR CD79B, y en menor frecuencia CD79A[94]. Estas mutaciones eluden el *feedback* negativo de los circuitos que atenúan la señalización del BCR, manteniéndolo activo de forma crónica.

En aproximadamente el 9% de los LBDCG ABC (y un pequeño grupo del LBDCG CGB), mutaciones oncogénicas de *CARD11*[31], inducen la activación constitutiva del BCR y NF-κB, en ausencia de señales de los receptores de antígeno (como CD40-CD40L). CARD11 es uno de los principales componentes (junto con BCL10 y MALT1) del complejo proteico, cuyo reclutamiento coordinado es necesario para la correcta señalización del BCR[96]. Como la activación del BCR, puede desencadenar múltiples cascadas de señalización (PI3K, ERK/MAP kinasa), además de NF-κB, es de esperar que varias vías contribuyen a la transformación neoplásica de las células el LBDCG ABC, como sugieren experimentos iniciales que demuestran la toxicidad cooperativa de los inhibidores de NF-κB y PI3K[31].

d. Mutaciones activantes de la vía de TLR (Toll-like receptors)

El gen codificante de MYD88, una proteína adaptadora que media la señalización de los receptores *toll* e interleuquina-1, está mutado hasta en un tercio de los LBDCG ABC[95,97]. Las mutaciones de *MYD88*, representadas comúnmente por la mutación L256P en un residuo invariable en el dominio TIR (Toll/IL1 receptor), e inducen la actividad de IRAK4 kinasa, lo cual promueve sucesivamente no sólo la activación de la vía NF-κB, sino también la activación de la vía JAK-STAT3.

e. Mutaciones que inactivan los reguladores negativos de NF-κB

TNFAIP3 está mutado o delecionado en el 20-30% de todos los LBDCG, aunque la mayoría ocurren en el subtipo ABC, y también se encuentra inactivado en otros linfo-

mas[93,98]. *TNFAIP3* codifica para A20 un regulador negativo de NF-κB, implicado en la finalización de la respuesta desencadenada por la estimulación de BCR y TLR.

En resumen, múltiples alteraciones, desregulan la cascada NF-κB, a diferentes niveles, lo que ofrece una oportunidad para desarrollar estrategias terapéuticas; así como el inhibidor de BTK (*Brutn tyrosine kinasa),* que sería eficaz en los casos con mutaciones en la vía del BCR por encima de BTK[99].

f. Lesiones genéticas en genes del ciclo celular

La deleción del locus 9p21 *INK4/ARF* es más común en el LBDCG ABC, y se objetiva en el 24-30% de los casos[59,100,101]. El locus *INK4/ARF* codifica tres supresores de tumores (*CDKN2B, ARF y CDKN2A*) y está entre los loci más frecuentemente inactivados en los cánceres humanos. CDKN2A, junto con CDKN2B, CDKN2C y CDKN2D, inhibe selectivamente las quinasas asociadas a ciclina-D, dando lugar a un bloqueo del ciclo celular en la fase G1-S. ARF es una proteína que estabiliza TP53, ya que previene su degradación por el proteasoma, mediada por MDM2. Los casos de LBDCG con del(9p21) parecen asociarse con peor pronóstico en el subtipo ABC. Además esta lesión parece contribuir a la transformación de síndromes linfoproliferativos indolentes a LBDCG, tanto la LLC (síndrome de Richter)[7,102], como el linfoma folicular[103,104].

LESIONES COMUNES A LINFOMA B DIFUSO DE CÉLULA GRANDE ABC Y CGB

a. Regulación de la transcripción y de la remodelación de la cromatina

Las mutaciones en genes implicados en la remodelación de la cromatina son recurrentes en el LBDCG (*CREBBP*, *EP300*, *MLL2* y *MLL3)*. Ninguna de estas mutaciones es específica de enfermedad o subtipo, siendo detectadas también en otros linfomas y ocurriendo tanto en LBDCG CGB como ABC, aunque en frecuencias diferentes. Los genes de las acetiltransferasas CREBBP y EP300, que modulan la actividad de gran número de factores de transcripción, son inactivados por mutaciones o pérdidas en aproximadamente el 15% y 5% respectivamente[78,83], contribuyendo a la activación de BCL6 y a la inactivación de TP53 en el LBDCG. La actividad balanceada de estas proteínas es clave en la regulación del daño del ADN durante la recombinación de las inmunoglobulinas en

el CG[105]; así, una consecuencia de la actividad de bcl6 predominante sobre p53 sería un aumento de la tolerancia a la inestabilidad genética en un contexto con respuestas apoptóticas dañadas. Estas mutaciones habitualmente son heterocigotas, permaneciendo la expresión del alelo "*wild type*", lo que sugiere su papel como genes supresores de tumor haploinsuficientes. De hecho la perdida de solo unos de los alelos *CREBBP (*o *EP300)*, es la casusa de una rara enfermedad congénita que cursa con el desarrollo de defectos y predisposición a tumores (síndrome de Rubinstein-Taybi) lo que proporciona evidencia sobre el efecto patogénico dosis dependiente de estos genes[106].

En el 10-18% de los LBDCG hay mutaciones recurrentes del gen que codifica el factor de transcripción MEF2B, y recientemente se ha demostrado que este gen activa la transcripción de *BCL6* en las células B normales del CG[107].

Al menos un tercio de los LBDCG presentan mutaciones características en *MLL2* (*mixed lineage lymphoma/leukemia 2*)[83]. *MLL2* codifica para una metiltransferasa, a pesar de que las consecuencias de las mutaciones de *MLL2* en el LBDCG, no son bien conocidas, parece que pueden llevar a generar una proteína truncada, que pierde el dominio catalítico SET, necesario para su actividad metiltransferasa. Así las mutaciones de *MLL2* tienen probablemente una repercusión en la regulación de la cromatina, que podría contribuir a la linfomagénesis, mediante la reprogramación del epigenoma de la célula tumoral.

Es importante destacar, que las mutaciones inactivantes de *CREBBP* y *MLL2* se observan también en el LF (40% y 89% respectivamente)[9,78,83], siendo una de las alteraciones genética más frecuentes reportadas en el LNH-B hasta la fecha. Además, estudios recientes, que reconstruyen la historia de la evolución clonal de LF a LBDCG sugieren que estas mutaciones representan eventos tempranos comunes en un clon ancestral, previo a la transformación del LF (tFL)[103,108]. Las alteraciones en los modificadores epigenéticos podrían facilitar la transformación en estadios iniciales, creando un microambiente permisivo para la proliferación y supervivencia del clon tumoral.

b. Alteraciones en la regulación de BCL6

La desregulación de la actividad de BCL6 es un mecanismo clave en la tranformación del LBDCG, producida por múltiples mecanismos. Los reordenamientos de *BCL6*

caracterizan el 35% de los LBDCG, siendo más frecuentes en el subtipo ABC[89]. Estas recombinaciones producen una desregulación de la expresión de la proteína intacta, impidiendo su regulación negativa, durante la reacción del CG. La secuencia 5'de *BCL6* sufre también múltiples mutaciones en más del 70% de los casos[109,110], aunque esto refleja la actividad fisiológica de la hipermutación somática (SHM, *somatic hypermutation*) en el CG, hay un grupo de mutaciones que tienen como diana el primer exón no codificante de *BCL6* que están restringidas al linfoma, desregulando la expresión de BCL6.

c. Vigilancia inmunológica

El escape de la vigilancia inmunológica es un hecho necesario para el desarrollo del tumor[111]. Las células del LBDCG son conocidas por presentar regulación negativa del complejo mayor de histocompatibilidad (MHC, *major histocompatibility complex*), tanto MHC I y II, en la superficie de la célula. Mecanismos genéticos y epigenéticos regulan este proceso[112–114]. β2M (beta-2 microglobulina) forma el MHC I con la cadena pesada del antígeno leucocitario humano (HLA). El gen de la β2M está inactivado en un tercio de los casos de LBDCG[9,60,115] y en un porcentaje mucho mayor de los casos de LBDCG (75%) la proteína no está expresada o tiene un patrón de expresión aberrante[115]. Las lesiones genéticas de los genes que codifican para MHC II son comunes en los casos de LBDCG que se originan en santuarios inmunes como los testículos y el sistema nervioso central[112,114,116].

La regulación negativa del transactivador CIITA del MHC clase II, parece ser un mecanismo importante que contribuye a silenciar el MHC II[112,114,117]. La ruptura genética del gen *CIITA* por translocaciones cromosómicas es común en el linfoma primario mediastínico (el 38% muestran translocaciones que afectan dicho gen), pero las traslocaciones y pérdidas de DNA son raras en el LBDCG (<5%)[117,118]. Sin embargo el MHC, que se encuentra en 6p, es diana de pérdidas de heterocigosidad con número de copias normal (CNN- LOH, *copy neutral number-loss of heterozycosity*)[119]. La pérdida de heterocigosidad de 6p es mucho menos común en los LBDCG de los pacientes con una inmunodeficiencia adquirida, así como en los pacientes con VIH o en los receptores de trasplante de órgano sólido[120,121], situaciones en las que la vigilancia inmune mediada por células T está ya reducida.

Otras lesiones genéticas que podrían contribuir al escape inmunológico son: la inactivación y/o desregulación de CD58 (receptor de CD2)[9,115]; la desregulación de los ligandos del receptor de la molécula de muerte celular programada 1 (CD274/PDL1 y CD273/PDL2) por diferentes mecanismos[78,117], o alteraciones en la superfamilia de los receptores de TNF (TNFSF9, TNFRSF14, FAS/TNFRSF6/CD95, TRAIL/TNFSF10)[9,60,122–126].

La comparación entre los pacientes inmunocompetentes con LBDCG y los pacientes inmunocomprometidos nos puede ayudar a reconocer qué lesiones específicas tienen un papel importante en el escape inmune, del mismo modo que la CNN-LOH en 6p. Por ejemplo, la deleción de 13q14.3 que principalmente afecta miR-15A y miR-16 y generalmente abarca el gen supresor *RB1*, casi nunca se detecta en pacientes inmunocomprometidos[120,121], lo que sugiere que la del(13q14.3), podría contribuir a evitar el control del sistema inmune en las células linfomatosas, probablemente a través de FAS[127]. O la baja frecuencia ganancias de 3q, que se observan en pacientes sometidos a trasplante de órganos[121], o de 18q en pacientes con VIH[120], indican que estas lesiones podrían proporcionar mecanismos de escape a la vigilancia inmunológica antitumoral en estas situaciones particulares de inmunosupresión. Por otro lado, la inactivación de moléculas implicadas en la vigilancia inmune como FAS o los receptores TRAIL, protege a las células linfomatosas no sólo de las células del sistema inmune, sino también de la respuesta apoptótica que sigue a la quimioterapia, y esto podría explicar por qué algunos de estos defectos genéticos se han asociado con pobre respuesta al tratamiento[128,129].

d. Desregulación de MYC y linfomas double/triple-hit

Las translocaciones cromosómicas que desregulan *MYC* (8q24) ocurren en el 5-15%[130] de los casos con LBDCG, más frecuentemente en el LBDCG CGB. La t(8;14)(q24;32) desregula MYC por juntarlo con el promotor IGHV. Es una característica citogenética del linfoma de Burkitt[131], pero el gen *MYC* puede también translocarse a otros loci. Aunque se han propuesto varios *scores* inmunohistoquímicos y algunos paneles inmunohistoquímicos pueden predecir las traslocaciones de *MYC*[132], necesitan validarse en series más grandes[42,43,133]

La translocación de *MYC*, y no necesariamente la presencia de una copia extra, está asociada con un peor pronóstico[49,134–142] y a un subtipo de LBDCG caracterizado por

la presencia de múltiples translocaciones, que en su mayoría afectan también a *BCL2*, *BCL6* o *CCND1* (ciclina D1). Estos son llamados *double* o *triple-hit* linfomas y ocurren en un bajo porcentaje de casos de LBDCG (hasta el 5%) pero tienen un pronóstico muy adverso, a pesar de tener un fenotipo CGB[143,144]. Estos linfomas generalmente tienen morfología de LBDCG en el examen histológico, y sólo los análisis genéticos (FISH) son capaces de separar estas entidades. Una serie de 371 linfomas *double* y *triple hit* ha sido recientemente descrita[145]. La histología en el 50% era de LBDCG, el 48% tenía un diagnóstico de célula B no clasificable con características de intermedio entre LBDCG y linfoma de Burkitt, y el 2% de LF. La translocación *MYC* se produjo en el 87% de los casos con *BCL2*, con *BCL6* en el 5%, con ambos en el 7%, y el 93% expresaba el marcador de centro germinal CD10. La gran mayoría de los pacientes presentaban fallo al tratamiento y habían fallecido, excepto para un pequeño grupo de pacientes (7%) con LDH normal, sin leucocitosis, ni infiltración del SNC, estadio I y II, con una supervivencia del 91% a los dos años. Esta larga serie retrospectiva sugiere una posible ventaja para las dosis altas de quimioterapia seguidas de trasplante en comparación con el estándar de tratamiento R-CHOP.

e. Otras lesiones

Las mutaciones somáticas en el gen que codifica para el factor de transcripción FOXO1, implicado en la vía de señalización PI3K para la activación del receptor de célula B en el desarrollo del CG, se han encontrado en el 8.6% de los casos con LBDCG. Su significado no está todavía bien definido. Las mutaciones de *TP53* son también lesiones patogénicas importantes en el 20% de los LBDCG[9,101], también en los LBDCG secundarios a la transformación de LF[103].

1.1.4. Factores pronósticos en el LBDCG

FACTORES PRONÓSTICOS CLÍNICOS

El Índice Pronóstico Internacional (IPI) es la primera y más robusta herramienta pronóstica para predecir la supervivencia en el LBDCG[146]. Son cinco variables clínicas (edad> 60 años, estado general mediante la escala "*Eastern Cooperative Oncology Group*" (ECOG)>1, estadio de Ann Arbor avanzado, III-IV, lactato deshidrogenasa (LDH) elevada

y número de localizaciones extraganglionares >1) que establecen 4 grupos de riesgo, con supervivencia global a 5 años de 26-73%. En los pacientes con ≤ 60 años, el IPI ajustado a la edad (aaIPI) se calcula con tres de estas variables no favorables: que incluyen ECOG >1, estadio Ann Arbor avanzado, y LDH alta. La utilidad del IPI ha sido reevaluada en pacientes tratados con la adición de rituximab a la quimioterapia[147–149]. Uno de los cambios sugeridos ha sido la reorganización de los mismos parámetros en el R-IPI, que estratifica en tres los grupos de riesgo: muy favorable (ningún parámetro adverso), favorable (1-2 parámetros adversos) y desfavorables (3-5 parámetros adversos)[150]. Los niveles de beta-2 microglobulina en suero son también un factor de mal pronóstico según varias publicaciones independientes, tanto en la era pre como post rituximab[151–153]. El IPI clásico permanece como factor pronóstico en algunos estudios[147,148], aunque no en otros[149,154]. Aunque el IPI parece conservar su valor pronóstico, el rango de superviencia entre los grupos se ha reducido con 53% de supervivencia libre de evento (SLE) a los 4 años en pacientes con 4-5 factores de riesgo y 94% en pacientes sin factores de riesgo, y una supervivencia global (SG) de 55% a 4 años en el grupo de peor pronóstico y de 94% en los de mejor pronóstico[150], que han mejorado su pronóstico de forma significativa[155]. Se han analizado otros factores en la era del rituximab. Así, el recuento absoluto de linfocitos (ALC, *absolute count of lymphocite*),[156,157] la infiltración de la médula ósea, así como el patrón de la misma[158], y la proteína C reactiva (PCR)[159], parecen ser factores pronósticos independientes en algunos estudios, identificándose otros modelos como el ALC/R-IPI (Índice Pronóstico Internacional revisado, *The revised International Prognostic Index*). Sin embargo, el IPI o R-IPI, o el ALC/R-IPI[160], no son suficientes para identificar a los pacientes que tendrán una peor supervivencia y considerar un tratamiento diferente al estándar (R-CHOP), ya que todos los grupos de riesgo tienen al menos el 50% de probabilidad de cura[161]. Tampoco son capaces de diferenciar grupos pronósticos en estos pacientes[162].

Sin embargo, sí que identifica un grupo en el que la supervivencia global a 5 años es < del 50%, y en los que es prioritario investigar factores biológicos que puedan ayudar a predecir el pronóstico y a identificar nuevas dianas terapéuticas.

Es necesario validar nuevos factores pronósticos clínicos, como la presencia de masa *bulky* en pacientes jóvenes de buen pronóstico[163], la infiltración medular y su morfología[158], la influencia de factores farmacogenómicos y farmacodinámicos en relación

con el sexo del paciente[164], nuevos marcadores séricos como las cadenas ligeras libres en suero[165], o la evaluación precoz de la respuesta en función de la Tomografía de emisión de positrones (PET)[166–168].

FACTORES PRONÓSTICOS BIOLÓGICOS

Algo similar ocurre con los factores pronósticos biológicos, ya que excepto con la translocación simultánea de *MYC* y *BCL2*[49] (linfomas *double hit*), no hay parámetros biológicos validados que puedan ser usados para guiar la terapia en la práctica diaria. La aparición de nuevos fármacos y su combinación con el estándar actual rituximab + quimioterapia, y la descripción de nuevos subgrupos de pacientes según sus alteraciones genómicas, hacen necesaria esta validación de posibles factores biológicos. Por tanto la determinación del riesgo es un objetivo en movimiento[162].

a. Biomarcadores pronósticos individuales

En pacientes tratados con R-CHOP los siguientes rasgos biológicos parecen conferir pronóstico adverso: la expresión de BCL2 (en el LBDCG de CG), MYC (junto con BCL2), VEGFR2, Ki-67 "*high microvessel density*", CD5, Indoleamine 2,3-dioxygenase, Skp2, mutaciones de p53. Mientras que se asocian a mejor pronóstico los siguientes: LMO2, HIF-1∞ y p21[26,43,58,158,169,170]. La mayoría de estos factores todavía requiere una validación independiente, y el valor pronóstico es controvertido en diferentes estudios.

– *BCL2* codifica una proteína antiapoptótica, ejemplo de la dificultad de establecer el valor pronóstico de marcadores individuales en el LBDCG. Se ha descrito sobreexpresión de BCL2 en el 40-60% de los LBDCG y se ha asociado en diferentes estudios con un pronóstico desfavorable[130,171]. Sin embargo, la introducción de rituximab parece vencer el mal pronóstico de la sobreexpresión de *BCL2*, ya que no se ha visto una correlación con la supervivencia en los pacientes que reciben inmunoquimioterapia[172]. Esto se debe, a que la expresión de *BCL2* en pacientes tratados con R-CHOP es factor de mal pronóstico en los LBDCG CGB pero no en los ABC,[58] que es lo opuesto a lo que se había observado previamente en pacientes tratados sólo con CHOP[173]. Este hallazgo puede ser expli-

cado por el diferente mecanismo que produce la sobreexpresión de BCL2 en los diferentes subtipos y por la acción del rituximab[26]. En el subtipo CG suele ser debida a la t(14;18), mientras que en el LBDCG ABC, la hiperexpresión puede ser mediada por la activación de NF-κβ[174,175].

– *TP53* es un gen supresor de tumores que actúa como factor de transcripción multifuncional. Está implicado en la parada del ciclo celular, la apoptosis, la diferenciación, la replicación, la reparación del ADN y la estabilidad genómica[176,177]. Se han descrito mutaciones de *TP53* en 18-30% de pacientes con LBDCG[178]. Recientemente Young *et al*[179,180], identificaron mutaciones de *TP53* en el 21% de los pacientes con LBDCG, en los que la supervivencia global (SG) fue significativamente peor que en aquellos con el gen *wild-type*. El mayor predictor de supervivencia global pobre fue la presencia de mutaciones de *TP53* en los sitios de unión al ADN. Así las mutaciones de p53 parecen tener un factor pronóstico adverso[179,181,182], sin embargo, otros estudios no han podido demostrar el impacto pronóstico de p53[183]. La forma mutada de p53 presenta una vida media prolongada, a diferencia de la forma germinal, que le permite ser detectada mediante inmunohistoquímica. Sin embargo, la correlación entre las mutaciones de y la detección inmunohistoquímica es inexacta[178].

– *TP21 - CDKN1: cyclin-dependent kinase inhibitor 1A:* regula negativamente la progresión del ciclo celular e inhibe la proliferación celular, bajo el estrecho control de p53[184]. Su expresión puede ser controlada por mecanismos dependientes o independientes de TP53. Múltiples estudios han cuantificado por IHQ la expresión de TP53 y TP21. La expresión fuerte de TP53 en el núcleo sin expresión de TP21, se ha asociado con alteraciones genómicas de *TP53*, y se ha usado como un imperfecto sustituto de las mutaciones de *TP53*[185]. Como se ha comentado la IHQ de TP53 no tiene un claro valor pronóstico, pero la adición de TP21 al panel IHQ podría aumentar su valor pronóstico[185,186].

– *Moléculas MHC:* la pérdida de expresión de moléculas de clase I y II, se ha relacionado con mal pronóstico en pacientes con LBDCG [112,187]. El mecanismo para la pérdida de expresión no está claro ya que a nivel molecular no existe pérdida del gen, sugiriéndose que la falta de expresión es un mecanismo de es-

cape inmunológico al no ser reconocidas las células tumorales por las células inmunes del paciente. Curiosamente, la mayor incidencia de casos de LBDCG HLA-clase I y II negativos corresponde a linfomas que afectan a sitios "inmuneprivilegiados"[112].

– *PD1-PDL1 (Programmed Death 1/Programmed Death Ligand 1)*: La vía de PD-1 receptor (CD279) y su ligando PD-L1 y PD-L2 (B7-DC; CD273) juega un papel vital en la tolerancia[188] periférica pero también media las señales inhibitorias que comprometen la inmunidad antitumoral. La inhibición de la vía PD-1/PD-L1 ha mostrado eficacia en los tumores sólidos[189], pero también en linfomas. En tumores sólidos, la expresión de PD-L1 en las células tumorales es un pronóstico adverso. Estudios recientes en LBDCG han analizado el valor de la expresión de PD-L1 en las células linfomatosas (LBDCG PD-L1$^+$ o en otras células del microambiente tumoral, mPD-L1$^+$). Parece que la expresión de PD-L1 está asociado con el LBDCG tipo no-CG. Además, parece que la expresión de PD-L1 en las células del linfoma, pero no en otras células del microambiente del LBDCG tienen valor pronóstico, asociándose con un peor pronóstico los LBDCG PD-L1$^+$[190].

– *CXCR4 (CXC-chemokine receptor 4) o CD184* es un receptor ubicado en la superficie de las células madre y otras células inmunológicas, el cual es específico para un compuesto llamado Factor 1 Estromal-Derivado (SDF-1) o CXCL12. La vinculación del SDF-1 al CXCR4 conduce a la expresión de moléculas de adhesión. CXCL12/CXCR4 es crucial para la retención de los precusores B en la médula ósea, y para el *homing* de los linfocitos B a los ganglios linfáticos. El CXCR4 se ha implicado en la migración de las células B malignas en diversas neoplasias hematológicas. Se ha visto sobreexpresión de CXCR4 en tumores hematológicos. Según estudios recientes la expresión de CXCR4 parece aumentar la supervivencia de las células del LBDCG y su diseminación, disminuyendo la supervivencia de los pacientes con LBDCG, siendo factor pronóstico independiente en pacientes tratados con rituximab y quimioterapia.[191,192].

En resumen, el valor pronóstico de muchos de estos marcadores biológicos-moleculares es controvertido y deben de ser validados en series independientes bajo un es-

tándar de tratamiento actual, sin olvidar que los marcadores pronósticos, pueden cambiar con la introducción de nuevas terapias.

Estos marcadores biológicos deben ser además aplicables en la rutina (facilidad de obtención y conservación de muestra adecuada) y cumplir criterios técnicos de estandarización y reproducibilidad. Como ejemplo, la inmunohistoquímica y la morfología, a pesar de su bajo coste y alta disponibilidad, y de constituir el pilar del diagnóstico anatomopatológico, tienen en ocasiones pobre reproducibilidad y difícil cuantificación, debido a diferencias en el procesamiento de los tejidos, los protocolos de tinción, y la variabilidad del observador.

Los métodos moleculares como el perfil de expresión génica (GEP), la hibridación genómica comparada, los SNPs arrays, la secuenciación masiva, parecen técnicas prometedoras revelando la patogénesis del LBDCG, sin embargo no están siempre disponibles debido al coste o la complejidad.

En cuanto al tejido de la muestra, para el GEP y los SNPs arrays se precisa muestra en fresco o congelada, de la que no siempre se dispone, lo que en muchos casos suponía una limitación. En cuanto al GEP, actualmente parece que se puede establecer un pronóstico con muestras parafinadas (FFPE, fresh-frozen paraffin-embedded) con GEP[193], pero sigue ofreciendo dificultades técnicas. Respecto, a los CN *arrays*, se disponen de plataformas, como MIP (*molecular inversion probe*) *assays* para el estudio de alteración en el número de copias génicas en muestras con parafina.

1.1.5. Tratamiento del LBDCG

El LBDCG es el linfoma agresivo más frecuente de los LNH. Aunque la mayoría de los casos de LBDCG se curan con R-CHOP (rituximab, ciclofosfamida, doxorrubicina, vincristina y prednisona), aproximadamente del 30 al 40% presentarán refractariedad al tratamiento o recaída[147]. El papel de la intensificación con trasplante autólogo en la supervivencia de los pacientes con mal pronóstico no está claro[194–196].

Otros esquemas se han testado en diferentes ensayos clínicos, EPOCH-R (*etoposide, prednisone, vincristine, cyclophosphamide and doxorrubicin with rituximab)*, entre otros, podrían ser de interés en los pacientes *doble hit* o con LBDCG con pronóstico des-

favorable[197,198]. Sin embargo, es necesario realizar nuevos ensayos randomizados, con datos moleculares e inmunohistoquímicos para extraer conclusiones firmes.

Históricamente, los pacientes han sido considerados a tener un mayor riesgo de fallo al tratamiento basados en el IPI o el aa-IPI[146,150]. La aparición del GEP, parecer establecer diferente OS para los pacientes según la COO, presentando un peor pronóstico los pacientes con subtipo ABC[25]. Además esta clasificación del LBDCG establece subgrupos que tienen una biología característica, así como exclusiva sensibilidad a la quimioterapia, mediante la que podrían estratificarse pacientes para recibir un tratamiento concreto.

Entre los más recientes avances en la terapia dirigida por marcadores moleculares están los fármacos dirigidos a las siguientes vías: Ver Figura 3.

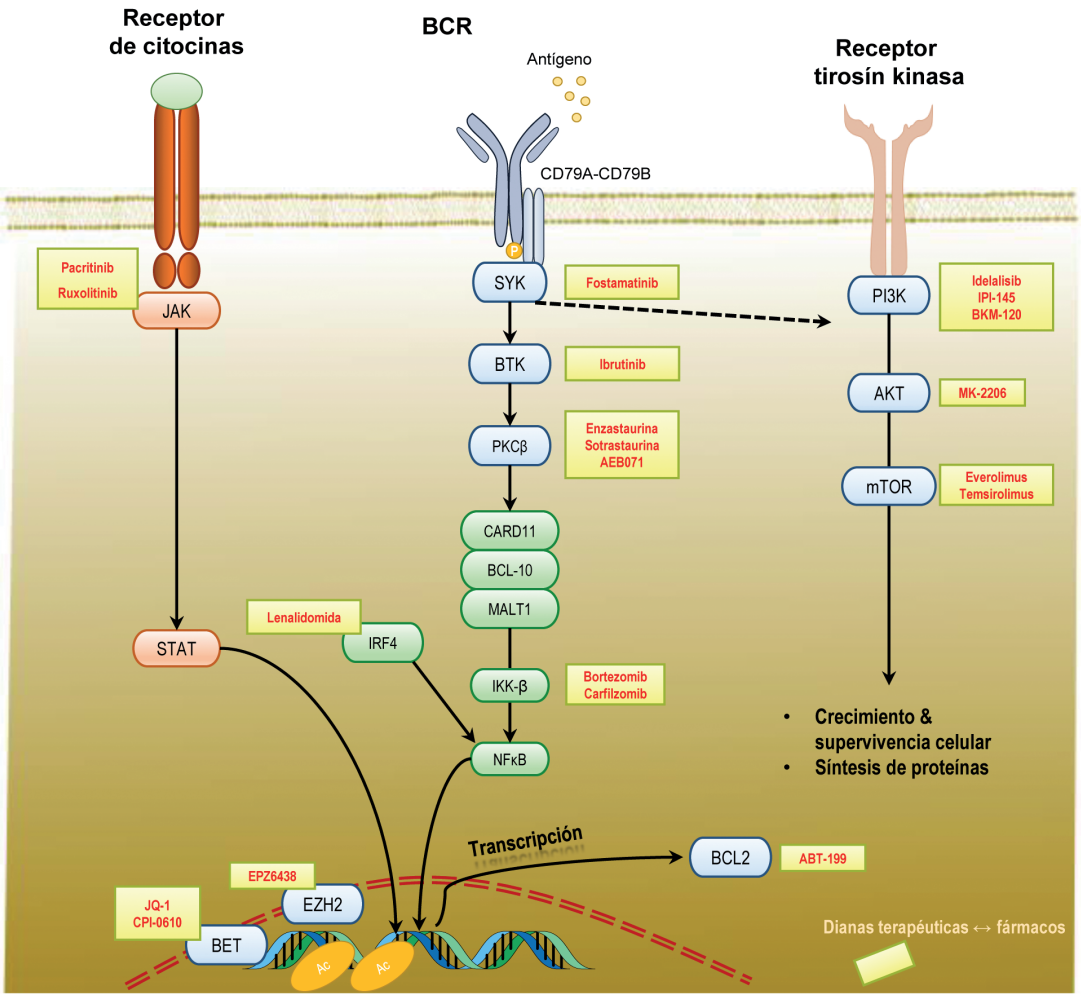


Figura 3. Vías claves de señalización en el LBDCG. Dianas terapéuticas. Adaptada de Mentha-Shah & Younes A. Sem Hematol 2015.

– Inhibición de la señalización del BCR: inhibidores de SYK (spleen tyrosine kinase), inhibidores de BTK (Bruton's Tyrosine Kinase), inhibidores de PKC-β.

– Vía de NF-κB: inhibidores del proteasoma, lenalidomida

– Inhibición de la vía PI3K/AKT/mTOR

– Inhibición de BCL2

– Inhibición JAK-STAT

– Nuevas dianas emergentes y estrategias terapeúticas: inhibidores de EZH2, inhibidores de BET (bromodomain and extra terminal family), inhibidores de PD1 (*Programmed Death*-1) y PDL1 (*Programmed Death Ligand* 1), células T con receptores antigénicos quiméricos (CAR, *chimeric antigen receptor*).

El LBDCG ABC, se caracteriza por activación constitutiva de NF-κB por diversas alteraciones genéticas. Mientras que en el subtipo CGB, BCL6 es una diana terapéutica ideal, dado que la supervivencia de las células en el LBDCG CGB depende en gran medida de dicha proteína. Sin embargo, la inhibición de BCL6 puede ser tóxica, por lo que se están diseñando nuevos inhibidores de dominios selectivos de BCL6, que parecen tener actividad in vitro en el LBDCG y que deberán ser testados en la clínica[199]. Además, se han testado moléculas inhibidoras de *EZH2* con datos preclínicos prometedores[81,82], por lo que se acaban de abrir ensayos clínicos con estas moléculas para el tratamiento de pacientes con LNH[99].

En el subtipo ABC, los inhibidores del proteasoma, como el bortezomib, parece que puedan ser eficaces en combinación con R-CHOP, aunque todavía hay estudios discrepantes[30,33]. Un potente inhibidor de BTK, ibrutinib, podría ser un potencial fármaco sobre todo en el subtipo ABC[200]. Así como inhibidores de MALT1[201], o IRAK4, y los inhibidores de PI3K (AL-101 o GS-1101)[202]. El agente inmunomodulador lenalidomida ha mostrado eficacia tanto como agente único como en combinación, y podría ayudar a vencer el efecto adverso del LBDCG ABC en primera línea[203,204].

En resumen, el IPI es un índice robusto, pero tras la introducción del rituximab, que aumentó claramente la supervivencia del LBDCG, y a pesar del intento de perfeccionar los índices clínico-biológicos, éstos han evolucionado poco desde su aparición en 1993, y además no logran identificar a aquellos pacientes que van a tener un pronós-

tico muy adverso. Por el contrario los avances moleculares en la última década han sido muy importantes. La traducción a la práctica clínica del COO del GEP a la IHQ tiene todavía problemas. Sin embargo, la COO, determina la vía de desarrollo de dianas terapéuticas, como BCL6 o EZH2 en el LDCG CGB, y la vía del BCR o NF-κB en el subtipo ABC. Además, la combinación de fármacos explorando diferentes vías, está aumentando el interés en este campo.

2.
EL RECEPTOR DE LA CÉLULA B. PAPEL EN LA ONTOGENIA DE LOS SÍNDROMES LINFOPROLIFERATIVOS Y EN EL LBDCG

Como se ha comentado previamente, se acepta la existencia de dos tipos principales de LBDCG NOS según la COO, el LBDCG CGB y al LBDCG ABC. El primero procede de un linfocito B que se ha quedado atrapado en el centro germinal, y conserva la capacidad de sufrir *ongoing mutation*. En cambio el subtipo ABC, procede de una célula postgerminal, que en principio ha perdido la capacidad de realizar SHM, aunque algunos autores han descrito variedad intraclonal en el receptor de la célula B lo que indicaría que persiste esa capacidad, aunque quizás en un menor grado.

Los linfocitos B maduros reconocen antígenos extraños mediante receptores de membrana, llamados receptores de la célula B (*BCR, B-cell receptor*). La diferenciación de los progenitores B en linfocitos B maduros tiene lugar en la médula ósea, reguladas por interacciones con el estroma medular[205,206]. Posteriormente, los linfocitos B viajan hasta los órganos linfoides secundarios (bazo, ganglios linfáticos y tejido linfoide asociado a mucosas) donde se produce la maduración dependiente de antígeno, principalmente en los centros germinales de los folículos linfoides[207,208]. La respuesta inmune que se produce en los órganos linfoides secundarios requiere de la presencia de macrófagos, células presentadoras de antígenos, linfocitos T y linfocitos B maduros[209,210]. En este ambiente se produce el contacto entre los linfocitos B y los antígenos, tras el cual aquellos linfocitos capaces de reconocer antígenos extraños proliferan y aumentan la afinidad de su BCR con el antígeno mediante el proceso de SHM[207,211]. Es también en este ambiente donde los linfocitos B pueden cambiar el isotipo de sus Igs con el fin de dispones de funciones efectoras diferentes. Tras la maduración dependiente de antígeno, los linfocitos B se diferencian a células de memoria o a células plasmáticas productoras de anticuerpos[212].

2.1. El Receptor de la célula B

Los linfocitos B presentan en su membrana celular receptores antigénicos de carácter proteico denominados Igs, que son responsables del reconocimiento antigénico

característico de la inmunidad específica. Aunque todas las células humanas tienen la dotación genética necesaria para su síntesis, sólo se expresan en los linfocitos B[213,214].

Hay 5 tipos de Igs (G, A, D, E y M), diferentes por su tamaño, carga, composición de aminoácidos y contenido de carbohidratos[215]. Las Igs estructuralmente se componen de dos cadenas polipeptídicas ligeras (IgL) idénticas y otras dos pesadas (IgH) también idénticas (de mayor peso molecular), unidas entre sí por puentes disulfuro[213]. Ambos tipos de cadenas están formadas por dos regiones distintas,- una constante (C) en la región carboxiterminal y una variable (V) en el extremo aminoterminal[216]. El tipo de Ig se define por la región constante de la cadena pesada (M, D, A_1, A_2, G_1, G_2, G_3, G_4 y E). Las regiones constantes de las IgH están formadas por 3 o 4 dominios globulares y en algunos casos contienen regiones bisagra, que proporcionan movilidad para adaptarse mejor al antígeno. Las regiones constantes de las cadenas ligeras (CL) también poseen pequeñas variaciones con respecto a los dos tipos de cadena: kappa (κ) y lambda (λ).

Funcionalmente, la región variable (Fab) está implicada en el reconocimiento antigénico mientras que la región constante (Fc) participa en las funciones efectoras de la célula, como las señales de transducción de membrana, y siendo responsables de la unión de las Igs a tejidos, a células del sistema inmune a proteínas del sistema de complemento[217].

2.1.1. Genes de las inmunoglobulinas: variabilidad genética

Cada una de las partes de la molécula de Ig, tanto la región constante como la variable, están codificadas por distintos genes. La región constante se codifica a partir de los segmentos génicos C, mientras que la región variable se obtiene de la combinación de varios segmentos génicos V (V*ariability*), D (D*iversity*, sólo en IgH) y J (J*oining*)[218].

La unión aleatoria de los diferentes segmentos génicos de la región variable V(D)J da una combinación específica y diferente para cada linfocito B. Esta recombinación es la responsable de la especificidad y variabilidad de cada receptor antigénico, y permite el ahorro de material genético. Así, la gran diversidad proteica se genera con unos pocos segmentos genéticos de una manera única y diferente en cada linfocito B, siendo un

marcador altamente específico, ya que la probabilidad de encontrar dos linfocitos con el mismo reordenamiento es prácticamente nula[218,219].

Hay 3 genes que codifican la síntesis de Igs, localizados en distintos *loci* cromosómicos: *IGH* en el cromosoma 14 (14q32) y para la cadena ligera κ en el cromosoma 2 (2p12) y l en el cromosoma 22 (22q11)[220,221]. El locus de IGH comprende una región de 1.2 MB en la que se definen entre 123 y 129 segmentos IGHV, 27 segmentos IGHD y 6 segmentos IGHJ[222–224], aunque no todos son funcionales, 38-46 de los segmentos IGHV son funcionales y se agrupan en 7 familias. Estos segmentos IGHV, IGHD e IGHJ preceden a nueve genes funcionales (μ, δ, $\gamma_{1\text{-}4}$, $\alpha_{1\text{-}2}$ y ε) y a 2 pseudogenes de la región constante (C) que codifican para cada uno de los isotipos de la cadena pesada. Además, todos los segmentos CH, excepto Cδ, tienen en 5' un segmento génico S (*switch*) necesario para el cambio de isotipo de cadena pesada que tiene lugar en la diferenciación linfoide[225,226].

El objetivo del proceso de diferenciación B es la síntesis de una Ig capaz de reconocer específicamente a un antígeno. Esta diversidad se consigue mediante una triple estrategia: 1) recombinación V-(D)-J, en los estadios más tempranos de la diferenciación B; 2) deleciones e inserciones aleatorias durante el proceso de recombinación que incrementan la variabilidad de la unión, 3) SHM en la fase de maduración de afinidad en la que se consigue incrementar aún más la variabilidad de las regiones V.

2.2. Diferenciación linfoide normal

Durante la diferenciación linfoide B ocurren una serie de procesos de recombinación en los genes de la Ig con un objetivo final, la formación de un reordenamiento capaz de codificar una molécula de Ig funcional que pueda reconocer antígenos extraños. Este reordenamiento está regulado por una serie de proteínas que aumentan o disminuyen en los diferentes estadios madurativos. Hay dos tipos de diferenciación de célula B según si se ha producido o no el contacto con el antígeno. Maduración temprana, previa al contacto con el antígeno, que ocurrirá desde los estadios más inmaduros de la célula B (estadio pro-B) hasta formación de un linfocito B competente (expresa en su superficie una única inmunoglobulina funcional). Y maduración tardía, existe contacto con el antígeno y el linfocito B competente aumenta su afinidad por dicho antígeno y se transformará bien en célula de memoria, o bien en una célula plasmática.

2.2.1. Maduración independiente del antígeno

En esta fase se produce la recombinación entre los diferentes segmentos genéticos de IGH e IGL. Este proceso ocurre en los estadios más precoces de la diferenciación linfoide B y determinará la variabilidad del repertorio primario de las Igs. Esta maduración tiene lugar en la médula ósea y termina cuando el linfocito B maduro la abandona para entrar en el torrente circulatorio hasta llegar a los órganos linfoides secundarios.

RECOMBINACIÓN SECUENCIAL DE LOS GENES DE LAS IGS DURANTE LA DIFERENCIACIÓN LINFOIDE B

La recombinación de los genes de las Igs comienza en el estadio pro-B con el reordenamiento de IGH, en que un segmento IGHD se une con un IGHJ[227] (Figura 5). En el estadio pre-B, un segmento IGHV se reordena con el D-J anteriormente constituido[228]. Puede ocurrir que este segundo paso no tenga lugar en uno o en los dos alelos, en cuyo caso el (los) reordenamiento(s) se denominan incompleto(s). Una vez que se ha producido un reordenamiento IGH funcional en unos de los dos alelos, el proceso de recombinación se detiene, la cadena IGH se une a φ-LC (pseudo cadenas ligeras) y a CD79, y el complejo se expresa en la superficie externa de la membrana celular. Si el reordenamiento de los genes IGHV-IGHD-IGHJ producido es funcional se inicia la producción de cadenas pesadas m aunque también pueden producirse cadenas δ por un fenómeno de *splicing* alternativo. Así, la detección de la cadena m citoplasmática se considera uno de los primeros marcadores de la célula B[229].

La ausencia de síntesis de cadenas ligeras impide el ensamblaje de la Ig completa, por lo que las cadenas pesadas μ se acumulan en el citoplasma. No obstante, algunas cadenas μ se unen a otras proteínas Vpreβ-λ5 que forman las pseudo-cadenas ligeras (φ-LC)[230]. Estas proteínas Vpreβ-λ5 tienen una homología significativa con los dominios variables y constantes de las IgL pero los genes que las codifican no sufren procesos de reordenamiento[231–233]. El complejo formado por las dos cadenas Igμ asociadas covalentemente con las pseudo-cadenas ligeras (φ-LC) se conoce como complejo pre-BCR[234]. Se cree que la llegada a la superficie celular del complejo pre-BCR permite a la célula recibir una señal desde el exterior que, por un lado, bloquea el proceso de reordenamiento de los genes de las cadenas pesadas del cromosoma homólogo (mecanismo de exclusión alélica), y por otro, activa el reordenamiento también jerárquico y secuencial de los genes

de las cadenas ligeras κ y λ. Si el primer reordenamiento IGHV-IGHJ no es funcional, no habrá cadenas μ ó δ y por tanto tampoco complejo pre-BCR. Sin dicho complejo, la célula no recibirá señales y no se bloqueará el proceso de reordenamiento, que continuará con el segundo alelo. Si éste ya es funcional, el complejo se expresará en superficie y se inducirá el reordenamiento de cadenas ligeras, continuándose la diferenciación. Sin embargo, si este segundo alelo tampoco es funcional, la célula no expresará nunca el receptor, no podrá recibir las señales que le hagan avanzar en la diferenciación y entrará en apoptosis[235].

Si el proceso anterior ha concluido con éxito y se ha activado el reordenamiento de los genes de las cadenas ligeras, éste tiene lugar de forma similar al de la cadena pesada. En primer lugar se reordenarían los genes κ. Si este reordenamiento es funcional, la célula comenzaría la transcripción de las cadenas ligeras κ que se ensamblan con las cadenas pesadas μ o δ citoplasmáticas formando una Ig completa (IgM-κ o IgD-κ) que ya puede expresarse en la superficie celular. La expresión en superficie de una Ig completa es característica de los últimos estadios de la diferenciación B antígeno independiente y bloquea la formación de nuevos reordenamientos en los genes de *IGL*. En caso de que el reordenamiento del primer alelo κ no resultara funcional, se procederá al reordenamiento del segundo, y si no se consigue la formación de una IgL funcional entonces se procede a reordenar el gen de la cadena l[227]. Este orden es el que justifica que en poblaciones linfoides B normales la relación κ/λ sea aproximadamente 2:1. Además, en caso de reordenamiento de los genes IGL, se produce la deleción del reordenamiento improductivo IGK[236–238]. De esta forma se evita la competencia entre la cadena aberrante Igκ y la Igλ funcional por la unión con la cadena IgH.

Si la cadena ligera es capaz de ensamblarse con la cadena IgM previamente formada, la célula entonces alcanza el estadio de LB inmaduro inmunocompetente Igμ+. Aquellos LB inmaduros que sean autorreactivos son eliminados por apoptosis o anergia, pero también pueden ser rescatados a través del proceso de edición del receptor[238–240]. Finalmente, la molécula de Ig sintetizada se une al heterodímero CD79 que permite su anclaje a la membrana y la transmisión de señales al interior de la célula. A este nivel de la diferenciación nos encontramos con un LB precoz inmaduro pero inmunocompetente, pues ya es capaz de reconocer AGs a través de su receptor antigénico. A partir de este momento el proceso madurativo estará orientado a aumentar la afinidad por el antígeno.

MECANISMO DE RECOMBINACIÓN IGHV-(IGHD)-IGHJ

La recombinación IGHV-(IGHD)-IGHJ está mediada por la expresión regulada de varias proteínas, entre las que se encuentran las codificadas por los genes activadores de la recombinación *RAG1* y *RAG2*, así como la proteín-kinasa dependiente de ADN (ADN-PK) y el complejo heterodimérico formado por las proteínas de unión al ADN Ku-70 y Ku-80[241,242]. El proceso de recombinación IGHV-(IGHD)-IGHJ se produce en 3 fases: primero se produce un corte en una de las cadenas del extremo 5'del heptámero, luego este corte se convierte en una estructura en horquilla en la parte codificante, quedando un extremo libre en la región no codificante[243]. Tanto el corte como la formación de la horquilla requieren de la secuencia señal de recombinación (RSS) y las proteínas RAG1 y RAG2. Los genes IGHV, IGHD, IGHJ están flanqueados por RSS, que son un grupo de enzimas conocidas en conjunto como VDJ recombinasa. Estas RSS se componen de 7 nucleótidos conservados (heptámero) localizados cerca del gen que codifica la secuencia seguidos por un espaciador (que contiene entre 12 y 23 nucleótidos no conservados) seguidos por un nonámero (9 nucleótidos). Las RSS se localizan en el lado 3'de una región V y en el lado 5' de la región J, que son los extremos implicados en el empalme. Sólo se recombinan eficientemente un par de RSS espaciadoras que no sean idénticas (por ejemplo, una con un espaciador de 12 nucleótidos se recombinará con otra con un espaciador de 23 nucleótidos). A esto se le conoce como regla 12/23 de la recombinación. Los extremos de las regiones no codificantes se unen formando las denominadas uniones señal. La apertura de la estructura en horquilla y la unión de los extremos codificantes están mediadas por las proteínas de unión al ADN Ku-70 y Ku-80 y la ADN-PK[244–246]. Durante la apertura y unión de los extremos codificantes, se produce la inserción al azar de nucleótidos no complementarios (N) por acción de la enzima TdT[247], mientras que la deleción de nucleótidos germinales de los extremos de los segmentos génicos que se están reordenando y las pequeñas adiciones de nucleótidos autocomplementarios (P) son producidos por nucleasas y polimerasas, respectivamente. Estas uniones "incorrectas" permiten aumentar la variabilidad de los genes de las Igs.

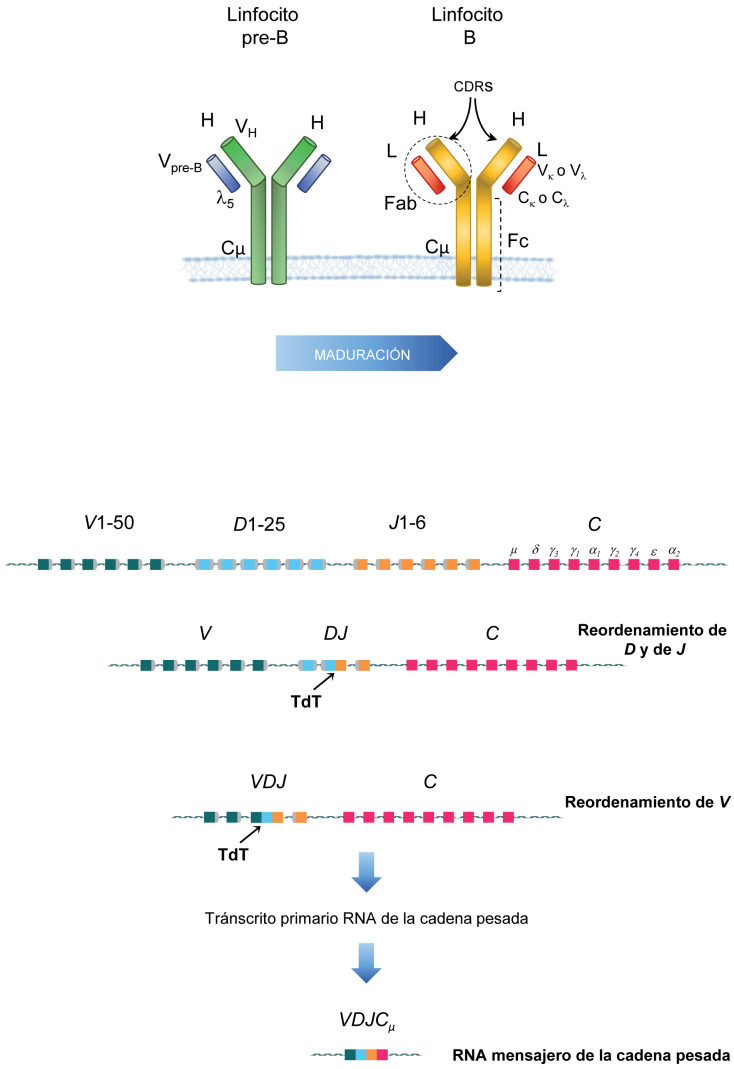


Figura 4. Esquema del proceso de reordenamiento y transcripción del gen de la cadena pesada de las inmunoglobulinas. Adaptado de Hematology: Basic Principles and Practice, Hoffman et al (6th edition).

ESTRUCTURA DE LA UNIÓN IGHV-IGHVJ

Dentro de la región variable de las Igs existen 3 zonas llamadas hipervariables flanqueadas por otras regiones relativamente conservadas (FR o *framework*). Las regiones hipervariables o CDR (*Complementarity Determining Region*) confieren a las Igs la afinidad específica por un determinado antígeno. De las 3 regiones, CDR1 y CDR2 se componen de secuencias de los genes VH mientras que CDR3, la más variable y la responsable de la especificidad de la Ig, está formada por la región de unión de los segmentos V-(D)-J.

EDICIÓN DEL RECEPTOR

Si los linfocitos B maduros que salen de la médula resultan ser reactivos frente a autoantígenos se transforman en "tolerantes", lo que significa que se produce un bloqueo en la cascada de transducción se señales mediadas por la sIg y en consecuencia la célula es incapaz de activarse[248]. Sin embargo, durante la diferenciación B las células B inmaduras pueden evitar el reconocimiento de autoantígenos alterando las regiones de unión al antígeno de sus Igs mediante un proceso de edición[249].

La edición del receptor antigénico puede ser llevada a cabo cambiando tanto las IgH como las IgL[238,250]. Este proceso de edición tiene lugar fundamentalmente en el locus κ mediante reordenamientos secundarios V-Jk aunque también puede producirse en los loci l y en IGH. La mayoría de los segmentos IGHV contienen en su extremo 5' una secuencia homóloga al heptámero de las RSS[251]. Los reemplazos pueden tener lugar por dos mecanismos: en el primero un segmento VH reemplaza al que se encuentra formando parte del reordenamiento V-D-J mediante el heptámero interno[215,251], en el segundo un segmento *IGHV* puede reordenarse al complejo IGHD-IGHJ preexistente en el alelo no funcional siguiendo el mismo esquema de los reordenamientos V-D-J normales[252]. Las regiones de unión de estos reordenamientos secundarios poseen las mismas inserciones y deleciones de nucleótidos que los reordenamientos V-D-J primarios. De hecho, la enzima TdT parece reactivarse durante este proceso de edición del receptor[253]. Dado el corto período de vida de las células B inmaduras, la capacidad de edición del receptor es limitada. Si no se logra una edición satisfactoria, las células mueren por un mecanismo de apoptosis.

2.2.2. Maduración dependiente del antígeno

HIPERMUTACIÓN SOMÁTICA (SHM, SOMATIC HYPERMUTATION)

La hipermutación somática (SHM) es un mecanismo genético que tiene lugar en los centros germinales tras la activación de las células B inducida por el antígeno. Se ha demostrado que es un proceso acoplado a la transcripción que se produce en una región de 1,5-2 Kb desde el promotor y que requiere tanto la presencia de un promotor como de potenciadores de la transcripción de los genes de las Igs[254–257]. Funcionalmente representa la base molecular de la maduración de afinidad de los linfocitos B naïve tras la

exposición al antígeno[258] pudiendo llegar a aumentar dicha afinidad entre 10 y 100 veces. Las mutaciones somáticas son principalmente mutaciones puntuales, aunque se han descrito también deleciones e inserciones[259] y pueden ocurrir en ambas cadenas de la doble hélice de ADN[260]. Normalmente los cambios afectan más a purinas que a pirimidinas y las transiciones (cambio de una base púrica o pirimidínica por otra base púrica o pirimidínica, respectivamente) son más frecuentes que las transversiones (cambio de una base púrica por una pirimidínica o viceversa)[261].

La mayoría de los SLP-B maduros suelen proceder de células que han pasado por el centro germinal (todos los MM, la mayoría de los LNH-B y la mitad de las LLC-B), debido a esto sus genes de las Igs suelen presentar mutaciones somáticas.

CAMBIO DE ISOTIPO

Tras la activación inducida por el antígeno, los linfocitos B proliferan y se diferencian produciendo diferentes tipos de Ig pero manteniendo la afinidad del BCR. El proceso se produce a través de otro reordenamiento en los genes IGH denominado switching o cambio de isotipo, en el que se produce recombinación entre las secuencias repetitivas localizadas en el extremo 5' de cada uno de los genes que codifican las regiones constantes de las cadenas pesadas de las Igs[262,263]. Al producirse la recombinación se delecionan algunas de las regiones CH de la configuración germinal, de manera que se sustituye la región CH más cercana a la región VH-JH por la región CH correspondiente[264]. En los estadios precoces, el isotipo IgH que expresan todos los linfocitos es Igμ o Igδ. Luego, ese mismo linfocito B puede expresar cualquiera de los otros isotipos o seguir expresando el isotipo Igμ según si se produce o no reordenamiento de esta región, sin que en ningún caso se produzcan cambios en la especificidad de la región de reconocimiento antigénico[265]. Lo cual implica que la mayor parte de los SLP-B maduros, que suelen haber pasado por el CG expresarán Ig de superficie distintas a IgM (excepto las LLC-B y la MW).

2.2.3. Detección de reordenamientos clonales

Todas las neoplasias linfoides proceden de la transformación de una célula linfoide que puede producirse en diferentes etapas del desarrollo normal de la célula B. En los

linfomas esta transformación suele ocurrir en una fase posterior a la estimulación antigénica de una célula B madura. En el proceso de linfomagénesis, a partir de este momento se genera una proliferación incontrolada de células que derivan todas de la antecesora común, fenómeno conocido como expansión clonal. Por ello todas estas células tienen el mismo patrón de reordenamiento de uno o varios de los genes de los receptores antigénicos, formando una población clonal, que predomina sobre el resto de linfocitos normales. Gracias a ello, el análisis molecular de los genes de los receptores antigénicos permite identificar patrones de reordenamiento comunes a muchas células, ayudando a establecer el diagnóstico de clonalidad en la mayoría de los SLPs.

TÉCNICAS MOLECULARES PARA LA DETECCIÓN DE CLONALIDAD

Los reordenamientos clonales de los genes de Ig, así como de los genes del receptor de la célula T (TCR, *T cell receptor*) pueden ser detectados por las técnicas de Southern blot (SB) y reacción en cadena de la polimerasa (PCR, *polymerase chain reaction*). En la actualidad el SB no se emplea, ya que es una técnica muy laboriosa y precisa cantidad y calidad de ADN importante y además es poco sensible (límite de detección 5-10%), por tanto la técnica de elección es la PCR porque es rápida, y requiere pequeñas cantidades de ADN (no necesariamente de alta calidad). Además en algunas técnicas de PCR, se emplean cebadores oligo específicos, son altamente sensibles, dependiendo el tamaño de la PCR de la posición relativa de los cevadores y del tamaño de la región de unión de los segmentos genéticos reordenados.

a. Detección de clonalidad y análisis de los productos de PCR

Las técnicas habituales de PCR para el estudio de clonalidad se basa en la amplificación de las regiones hipervariables (CRDs) de los genes de las Ig/TCR, mediante el empleo de cebadores consenso de las regiones V y J que flanquean dichas regiones hipervariables (Figura 6). El producto de PCR se analiza mediante electroforesis en agarosa y visualización con bromuro de etidio. Si existe una población clonal, el producto de PCR migrará como una banda única y definida con un tamaño concreto. Por el contrario, si no existe clonalidad o representa un porcentaje muy pequeño de la muestra (<1%), el producto de PCR se visualizará como múltiples bandas de escasa intensidad imposibles

de distinguir (*smear*). Esta técnica sencilla, que es la utilizada en la mayoría de los trabajos publicados, tiene sus limitaciones: existencia de falsos negativos debida a la hibridación incorrecta de los cebadores, bien por utilizar oligonucleótidos consenso (que no hibridan con todos los segmentos genéticos), o bien por la existencia de mutaciones en la zona de hibridación de los mismos. El principal problema es la discriminación entre policlonalidad y monoclonalidad, y por tanto los falsos positivos[266–268]. Para resolver el problema de los falsos positivos, es preciso realizar un análisis posterior (técnica heterodúples, geneScan) del producto de PCR por secuenciación. El empleo de diferentes estrategias como es el uso de cebadores frente a las 7 familias IGHV, ha incrementado el % de casos amplificables[269–271]. Así se obtienen productos de PCR clonales en el 65-85% de los SLP-B[272–274].

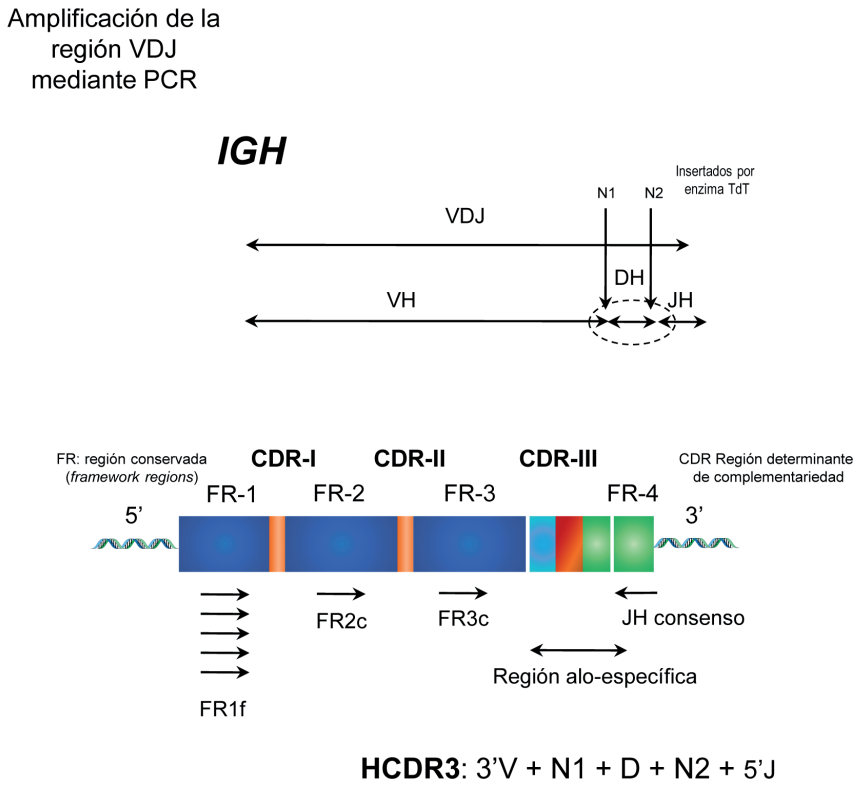


Figura 5. Amplificación de la región VDJ mediante PCR. Región de unión V-D-J: HCDR3.

Desde la aparición de los secuenciadores automáticos, se extendió el uso de la tecnología geneScan, que permite distinguir productos de PCR que se diferencian en al menos 1 pb. Tras la amplificación por PCR del reordenamiento, empleando un primer

fluorescente, se analiza el producto en un secuenciador automático equipado por un láser que detecta dicha fluorescencia. De este modo, un producto clonal aparecerá como una única banda de intensidad elevada, mientras que los productos policlonales se visualizarán como una distribución gaussiana de bandas[274,275].

Las ventajas del geneScan son mayor sensibilidad, la presencia de linfocitos policlonales no afecta tan drásticamente al diagnóstico de clonalidad como el caso de análisis de heterodúplex. Por otra parte, al incluir un marcador de peso molecular marcado con un fluorocromo diferente al de la PCR, el sistema nos permite conocer el tamaño exacto del producto de PCR, lo cual puede ser muy útil a la hora de evaluar diferentes muestras de un mismo paciente.

b. Detección de clonalidad por secuenciación masiva. Variación intraclonal.

Actualmente se están desarrollando métodos para la detección de clonalidad por NGS (*next generation sequencing*), metodología mucho más potente, pero que a su vez presenta la dificultad del análisis. En los últimos años también se ha descrito el concepto de variedad intraclonal. Así, la heterogeneidad intratumoral es importante en el origen, recaída, así como en el tratamiento del tumor. La célula de origen del tumor parece ser la misma "clon fundador", pero se han descrito subclones en la evolución basados en la aparición de mutaciones somáticas, y se ha observado que solo algunas de las mutaciones son importantes para la transformación o la refractariedad tumoral "*drivers*", mientras que otras son solo "*passengers*". Diferentes modelos evolutivos se han descrito[276,277]. Tanto para las distintas mutaciones del genoma como para la secuencia de las imunoglobulinas de cada linfoma, la NGS tiene todavía mucho para enseñarnos. Es posible que las mutaciones de IGHV pudieran variar en los diferentes subclones de un mismo linfoma y puede que la heterogenedidad genética pueda afectar a la evolución del tumor y a la progresión del mismo, sin embargo, estas hipótesis no han sido totalmente demostradas. Así, la AID se expresa tanto en las células del LBDCG CGB como en el subtipo ABC, además su expresión no se ha correlacionado en algunos estudios con el patrón de variedad intraclonal, hipotetizando que quizás otros factores hagan falta para la SHM en estos tumors[278]. En concreto en el LBDCG hay estudios que describen su heterogenicidad por NGS[84,279], pero muy pocos que analicen la evolución clonal[280]. En relación al

gen de las inmunoglobulinas, hay estudios que han tratado de analizar la clonalidad mediante NGS, especialmente del TCR[281–283], y hay algún estudio de VDJ, aunque escasos[283,284]. Sin embargo, ya se han desarrollado estrategias de análisis más robustas[285], por lo que hay diferentes estudios en marcha para analizar la historia natural evolutiva de estos tumores. En este sentido, recientemente se han publicados dos estudios, uno en leucemia linfoblástica aguda[286] y otro en linfoma folicular[287] que plantean diferentes mecanismos evolutivos y de selección clonal en cada enfermedad.

2.2.4. El gen de las inmunoglobulinas en el diagnóstico de los SLP-B

El gen de las inmunoglobulinas es en un marcador ideal para la detección de clonalidad y el diagnóstico de los síndromes linfoproliferativos, especialmente en aquellos casos en los que la anatomía patológica no es concluyente. En el LBDCG es útil para diferenciar aquellos casos en los que podría tratarse de una proliferación reactiva, en algunos casos de diagnóstico intermedio, o cuando el material obtenido no es suficiente. En la amplificación del reordenamiento IGH de los síndromes linfoproliferativos, la detección del mismo por PCR puede ser difícil por diversos motivos, como la calidad de la muestra (muestras de ADN fijado en formol e incluido en parafina, muestras de punción con aguja fina – PAAF), o la presencia de mutaciones somáticas (al ser un síndrome linfoproliferativo altamente mutado los cebadores diseñados pueden no hibridar y dar falsos negativos). Así, el reordenamiento incompleto D-J, aunque no presente en todos los casos, puede ser una alternativa factible en algunas muestras.

El reordenamiento de IGH, tanto completo como incompleto (en algunos casos), es también en una diana ideal para el seguimiento de la enfermedad mínima residual (EMR) en el caso de la célula tumoral, ya que la presencia de estos reordenamientos identifica específicamente a dicha célula y la distingue de las células normales, que carecen del reordenamiento IGH o tienen otro diferente. El empleo de cebadores que hibriden específicamente con la región HCDRIII, cebadores alelo específicos, permiten la detección específica de un reordenamiento clonal entre un millón de reordenamientos policlonales (técnica ASO-PCR). En el caso del LBDCG sería útil si hay infiltración de la médula ósea, y actualmente se está buscando también el reordenamiento con PCR específica en la sangre periférica.

2.3. Centro germinal: fisiología y oncogénesis de la célula B

Las células B que han sufrido el proceso de recombinación VDJ de forma exitosa migran como células B naive desde la médula ósea a los órganos linfoides secundarios. Tras el encuentro con el antígeno, las células naive se activan por su interacción con los linfocitos T CD4, en los tejidos linfoides y se agregan en folículos primarios para formar los Centros Germinales (GCs).

Los CGs son estructuras histológicas encargadas de la generación y de la selección de células B que producen anticuerpos de alta afinidad[288–290]. La función inmunológica fundamental del CG tiene un peligroso inconveniente, ya que los mismos mecanismos que permiten el desarrollo de los receptores o inmunoglobulinas de alta afinidad, son los mismos que están implicados en la transformación maligna de las células B[291,292].

La estructura del CG está compuesta por una zona oscura (*dark zone*), la que contiene casi exclusivamente células B en proliferación que sufren SHM (centroblastos), y una zona clara (*light zone*), en la que las células B son seleccionadas en base a su afinidad de su BCR por el antígeno y sufren el cambio de clase de isotipo CSR, *class-switch recombination*), para dar lugar a células con la capacidad de producir anticuerpos de alta afinidad y con diferentes isotipos (centrocitos). En esta zona clara los linfocitos B están entremezclados con linfocitos T, macrófagos, y células dendríticas. El resultado de la reacción de CG es la generación tanto de células plasmáticas como de células B de memoria.

El uso de las técnicas de imagen intravital ha contribuido a una mejor caracterización de las dos zonas, permitiendo la identificación de CXCR4 como un marcador funcionalmente muy importante para los linfocitos B de la zona oscura. Estas técnicas han demostrado también la polarización funcional de las dos zonas de los CGs[293–297]. Sin embargo, la maduración de las células B no siguen una vía unidireccional entre los dos compartimentos, sino más bien implica reentradas cíclicas en las que las células del CG, tras la selección en la zona clara regresan a la zona oscura para someterse a división celular y SHM adicionales. Estas vías parecen estar alteradas en los linfomas como resultado de alteraciones genéticas; muchas de ellas descritas por la secuenciación del genoma de linfomas humanos[9,60,83,84,298]. Así las vías de la reacción del CG están implicadas en la pa-

togénesis de la mayor parte de los linfomas. Además, juegan un papel importante los virus y el microambiente[299].

2.3.1. Iniciación del CG y formación de la zona oscura

Durante la reacción del CG, se producen las dos modificaciones del ADN de las células B que alteran el BCR: SHM y CSR, como se ha descrito anteriormente, ambas requieren de la enzima desaminasa de citidina inducida por activación, AID, *activation-induced cytidine deaminase*[300]. La CSR cambia la clase de la IGH de IgM a IgG, IgA o IgE, mientras la SHM implica a las mutaciones de IGHV, que crean una población de células B con aumento o disminución de la afinidad por un antígeno particular.

Además, una serie de factores de transcripción dan forma al fenotipo de las células B en los diferentes estadios del CG (Figura 7). Así, los factores que contribuyen a los estadios tempranos de la formación del CG son:

– ***MYC*** fue el primer gen asociado con los linfomas derivados del CG[301,302]. Pero su papel fisiológico no se conocía. MYC tiene un patrón de expresión bimodal durante la reacción del CG: aumenta de forma precoz en el inicio de la reacción del CG, es suprimida transcripcionalmente en la zona oscura y se reexpresa en un subgrupo de células B de la zona clara que están destinadas a reentrar en la zona oscura. Así, la formación de los CGs requiere la expresión de *MYC* durante los primeros días de desarrollo de CG, pero su expresión es silenciada de forma inmediata como resultado de la represión transcripcional de BCL-6 (*B cell lymphoma 6*)[63,303], que es el principal regulador de la reacción del CG. El aumento inicial de la expresión de *MYC* implica la inducción o amplificación[304] de los programas de transcripción que están asociados con muchas de las funciones de MYC, y que incluyen la estimulación del metabolismo, la función de la telomerasa y la replicación del ADN.

– ***BCL6*** se clonó a partir de un punto de ruptura de una translocación cromosómica en LBDCG[305]. La proteína BCL-6, es el principal regulador de la reacción del CG, se expresa específicamente por las células B del CG y se requiere para su formación. BCL-6 es un represor transcripcional que controla procesos que

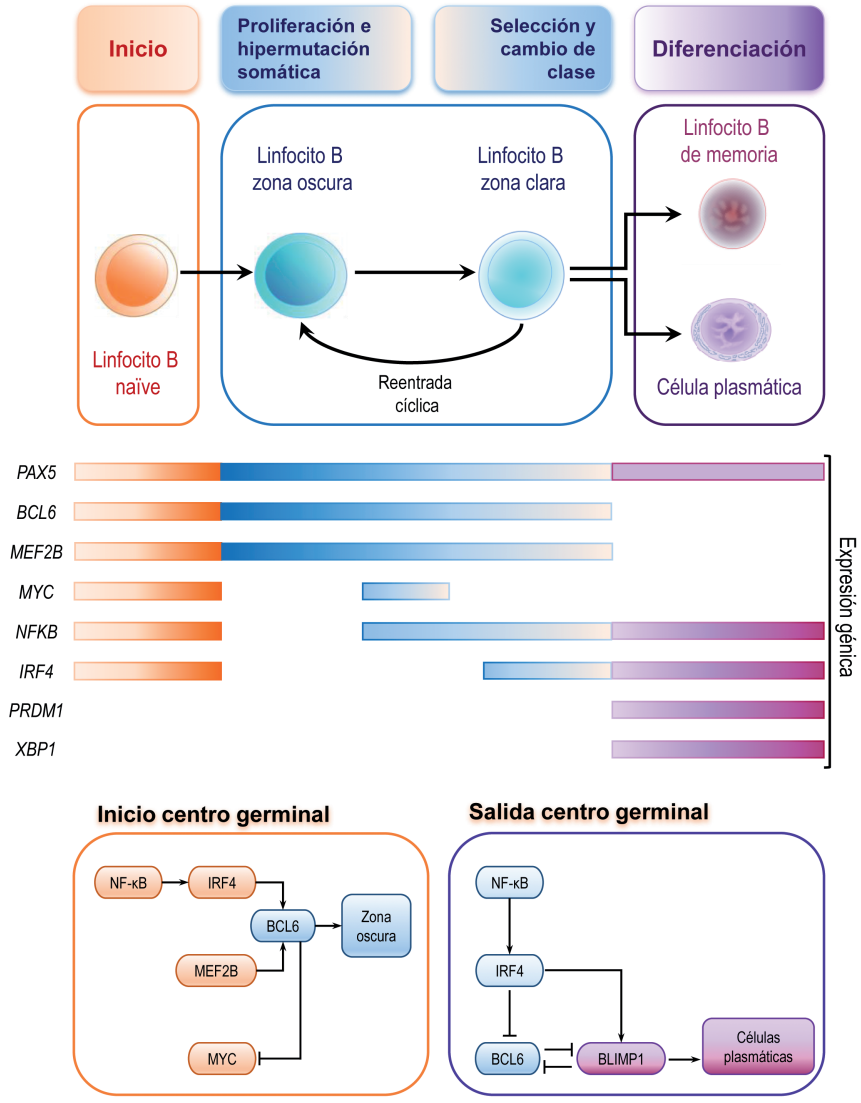


Figura 6. Redes transcripcionales que impulsan la reacción del CG. (Modificado de Basso K. & Dalla-Favera R. Nat Rev Immunol 2015) Múltiples factores transcripcionales están implicados en la iniciación de la reacción del centro germinal (CG). PAX5, paired box 5, se expresa a lo largo de la vida de las células B maduras- excepto en aquellas comprometidas a la diferenciación plasmátcia. Mientras que la expresión de BCL6 y MEF2B está restrigida al CG. MYC es necesario en la iniciación del CG y durante la reentrada en la zona oscura desde la zona clara, pero esta ausente en la mayoría de las células B del CG. NF-κβ e IRF4 se expresan durante el inicio del CG, están ausentes en la zona oscura y reaparecen en las células B de la zona clara. PRMD1 que codifica par BLIMP1 y XPB1 se inducen en las células que están comprometidas a la diferenciación a célula plasmática. Estos factores de transcripción están con frecuencia implicados en la regulación entre ellos. NF-κβ activa a IRF4, que contribuye a inducir a BCL6 durante el inicio del CG, pero promueve la regulación negativa de BCL6 y la expresión de PRMD1 a la salida del CG. BCL-6 regula negativamente la expresión de MYC y PRMD1.

llevan a la formación y al mantenimiento del CG[306]. El primer aumento de expresión de BCL-6 se detecta en la zona interfolicular en un pequeño subgrupo de células B naive, que se ha comprometido de forma exitosa por las interacciones con el antígeno y las células T[307,308]. IRF8 (*interferon-regulatory factor 8*), IRF4 (in *interferon-regulatory factor 4*)[309], y MEF2B[107] contribuyen a la in-

ducción de la expression de *BCL6*. NF-κB e IRF4 se expresan en el inicio de la reacción del CG, están ausentes en la zona oscura y reaparecen en las células B de la zona clara. NF-κB activa IRF4, y contribuye a la inducción de *BCL6* durante el inicio del CG, pero promueve la inhibición de la expresión de *BCL6*. Además, tras la inducción los niveles de expresión de BCL-6 hay un mecanismo de autorregulación por el cual BCL-6 se une a su propio promotor para regular negativamente su transcripción. BCL-6 tiene múltiples genes diana, que están implicados en promover y modular la reacción del CG. Estos genes están implicados en promover y modular la reacción del CG, principalmente en la represión de procesos celulares en la hipermutación de los centroblastos en la zona oscura: inhibiendo la parada del ciclo celular, modulando la apoptosis, haciendo que las células B del CG menos sensibles al daño del ADN, proporcionando un ambiente tolerante para que la remodelación fisiológica del ADN se produzca; e inhibiendo las vías implicadas en la diferenciación de las células B de memoria (para evitar la salida precoz del CG) y la diferenciación plasmática[310,311]. A pesar de que BCL-6 es un represor transcripcional, puede también promover la expresión génica de forma indirecta a través de la regulación negativa de los microRNAs[312]. El papel esencial de BCL-6 para una respuesta efectiva de la respuesta inmune fue demostrada en ratones deficientes de BCL-6, en los que la formación de CG se encontraba abolida[313,314]. Al contrario, la desregulación constitutiva de la expresión de BCL-6, conduce a la linfomagénesis en ratones y humanos[315].

– Otro gen que interviene en la iniciación del CG es el potenciador de la histona metiltransferasa, ***EZH2*** (enhancer of zeste homologue 2), que contribuye a la represión transcripcional. Su expresión en las células B maduras del CG y es necesario para el completo desarrollo del CG. Además está implicado en la linfomagénesis[80].

2.3.2. Zona clara, reentrada en la zona oscura y salida del CG

Después de la proliferación y la SHM en la zona oscura, las células B se desplazan a la zona clara, donde se dan los siguientes procesos: selección de células B que produ-

cen anticuerpos de alta afinidad, el proceso de CSR, y el inicio de la diferenciación de los centrocitos en células plasmáticas o células B de memoria.

Entre los genes importantes en la zona clara y especialmente en la reentrada a la zona oscura se encuentra *MYC*. La expresión de MYC es reprimida por BCL-6 en la zona oscura pero reactivada en un pequeño subgrupo de células B de la zona clara que están programadas para la reentrada en la zona oscura y someterse a adicionales ciclos de proliferación[63,303].

La activación de la vía de NF-κB es esencial para la selección de alta afinidad: las células B de la zona clara que llevan antígenos con una alta afinidad por el BCR, reciben señales adicionales de receptores, como CD40 y TLRs. Todos ellos activan diferentes vías (MAPK, PI3K), confluyendo en el complejo de transcripción NF-κB[316]. En las células del CG de la zona oscura la firma de NF-κB está ausente, pero esta se detecta en un subgrupo de células de la zona clara. En las células de la zona clara, NF-κB activa IRF4, que a su vez inhibe BCL6, favoreciendo así la diferenciación pos-CG[317]. La relevancia de este eje es fundamental para la fisiología del CG, la deleción de IRF4 conduce a la pérdida de diferenciación plasmática en ratones[318], y la activación constitutiva de NF-κB produce su acumulación[319].

Además de la señalización a través del BCR para el desarrollo de células plasmáticas, la regulación negativa de *BCL6* es esencial para la salida del CG de las células B y para su diferenciación. La inhibición de la expresión de BCL6, lleva a la represión de sus dianas, entre ellas, BLIMP1 (*B lymphocyte-induced maduration protein 1*), represor transcripcional codificado por PRMD1 (*positive-regulatory-domain containing 1*), permitiendo de este modo la diferenciación plasmática[320,321]. Múltiples vías coordinan la regulación negativa de *BCL6* y promueven su degradación[306]. La represión transcripcional de BCL6 mediada por IRF4 ocurre por la activación de NF-κB que se produce por la señalización del BCR, CD40 y TLR.

La inducción de BLIMP1 lleva a la diferenciación a célula plasmática: la producción de células plasmáticas requiere el cambio de "programa transcripcional de célula B", que depende de varios factores transcripcionales, incluyendo PAX5 (*paired box 5*)[322–324] y BCL6, a un "programa transcripcional de célula plasmática", que se establece por BLIMP1,

IRF4 y XBP1 (X-box binding protein)[325]. La expresión de IRF4 inducida por NF-κB promueve la diferenciación de la célula plasmática mediante la activación directa de *PRMD1*[326], así como por la liberación de la represión en el promotor de *PRMD1* mediada por BCL6[321]. A su vez, la expresión de BLIMP1, reprime "el programa transcripcional de la célula B" manteniendo la represión de BCL6. BLIMP1 es necesario para la diferenciación plasmática, ya que reprime la proliferación celular e implanta "el programa específico de la célula plasmática", que incluye la adquisición de un fenotipo secretor de anticuerpos que depende de la expresión de XBP1[327].

2.4. Linfomagénesis y origen de la célula tumoral del LBDCG

Las señales celulares descritas previamente y los efectores transcripcionales implicados en el control fisiológico del CG son diana de alteraciones genéticas que alteran su función, dando lugar a los linfomas derivados del CG. La mayoría de los LNH-B, incluyendo el linfoma de Burkitt, el linfoma folicular, y el LBDCG, derivan de las células del centro germinal, como se muestra por la presencia de genes de las inmunoglobulinas somáticamente mutados en sus genomas[328,329].

En base al GEP, el linfoma de Burkitt parece derivar de células B de la zona oscura del CG, mientras que el linfoma folicular y el LBDCG corresponden a células "paradas" en su diferenciación" por alteraciones que ocurren en el tránsito del CG[22,294].

En particular el linfoma folicular y el LBDCG tipo CGB se asemejan a las células de la zona clara, mientras que el subtipo ABC parece derivar de las células del CG detenidas durante estadios tempranos de la diferenciación de célula plasmática pos-CG (plasmablastos)[22,294]. El linfoma de célula B mediastínico representa un supertipo diferente que se origina de las células B tímicas pos CG en el mediastino[24].

2.4.1. Mecanismos de lesiones genéticas en el CG

De forma análoga a la mayoría de tumores, el LBDCG también alberga lesiones genéticas como deleciones, amplificaciones, mutaciones puntuales que conducen a la activación de oncogenes, o a la inactivación de genes supresores de tumores.

Las modificaciones del BCR (SHM y CSR) que requieren de la AID, son esenciales para la respuesta inmune normal, sin embargo, por su capacidad para introducir roturas en el ADN, exponen a las células del CG a un riesgo constante, que puede convertirse en patológico. Además, estas reacciones tienen lugar en un ambiente, el CG, donde las células B se replican rápidamente y los *checkpoints* de daño del ADN están silenciados como resultado de la actividad del represor transcripcional BCL6. Así, los LNH-B, y específicamente el LBDCG, tienen dos tipos de alteraciones genéticas adicionales, que son las translocaciones cromosómicas y las SHM aberrantes (ASHMS), que derivan de errores durante una de estas dos reacciones[291]. La prueba de este modelo, vino de la demostración de que la ablación de la AID, enzima requerida para SHM y CSR en modelos de ratones propensos a desarrollar linfoma podía prevenir la formación de reordenamientos *MYC-IGH* y el desarrollo de LBDCG[330,331].

ASHM es un mecanismo de inestabilidad genómica, cuya diana son las secuencias 5' en genes activamente trascritos, como resultado del funcionamiento erróneo del proceso fisiológico del proceso de SHM[332]. En las células del CG, la SHM está restringida a unos pocos genes, incluyendo los genes de las inmunoglobulinas y *BCL6*[110,332], porque, aunque la AID puede unirse a múltiples secuencias de ADN[333], las mutaciones en genes fuera de los genes diana son normalmente reparadas con alta precisión[334], lo que previene la difusión de la actividad mutacional. En cambio, casi la mitad de los casos de LBDCG, muestran múltiples mutaciones somáticas en un elevado número de genes que se transcriben activamente, incluyendo los protooncogenes *PIM1* y *MYC*[332,335]. Aunque no se conoce de forma completa el daño genético producido por este mecanismo, probablemente contribuye a la heterogeneidad del LBDCG. La heterogeneidad clínica y biológica es ampliamente conocida, sin embargo la introducción de los estudios de GEP, han permitido el reconocimiento de distintos subtipos, basados en el estadio de diferenciación de la COO dentro del CG.

2.4.2. Subtipos de LBDCG derivados de distintas fases de la diferenciación B

La identificación de los circuitos moleculares que regulan la transición entre la zona clara (LZ, *light zone*) y oscura (DZ, *dark zone*) es primordial para comprender la linfomagénesis porque estos programas transcripcionales pueden ser reconocidos en

cierta medida en los dos subtipos de LBDCG basados en su COO. Ver Figura 8. Ambos, el LBDCG CG y el ABC pareen más relacionados con las células de la zona clara[294]; sin embargo el LBDCG de CG se define por su elevada expresión de bcl6 y CD10, y la ausencia de marcadores de las fase pos-CG (IRF4 y BLIMP1), y por genes de las inmunoglobulinas altamente mutados con mutaciones "*ongoing*"[336]. A la inversa, el GEP del LBDCG ABC se asemeja al de las células B activadas, sugiriendo que su supuesta célula de origen normal ha recibido señales para regular negativamente el programa del CG y está preparada para la diferenciación B terminal.

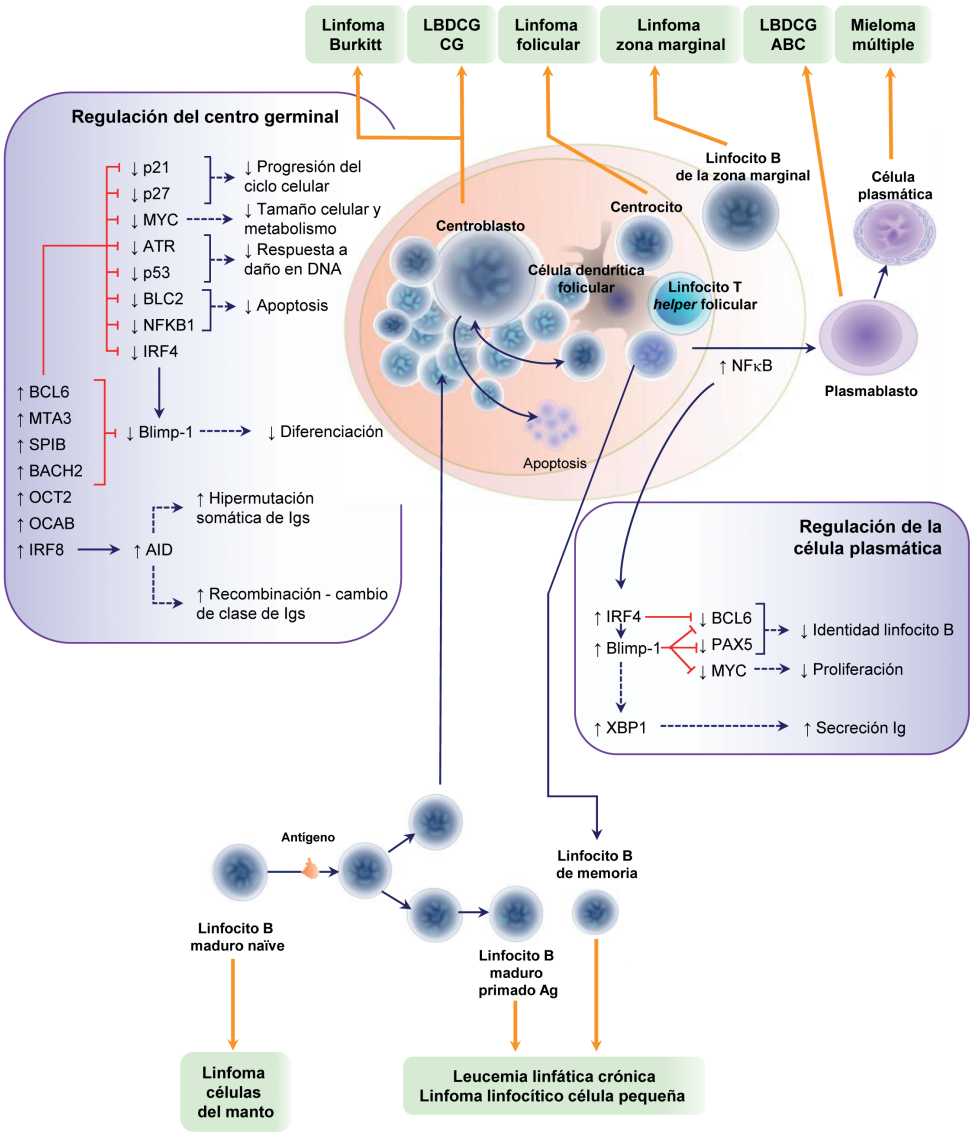


Figura 7. Diferenciación de la célula B y linfomagénesis. Modificada de Lenz G & Staudt L. N Engl J Med 2010. Representación esquemática de como los linfomas surgen apartir de diferentes estadios de la diferenciación normal de la célula B. Representados algunos mecanismos que intervienen en la fisiología del CG y de la diferenciación terminal, y cuya alteración puede dar lugar a la linfomagénesis.

Según esta teoría, la célula de origen del LBDCG ABC, no tendría posibilidad de mutaciones "*ongoing*" en los genes de las inmunoglobulinas[91,336], sin embargo, hay trabajos que muestran evidencia de mutaciones "*ongoing*" en este subtipo, proponiendo la hipótesis de un mecanismo de SHM adicional independiente del CG[337], aunque hasta la fecha esta última hipótesis no ha sido demostrada. Finalmente, hay estudios que sugieren CSR podría tener un papel en célula de origen del LBDCG, correlacionando el isotipo del BCR con el subtipo molecular, expresando es subtipo CG un isotipo secundario (IgG o IgA) y el subtipo ABC un isotipo primario (IgM o IgD), sugiriendo que la conservación del isotipo IgM podría ser necesario para la linfomagénesis en el subtipo ABC[338].

2.4.3. El gen de las Inmunoglobulinas en el LBDCG

El LBDCG, está por tanto constituido por células B que llevan un reordenamiento clonal de los genes de las IGs. El reordenamiento de IGH ocurre, como se ha visto previamente, durante la ontogenia de la célula B de una forma ordenada, llevando a la unión de distintos genes IGHV, IGHD y de unión IGHJ[219].

Una vez que un reordenamiento ha llegado a ser completamente funcional, la célula B naive pueda reaccionar contra antígenos en el CG de los órganos linfoides secundarios, experimentando una maduración de la afinidad por el proceso de la hipermutación somática y cambio de clase por recombinación.

Las características moleculares de los reordenamientos de IGH, así como las mutaciones somáticas de los mismos, pueden proporcionar información importante sobre la ontogenia de los linfocitos B. Por ello las características de los genes IGH se han analizado en profundidad. Tanto la SHM como el uso IGHV, ha sido analizado en las células B normales, en las enfermedades autoinmunes y en los síndromes linfoproliferativos. Algunos IGHV se han relacionado con algunos patógenos o se han vinculado con trastornos autoinmunes, o ciertos SLP-B. A pesar, de que el LBDCG es el LNH agresivo más frecuente, hay pocos estudios caracterizando el gen de las inmunoglobulinas, y la mayoría basados en poco pacientes o casos seleccionados[120,338–343]. Concretamente, se ha observado que el gen *IGHV4-34*, está sobrerrepresentado en el LBDCG con respecto a su uso en las células normales y en otros SLP-B[338,339,344–346]. Además se ha asociado la expresión de *IGHV4-34* con el subtipo no-CG[100,338]. El LBDCG es un SLP-B con SHM, ya que su cé-

lula de origen procede el CG. El uso de IGHV y el patrón de las mutaciones, nos pueden dar información sobre la patogénesis de este tipo de linfoma y el papel de la estimulación antigénica, no bien conocido hasta la fecha.

Se han identificado algunos patógenos como responsables de la linfomagénesis mediante la activación del BCR[347,348] o como eventos secundarios. La implicación de antígenos en la linfomagénesis podría apoyarse también en el estudio de los reordenamientos de las Igs. Importante, es el uso del gen *IGHV1-69*, que ha revelado una asociación entre la infección del virus de la hepatitis C (VHC) y la linfomagénesis[16,349] parece que el VHC infecta las células B mediante la unión de la proteína de cubierta con CD81, lo que lleva a una estimulación crónica. El EBV es otro patógeno bien conocido.

Así, la identificación de grupos estereotipados con una región VH CDR3 muy similar, apoyaría dicha hipótesis. Los BCR estereotipados se han estudiado profundamente en la LLC-B[350–352], y más recientemente en otros linfomas[352–356]. La existencia de BCR estereotipados en el LBDCG y su significación no se conoce todavía.

2.4.4. Concepto de Receptor Estereotipado: papel en los Síndromes Linfoproliferativos B (SLP-B) y específicamente en el LBDCG

Los criterios para definir secuencias estereotipadas del reordenamiento de IGH, aún no están bien establecidos[352]. Los criterios para la identificación de patrones estereotipados han sido previamente publicados y actualizados[352,357], e incluyen: 1) al menos, identidad del 50% de los aminoácidos; 2) como mínimo 70% de similitud entre secuencias; 3) IGHV perteneciente al mismo clan; 4) longitud idéntica de VH CDR3; y 5) localización exacta de la secuencia patrón en la región VH CDR3.

Los patrones estereotipados y la definición de subtipos han sido extensamente analizados en la LLC. Esto indicaría que ciertos antígenos comunes pueden dar lugar a la LLC, ya que el hecho de que los reordenamientos de LLC de diferentes individuos sean tan similares es prácticamente imposible por el proceso de reordenamiento y SHM. También como hemos comentado ha sido estudiado en el linfoma del manto[352,353], en el linfoma esplénico de la zona marginal[354], etc.

2.4.5. Concepto de N-Glicosilación y su papel en el LBDCG

Los sitios de potencial N-Glicosilación adquiridos se producen por la SHM en las regiones variables de la cadena pesada de la Ig (IGHV). El motivo para N-glicosilación es Asn-X-Ser/Thr, donde X es cualquier aminoácido excepto Pro, Asp o Glu. La N-Glicosilación se da en el 9% de las células B normales, y parece sestar relacionado con algunos SLP-B, como el LF (84%), LB endémico (82%) y LBDCG (41%), aparentemente asociado a t(14;18) y subtipo CGB. Estos datos sugieren que los sitios nuevos potenciales de N-Glicosilación, podrían estar relacionados con los SLP-B que derivan del CG. Además, el hallazgo de que estos motivos no son comunes en células B normales, somáticamente mutadas o en secuencias IGHV no funcionales indicaría que estos sitios son positivamente seleccionados.

Estos nuevos sitios hacen que interaccione con el microambiente del CG, y contribuye a la supervivencia de las células tumorales, quizás sin necesidad de un antígeno. Las inmunoglobulinas llevan oligosacáridos N-glicosilados localizados principalmente en las regiones constantes de la cadena pesada. Estos actúan como separadores para la molécula de inmunoglobulina manteniendo las funciones efectoras. La glicosilación de la región variable es menos común, pero se ha encontrado en las moléculas de algunos anticuerpos. Alguna secuencia germinal tiene un sitio potencial de N-glicosilación naturalmente, incluyendo VH1-08, VH4-34 y VH5a. Estudios de inmunoglobulinas monoclonales con actividad de anticuerpo han demostrado que la unión a antígeno puede estar aumentada o disminuida por la presencia de carbohidratos en la región V. Hay también evidencia de que la glicosilación de la región V puede alterar las propiedades y otras características atribuidas normalmente a la región Fc. Las secuencias no funcionales IGHV tienen pocos de estos sitios, lo que va a favor de selección positiva. Diversos estudios sugieren que el control sobre el sitio natural, es independiente de la adquisición de nuevos sitios.

Parece también ser importante el lugar donde se localizan estos nuevos sitios dentro de IGHV. Se piensa que estos cambios podrían contribuir a la linfomagénesis, independiente del antígeno, mediante la unión al BCR de manosas del microambiente del CG[358,359].

Así, la multitud de moléculas glicosiladas expresadas por las células B en el CG podrían argumentar contra un simple adhesivo papel de los oligosacáridos añadidos en la región IGHV de los SLP-B del CG.

3.
ESTRUCTURA DEL ADN Y POLIMORFISMOS GENÉTICOS

3.1. Organización del ADN

El ADN es la molécula que contiene toda la información genética del individuo. El conjunto de esta información presente en las células es el denominado genoma y, según su localización, podemos identificar un genoma nuclear, complejo, y un genoma mitocondrial, más simple. El genoma humano haploide está constituido por aproximadamente $3.3x10^9$ pares de bases (pb). Se pueden distinguir dos tipos de ADN, según la función biológica que desempeñen[360,361]

El ADN codificante del genoma humano, es solo una pequeña fracción, alrededor del 1.5% son exones, que codifican proteínas. Soporta gran presión selectiva, lo que se traduce en una variabilidad de regiones limitada. Mientras que más del 90% consiste en ADN no codificante de funciones diversas (como por ejemplo, promotores de genes o reguladores de transcripcion) y, en otros muchos casos, de función desconocida o sin función aparente. Este tipo de ADN, por ser altamente polimórfico, tiene un gran interés de cara a la identificacion de individuos[362].

3.2. Polimorfismos genéticos: polimorfismos de nucleótido único (SNPs, *Single Nucleotide Polymorphism*)

Dentro de la semejanza entre los individuos de una misma especie, estimada en un 99.9% del genoma humano, existen variaciones individuales en la secuencia del ADN, muchas, sin efecto alguno sobre la información genética, ya que afectan a secuencias no codificantes del ADN. Este 0,1% de variaciones de secuencias de ADN entre individuos es lo que se conoce como polimorfismos genéticos. Los polimorfismos son variaciones en la secuencia del ADN, que se detectan en al menos el 1% de los individuos de una población. En la práctica, para que un locus sea considerado polimórfico el alelo más común de dicho locus debe tener una frecuencia poblacional menor del 99% y, de

acuerdo con la ley de Hardy-Weinberg, al menos un 2% de la población debe ser heterocigota para ese locus. Desde hace casi un siglo se utilizan los polimorfismos genéticos para estudiar las variaciones entre individuos. Son caracteres estables que se transmiten por herencia mendeliana simple y constituyen una expresión de la diversidad genética entre individuos de la misma especie.

En el ADN codificante existe poca variabilidad individual, exceptuando la región HLA. Los polimorfismos pueden ir desde la modificación de una sola base hasta cambios en número y/o tamaño en la unidad de repetición. Podríamos denominarlos en función del tipo de cambio que se produce:

a. **Polimorfismos de secuencia.** Se producen por el cambio de uno o más nucleótidos en una secuencia del ADN, sin modificación de tamaño. Suelen ser poco polimórficos y son típicos del ADN codificante, como por ejemplo el sistema HLA.

b. **Polimorfismos de longitud.** Se producen por la inserción o deleción de uno o más nucleótidos. Este tipo es el más abundante en el ADN repetitivo, sobre todo en el ADN mini- y microsatélite.

Se puede hacer otra denominacion en base al número de alelos que presentan:

a. **Polimorfismos bialélicos.** Aquellos que se presentan únicamente con dos variantes posibles. Ejemplo de ellos son los polimorfismos de nucleótido simple (SNP, *Single Nucleotide Polymorphism*).

b. **Polimorfismos multialélicos.** Aquellos que presentan más de dos variantes para el mismo locus. Algunos ejemplos son el ADN repetitivo (minisatélites, microsatélites) o el sistema HLA. A continuación describiremos con mayor detalle los polimorfismos más variables y de mayor aplicación clínica en la actualidad.

3.3. Sistema de Histocompatibilidad

3.3.1. Generalidades

El sistema HLA (*Human Leukocyte Antigen*) constituye el complejo principal de histocompatibilidad humano cuya función es la presentación de antígenos. Está compuesto por un conjunto de proteínas denominadas antígenos de histocompatibilidad presentes en la membrana celular, que se caracteriza por su extraordinario grado de variabilidad de unos individuos a otros. Esta variabilidad depende de dos factores: 1) poligenia (diversos antígenos codificados por distintos genes) y 2) polimorfismo (cada uno de los genes presenta múltiples alelos). La combinación de ambos hace del sistema HLA el sistema genético más complejo del ser humano (Figura 9).

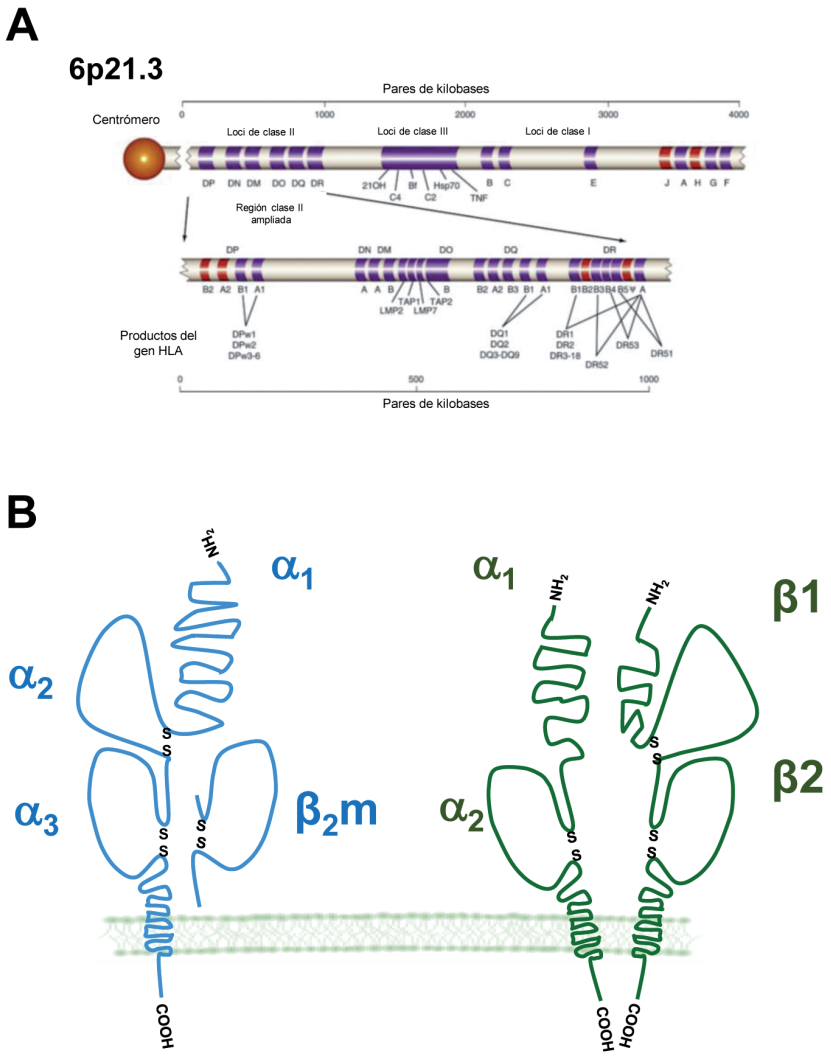


Figura 8. Mapa del complejo genético del HLA (A). Se ilustran los clusters de genes según la clase de productos codificados. El símbolo ψ representa los pseudogenes DRB, designados 7, 8 y 9. Esquema de las moléculas HLA de clase I y II (B) ilustrando su similitud estructural: dos dominios "inmunoglobulina" se sitúan próximos a la membrana celular (α3 y β2m para la clase I y α2 y β2 para la clase II). Los otros dos dominios se extienden hacia el medio extracelular con hélices α (α1 y α2 para la clase I y α1 y β1 para la clase II) y una plataforma de láminas β paralelas que forman un surco de unión a péptidos.

En general, el sistema genético que codifica a este tipo de antígenos, observado en casi todos los vertebrados, se conoce como complejo mayor de histocompatibilidad (MHC, de *major histocompatibility complex*). Las moléculas HLA juegan un papel central en la respuesta inmunitaria frente a antígenos extraños así como en la selección positiva y negativa de los linfocitos T inmaduros frente a péptidos propios (autoantígenos); los linfocitos T solo reconocen y pueden responder frente a antígenos peptídicos que se han generado en el interior de las células. Estos antígenos derivan de patógenos que se replican dentro de las células o de patógenos externos que luego las células internalizan. En todos los casos, y por mecanismos distintos, los patógenos sufren en el interior de la célula un proceso de degradación del que derivan fragmentos peptídicos. A medida que se van formando, estos péptidos extraños se combinan en el citoplasma con los antígenos de histocompatibilidad formando complejos que se transportan a la superficie celular para su presentación y reconocimiento por los linfocitos T, lo que constituye el punto de partida de la respuesta inmune.

3.3.2. Estructura, organización genética y función del complejo HLA

Todos los genes del sistema HLA se hallan situados en una región cercana al centrómero en el brazo corto del cromosoma 6 (6p21.31-6p21.32). La región abarca una distancia de unos 4.1 Mb y contiene más de 200 *loci*. Los *loci* están agrupados físicamente en tres regiones, llamadas clase I, clase II y clase III. Los *loci* de clase I se hallan más cerca de la zona telomérica que los de clase II. La región de clase I codifica genes de la cadena pesada de las moléculas clásicas de trasplante, HLA-A, -B, -C, los genes no clásicos HLA-E, -F, -G, MICA, MICB, un numero largo de pseudogenes de clase I y otros genes, algunos de los cuales tienen una función desconocida. La región de clase II contiene los genes que codifican las cadenas a y β de los cinco tipos de moléculas de clase II: HLA-DR, DQ, DP, DP y DM. En esta región hay también otros 4 genes que intervienen en el procesamiento de antígenos; PSMB8 y PSMB9 codifican componentes del proteasoma responsables de transformar proteínas en péptidos; TAP1 y TAP2 participan en el transporte de estos péptidos del citosol al retículo endoplasma tico para su unión a las moléculas de HLA y la formación del complejo proteico HLA-péptido. Finalmente, la región de clase III codifica varias moléculas estructural y funcionalmente distintas, incluyendo las fracciones del complemento C4, C2 y el factor Bf, la 21-hidroxilasa, el factor

de necrosis tumoral (TNFα) y la proteína de choque térmico Hsp70 (*Heat shock protein 70*). El gen del *locus* de la cadena asociada a la cadena β de clase I, la β-2-microglobulina (β2M), está en el cromosoma 15.

Este grupo de genes tiene herencia mendeliana como caracteres codominantes simples. La región genética de un cromosoma que contiene todos los genes HLA transmitidos en un mismo bloque se denomina haplotipo HLA. Los *loci* del HLA codifican proteínas heterodiméricas que se expresan en la superficie celular y juegan un papel crítico en el reconocimiento y en la respuesta inmune.

Las moléculas de clase I consisten en dos cadenas polipeptídicas, una larga de 346 aminoácidos (cadena pesada) y una corta (cadena ligera) de 99 aminoácidos (aa), la β-2-microglobulina. La cadena pesada consiste en 5 regiones principales o dominios: tres dominios extracelulares, designados como a1 (incluye el N'-terminal), α2 y α3; un dominio transmembrana donde la cadena polipeptídica pasa a través de la membrana de la célula; un dominio citoplasmático (donde está el C'-terminal) dentro del citoplasma celular. Los dominios externos α1 y α2 contienen dos segmentos de hélice alfa que forman dos cadenas con un surco entre ellas. Los péptidos (de un tamaño entre 8-19 aa) presentados por las moléculas de HLA de clase I son de origen endógeno y se une no covalentemente en el surco entre las dos cadenas alfa. La cadena ligera, constituida por la molécula β-2-microglobulina, se une mediante interacciones no covalentes al dominio α3. Este dominio contiene una región conservada de unión a la molécula CD8 del linfocito T.

Las moléculas de clase II, a diferencia de las de clase I, están constituidas por dos polipéptidos transmembrana, cada una con dos dominios extramembrana (α1 y α2, β1 y β2) unidos mediante enlaces no covalentes. Interactúan para formar un surco a la cara exterior (formado por los dominios α1 y β1) que, al igual que en las moléculas de clase I, es el lugar de unión del péptido, si bien estos son de mayor tamaño (12-24 aa).

Asimismo, el dominio β2 contiene una región conservada de unión a la molécula CD4 del linfocito T. Al contrario del origen endógeno de los péptidos unidos a moléculas de clase I, los fragmentos antigénicos (péptidos) unidos a moléculas de clase II derivan de antígenos de origen exógeno. Las moléculas extracelulares son internalizadas por endocitosis. Los endosomas se fusionan con lisosomas y su contenido es digerido parcialmente. Los fragmentos resultantes se colocan en moléculas de clase II y retornan a la superficie celular.

3.3.3. Presentación de antígenos

Las moléculas HLA de clase I se encuentran en casi todas, si no en todas, las células nucleadas. Por el contrario, las moléculas de clase II solo se expresan en ciertos tipos de células especializadas en el procesamiento y presentación extracelular de antígenos a los linfocitos T, tales como macrófagos y linfocitos B. Además, los fragmentos de antígenos mostrados por las moléculas de clase I se generan por macromoléculas sintetizadas dentro de la célula, mientras que las de clase II son moléculas adquiridas en el exterior de la célula.

Las moléculas HLA presentan fragmentos de antígenos a diferentes subpoblaciones de linfocitos T, la mayoría de los cuales pertenecen a uno de los dos subtipos CD4+ o CD8+. Ambos subtipos de linfocitos tienen un receptor de antígenos (TCR) que reconoce un epítopo del antígeno presentado. Además de la interacción de la molécula de HLA con el TCR, es necesaria la participación de las moléculas CD4 y CD8. Así, las moléculas CD8 de los linfocitos T CD8+ se unen a un sitio localizado únicamente en las moléculas de histocompatibilidad de clase I (dominio $\alpha 3$), mientras que las moléculas CD4 de los linfocitos T CD4+ se unen al dominio $\beta 2$, exclusivo de las moléculas de histocompatibilidad de clase II.

3.3.4. Nomenclatura

En la actualidad hay en uso dos formas de nomenclatura HLA. Serológica, establecida en 1975 y modificada en 1984, se basa en las especificidades (epítopos) de los productos de los genes HLA definidos por técnicas inmunológicas serológicas o celulares. O molecular, establecida en 1987 y basada en alelos definidos por secuenciación. Las técnicas serológicas y celulares son técnicas inmunológicas que detectan determinantes o epítopos de los productos genéticos del HLA. Cuando un epítopo se da en más de un producto genético HLA se definen grupos de alelos más que alelos individuales. Este fenómeno creó al principio un proceso denominado "*splitting*". Los *splits* en realidad son productos genéticos HLA concretos que parecen compartir epítopos comunes definidos serológicamente. Así, podemos definir al splitting como el descubrimiento de productos genéticos que previamente se creían eran únicos que en realidad son múltiples mediante el empleo de procedimientos moleculares genéticos o bioquímicos con mayor poder de discriminación. Este proceso todavía continua.

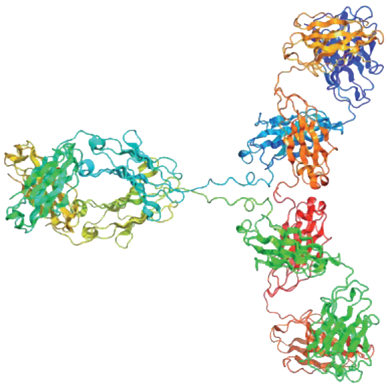
En la nomenclatura molecular los alelos específicos son designados mediante un asterisco seguido de un número de hasta 7 dígitos. Los dos primeros dígitos indican la especificidad serológica más estrechamente relacionada, si bien recientemente se utilizan para indicar la similitud estructural de nuevos alelos con los miembros de un grupo ya definido, incluso aunque estos puedan no ser serológicamente similares o la serología no esté clara. Los dígitos tres y cuatro definen la especificidad alélica. Por ejemplo, la especificidad serológica HLA-A1 en realidad comprende más de 45 alelos distintos, que se denominarían desde HLA-A*0101 hasta HLA-A*0145. Hay que destacar que estas variantes alélicas diversas, definidas por secuenciación, son análogas al proceso de "*splitting*" previamente descrito para las especificidades descritas por serología. Cuando se encuentra una variante alélica por alteración de la secuencia en una región nucleotídica no codificante los dígitos 02 se añaden a la designación, quedando la 01 como la normal y la 02 como la polimórfica. Por último, se pueden añadir una N o una L de forma opcional en algunos alelos, en los que la N indica que el producto no es expresado (N) o es expresado con un nivel muy bajo (L). Cuando se lleva a cabo un tipaje HLA molecular en ADN, la nomenclatura molecular debería asignarse de acuerdo con el nivel de discriminación del procedimiento empleado. Si el procedimiento no es capaz de discriminar entre variantes moleculares del HLA-A2, por ejemplo, lo apropiado es asignar al resultado la nomenclatura HLA-A*02. El estudio con los dos primeros dígitos se denomina baja resolución, ya que no llegan a discriminar entre alelos, pero permite definir entidades de similitud estructural. Los estudios de alta resolución permiten discriminar entre alelos, y se identifican con el tercer y cuarto dígito.

El sistema HLA es probablemente el sistema genético más polimórfico del ser humano, y prueba de ello son los más de 3.000 alelos descritos hasta la fecha. Si hubiera una distribución aleatoria de los alelos HLA, existirían unos 1023 genotipos únicos HLA-A, -B, -C, -DR y -DQ. Sin embargo, se sabe que los distintos alelos HLA se heredan en haplotipos, por lo que los alelos de diferentes *loci* no se asocian al azar. Así, se observa una asociación más frecuente de ciertos alelos con otros en comparación con la frecuencia que cabría esperar si la asociación fuera aleatoria en base a sus frecuencias alélicas, lo que se conoce como desequilibrio de unión. Este hecho permite la presencia de haplotipos más frecuentes. Además, la frecuencia de los distintos alelos varía entre poblaciones.

3.3.5. HLA y Linfoma

Diversos polimorfismos genéticos se han asociado con susceptibilidad o pronóstico en varios tipos de linfoma no hodgkin B (LNH-B)[363,364]. En los últimos años, estudios de asociación han identificado la región 6p21.3, como una región de riesgo para la susceptibilidad al desarrollo de diferentes tipos de linfomas, como el linfoma folicular[365–367] o el linfoma de hodgkin[368,369]. El sistema HLA localizado en esta región, juega un papel importante en la respuesta inmune antitumoral y en la apoptosis de las células tumorales[370], y así puede ser esencial para el control de la neoplasia. Estudios previos han mostrado relación entre los polimorfismos del HLA y la susceptibilidad para ciertas neoplasias hematológicas, como la leucemia linfática crónica, el mieloma múltiple y la leucemia linfoblástica aguda[369,371,372]. Sin embargo, hay poca información sobre la relación entre los polimorfismos del HLA y la susceptibilidad a desarrollar LNH-B o su evolución[373–377]. En el LBDCG se han descrito algunas asociaciones entre las especificidades del HLA y este subtipo de linfoma, así como con otros polimorfismos genéticos. Se ha observado una menor supervivencia libre de progresión, en pacientes con LBDCG con ausencia de HLA-DR2 o que llevan el polimorfismo TNF_{-308A}[374]. Además, la supervivencia global parece más corta en los pacientes con el haplotipo C*07-B*08-$LTA+_{252-}TNF_{-308A}$ o el HLA-C*07:01[377]. Sin embargo, estos estudios consideraron pacientes tratados antes de la introducción del rituximab. Como actualmente la mayoría de los pacientes reciben combinaciones de tratamiento con rituximab, la influencia de los polimorfismos del HLA en el pronóstico y susceptibilidad a padecer LBDCG no están claros.

Hipótesis de trabajo



El linfoma B difuso de célula grande (LBDCG) es un linfoma agresivo y representa el subtipo de Linfoma No Hodgkin más frecuente, constituyendo aproximadamente el 40% de los mismos. La mayoría son clasificados como LBDCG, NOS, según la OMS 2008.

A pesar de los múltiples estudios génicos publicados en el LBDCG, todavía no se conoce con exactitud ni la patogenia ni la biología de este linfoma. El diagnóstico molecular del LBDCG se basa en el origen clonal común de todas las células tumorales. El estudio exhaustivo del reordenamiento del gen de las inmunoglobulinas, investigado hasta ahora sólo en series cortas de pacientes con LBDCG, nos puede ayudar a comprender el origen y la patogénesis de este tipo de Linfoma, además de la posible influencia de ciertos antígenos en la selección de un clon tumoral.

Los estudios de expresión génica ("GEP: *gene expression profiling*") han permitido diferenciar el LBDCG, NOS, en dos subtipos según la célula de origen: LBDCG del centro germinal (CGB) y de célula B activada (ABC). El GEP, aunque publicado hace ya más de una década, debido a su coste y complejidad, no se ha implantado en la práctica rutinaria para diagnosticar y tratar a nuestros pacientes. Los métodos inmunohistoquímicos permiten clasificar el LBDCG, NOS, en CGB y non-CGB, y se han propuesto como sustitutos del GEP, sin embargo, la concordancia es variable y no está claro el valor pronóstico en la era del rituximab.

En los últimos años, con la implantación de las técnicas de secuenciación masiva NGS, se están descubriendo nuevas mutaciones génicas, que aparecen en un subtipo histológico concreto o en ambos. Pero todavía no se ha encontrado una mutación patognomónica del LBDCG, ni un patrón de mutación génico que pueda diferenciar ambos subtipos. Además, todos estos hallazgos, aunque valiosos, no han podido identificar a los pacientes con muy mal pronóstico. Los principales factores pronósticos siguen siendo clínicos, Índice Pronóstico Internacional –IPI–, pero no identifican a ese subgrupo de pacientes que va a ser completamente refractario a la terapia. Otro factor pronóstico que parece repetirse en las diferentes publicaciones, es la presencia de una translocación de *MYC* y *BCL2* y/o *BCL6*, simultáneamente, los llamados *double* o *triple hit* linfoma.

Además, los estudios de hibridación genómica comparada, CGH *arrays*, permiten identificar alteraciones cromosómicas, y más recientemente con los SNP (*single nucleotide polymorphism*) arrays o MIP (*molecular invesion probe*) *arrays*, se pueden detectar de-

leciones, ganancias o pérdidas de heterocigosidad en pequeñas regiones del genoma de hasta 100 kb o 4 MB, respectivamente. Estos análisis están aumentando el conocimiento sobre el LBDCG, describiendo alteraciones en el número de copias (CNA, *copy numer alterations*), que afectan a genes implicados en vías importantes de señalización, y que en ocasiones no pueden ser identificados por otra metodología. Estos estudios de CNA podrían tener un valor importante ya que no son tan costosos como la NGS, y ya se pueden realizar en parafina. Por último, el estudio de CNA tiene la ventaja respecto a la inmunohistoquímica de ser más objetivo, por ello nos planteamos su uso como una técnica complementaria que pueda ser empleada en la rutina.

Por otra parte, se sabe la importancia de la inmunidad y en concreto del Sistema HLA en el LBDCG, habiéndose descrito alteraciones en este sistema, que lleven a que las células tumorales puedan escapar a la vigilancia inmunológica.

Además, no se conoce con certeza todavía si realmente podrían existir factores genéticos, y medioambientales que puedan llevar a un aumento en el riego de desarrollar un LBDCG.

En este contexto, la hipótesis planteada en esta tesis doctoral es:

a. En el primer trabajo, la hipótesis planteada es, mediante el estudio exhaustivo del gen de la cadena pesada de las inmunoglobulinas, podemos aumentar el conocimiento sobre la patogenia, origen, y etiología del LBDCG. Asimismo, se pretenden identificar las diferentes características del receptor de la célula B (BCR) en los dos subtipos de LBDCG, NOS, CGB y non-CGB y la posible existencia de secuencias estereotipadas, que puedan implicar a antígenos en la patogenia del LBDCG.

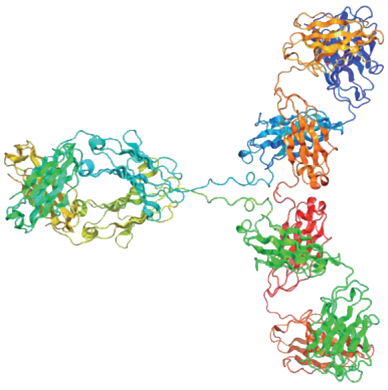
b. En el segundo trabajo, la hipótesis planteada fue describir alteraciones de CNA, que puedan clasificar, como sustituto o complemento de la inmunohistoquímica, el subtipo de LBDCG. Ello sería muy relevante dada la diferente reproductibilidad de los algoritmos inmunohistoquímicos existentes.

c. En este trabajo se planteó también identificar CNAs específicas que pudieran correlacionarse con el pronóstico del LBDCG para identificar pacientes con pronóstico adverso, no identificables con las escalas clínicas y clasificaciones bio-

lógicas vigentes. Ello sería importante para poder identificar aquellos pacientes que podrían beneficiarse de recibir tratamientos con nuevas moléculas en Ensayos Clínicos.

En el tercer trabajo, la hipótesis era identificar factores propios del huésped relacionados con el sistema HLA con impacto en el desarrollo y el pronóstico del LBDCG.

Objetivos



I. PRIMER TRABAJO: ESTUDIO DE LOS GENES DE LAS INMUNOGLOBULINAS EN LA CÉLULA TUMORAL DE MUESTRAS DE LBDCG

– Analizar las características de los reordenamientos completos e incompletos de la cadena pesada de las inmunoglobulinas en 165 pacientes con LBDCG, NOS, para profundizar en el conocimiento de la biología de esta enfermedad.

a. Caracterizar el receptor de la célula B en el LBDCG, NOS:

- Uso de segmentos génicos de las inmunoglobulinas.
- Patrón y distribuciones de las mutaciones somáticas.
- Definir la posible existencia de secuencias estereotipadas de las inmunoglobulinas en el LBDCG.
- Correlacionar la información obtenida con los subtipos inmunohistoquímicos de LBDCG
- Analizar las posibles implicaciones clínicas.

b. Analizar los reordenamientos incompletos D-J: frecuencia, uso de segmentos génicos y estado mutacional.

2. SEGUNDO TRABAJO: ANÁLISIS DE CNA EN MUESTRAS PAREADAS (NORMAL-TUMORAL) DE PACIENTES CON LBDCG

– Analizar las alteraciones en el número de copias (CNAs) y pérdida de heterocigosidad con número de copias normales (CNN-LOH) en muestras tumorales y no tumorales pareadas de 60 pacientes con LBDCG y sus relaciones clínico biológicas.

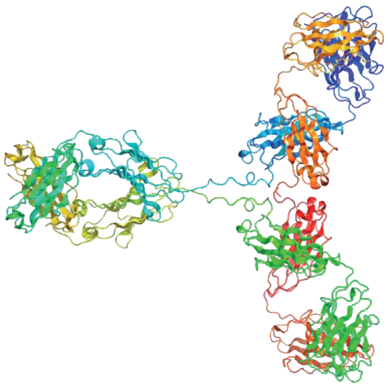
- Usar la información de CNA para identificar CNAs específicas que puedan predecir el subtipo de LBDCG según la célula de origen (CGB vs no-CGB), así como firmas pronósticas.
- Identificar pacientes de mal pronóstico, mediante el estudio de CNA, basándonos en CNAs específicas, para poder ofrecer a estos pacientes nuevas terapias dirigidas.
- Identificar alteraciones en las células no tumorales de pacientes con LBDCG, que pudieran tener un papel en el desarrollo de esta enfermedad.

3. TERCER TRABAJO: VALORAR LA IMPORTANCIA DE FACTORES GENÉTICOS DEL HUESPED (POLIMORFISMOS-HLA) EN EL DESARROLLO Y EVOLUCIÓN DE LA ENFERMEDAD (INCIDENCIA Y PRONÓSTICO)

– Analizar las especificadas antigénicas del Sistema Mayor de Histocompatibilidad (MHC) o Antígeno Leucocitario Humano (HLA) en nuestra serie de 250 pacientes con LBDCG.

– Correlacionar las especificidades con del HLA con el desarrollo y pronóstico del LBDCG.

Material, Métodos y Resultados



En esta sección de la memoria se describen los pacientes, material y métodos empleados, así como los resultados obtenidos en relación con cada uno de los objetivos planteados.

Para ello se incluyen los tres artículos científicos originales redactados y publicados como consecuencia del trabajo realizado.

Cada uno de los artículos referidos está precedido de un breve resumen en castellano que facilita una revisión rápida de la información contenida en los mismos.

1. PRIMER TRABAJO: ESTUDIO DE LOS GENES DE LAS INMUNOGLOBULINAS EN LA CÉLULA TUMORAL DE MUESTRAS DE LBDCG

Molecular characterization of immunoglobulin gene rearrangements in diffuse large B-cell lymphoma: antigen-driven origin and IGHV4-34 as a particular subgroup of the non-GCB subtype.

Sebastián E, Alcoceba M, Balanzategui A, Marín L, Montes-Moreno S, Flores T, González D, Sarasquete ME, Chillón MC, Puig N, Corral R, Pardal E, Martín A, González-Barca E, Caballero MD, San Miguel JF, García-Sanz R, González M.

Objetivo

El objetivo de este trabajo es analizar las características de los reordenamientos completos e incompletos de la cadena pesada de las inmunoglobulinas en 165 pacientes con linfoma B difuso de célula grande (LBDCG), *not otherwise specified*, NOS, para profundizar en el conocimiento de la biología de esta enfermedad.

Metodología

Pacientes

Para realizar este trabajo seleccionamos 165 pacientes diagnosticados de LBDCG, NOS, de acuerdo con la clasificación de la Organización Mundial de la Salud de 2008. Para diferenciar los subtipos biológicos, se utilizó el algoritmo de Hans.

Métodos

El DNA tumoral se obtuvo de las muestras en el momento del diagnóstico. En la mayoría de los casos (84%) las muestras procedían de adenopatías. Utilizamos tanto muestras de tejido fresco como muestras de tejido fijado en formol e incluido en parafina. Las muestras se testaron para la amplificación de los reordenamientos de la cadena pesada de las inmunoglobulinas (IGH) según los protocolos del proyecto BIOMED-2, en cuya estandarización nuestro grupo participó activamente. Después de la amplificación de los reordenamientos clonales completos e incompletos mediante Gene Scan, se procedió a la secuenciación de los mismos en un Secuenciador ABI3130 (Applied Biosystems).

Los genes germinales IGHV, IGHD e IGHJ, de los reordenamientos completos se identificaron usando la base IMGT/V-QUEST. Los reordenamientos incompletos se analizaron con BLAST. Se analizó el porcentaje de mutaciones somáticas, (*somatic hypermutation*, SHM) de IGHV, y se estudiaron sus características en base a los criterios de SHM canónicas. Además, se analizó la secuencia *VH CDR3* y la existencia de reordenamientos estereotipados según los criterios previamente publicados (*Darzentas N, Leukemia 2010* y *Agathangelidis A, Blood 2012*), para lo que utilizamos el software ClustalW/ClustalX 2.0. Las secuencias fueron comparadas con los patrones estereotipados previamente publicados (*Stamatopoulos K, Blood 2007* y *Agathangelidis A, Blood 2012*).

El análisis estadístico se realizó usando Graph-Pad y SPSS.

Resultados

Reordenamientos clonales

El porcentaje de detección de clonalidad, mediante amplificación de los reordenamientos completos V-D-J y/o incompletos D-J fue del 90%. La frecuencia de reordenamientos incompletos fue del 41%. El reordenamiento completo amplificó en 130 pacientes (79%), de los cuales 109 pudieron secuenciarse con éxito. Las familias IGHV3 e IGHV4 fueron las más predominantes, siendo *IGHV4-34*, *IGHV3-23* e *IGHV4-39* los más frecuentes. Con respecto a los reordenamientos incompletos, se observó un uso específico de IGHD, estando ausentes dos de los segmentos más usados en los reordenamientos completos. Esto sugiere que ciertos IGHD pueden favorecer el proceso de reordenamiento para obtener el reordenamiento completo funcional final.

Mutaciones somáticas

En el 91% de los reordenamientos completos se identificó la presencia de SHM. Todos los casos con el gen *IGHV4-34* estaban mutados. La mayoría de los reordenamientos incompletos (88%) fueron no mutados. El patrón de distribución de las mutaciones somáticas IGHV de nuestra serie de LBDCG cumplía los criterios de mutaciones canónicas: 1) un mayor índice de mutaciones reemplazantes en regiones CDR que FR,

2) mayor frecuencia de transiciones, 3) un porcentaje mayor de mutaciones de lo esperado en los hotspots RGYW.

Se encontraron cambios recurrentes en algunos aminoácidos, principalmente en los genes *IGHV4-34*, *IGHV4-39* e *IGHV1-18*. A pesar del alto porcentaje de mutaciones en los genes *IGHV4-34*, todos conservaban el motivo AVY y el W en la posición 7. Por el contrario, el sitio de N-Glicosilación de *IGHV4-34* estaba perdido en el 44% de los casos. Estos cambios podrían indicar cierto grado de selección en las mutaciones.

Secuencias CDR3

Tres secuencias de nuestra serie cumplieron los criterios de reordenamientos estereotipados. Dos de ellas, podían incluirse en 2 de los subgrupos estereotipados ya descritos en la LLC (grupos 5 y 7D), ambas usaban el gen *IGHV1-69* y no tenían mutaciones somáticas. El otro caso, podría formar un subgrupo provisional, con una secuencia pública de LBDCG (*Ruminy et al, Leukemia 2010*).

Correlaciones moleculares e inmunohistoquímicas

El uso y porcentaje mutacional de IGHV fue diferente en los 2 subtipos histológicos. Se observó un mayor porcentaje de SHM en el grupo CGB. Además, todas las secuencias no mutadas pertenecían a casos clasificados como no-CGB. El gen IGHV1-69 fue más frecuente en el subtipo no-CG. Y todos los casos con IGHV4-34, a pesar de estar altamente mutados, pertenecían al subtipo no-CGB.

Análisis comparativo de los reordenamientos IGH en el LBDCG con otros SLP-B y con células B normales

Para destacar las características específicas de los reordenamientos en el LBDCG, se compararon los resultados con las secuencias publicadas de otros SLP-B y de células B normales. El gen IGHV4-34 estaba más representado de forma significativa en el LBDCG que en células B normales, o en otros SLP-B como la LLC, la leucemia de células peludas, y especialmente con respecto al mieloma múltiple o la macroglobulinemia de Waldeström. Sin embargo, este gen se encontró en una proporción similar en el linfoma esplénico marginal o en el linfoma del manto. Por el contrario algunos genes, usados

frecuentemente en otros SLP-B y en células normales, como el gen IGHV3-30, no estaban presentes en nuestra serie.

Como se ha descrito previamente la mayoría de los casos fueron mutados, con un porcentaje de SHM similar al mieloma múltiple y mayor que la LLC.

Conclusiones

A pesar de que el LBDCG, es el linfoma no-Hodgkin más frecuente, el origen y biología de esta patología no se conoce con exactitud.

Mediante el estudio del receptor de la célula B, marcador de identidad del linfocito B, en nuestra serie de LBDCG, y con los resultados descritos previamente, podemos extraer las siguientes conclusiones:

– El estudio del BCR en el LBDCG, basándonos en el hallazgo de un uso "biased" de los segmentos génicos IGHV, mutaciones somáticas que cumplen los criterios canónicos, además de la presencia de mutaciones somáticas en aminoácidos recurrentes, sugiere que determinados antígenos podrían tener un papel en el desarrollo del LBDCG.

– Además, el estudio del BCR en los dos subtipos de LBDCG, indica que hay una diferente distribución de los segmentos génicos IGHV en los dos subtipos inmunohistoquímicos. Se confirmó en una serie más amplia que las publicadas previamente, la relación del gen *IGHV4-34* con el subgrupo no-CGB. Además se constató que estos linfomas presentan características biológicas particulares, pudiendo considerarse como un subgrupo dentro del subtipo no-CGB: presentando un alto porcentaje de mutaciones somáticas, pero conservando el motivo AVY y el aminoácido W en posición 7, lo que implica que estas células mantienen la capacidad de unirse a superantígenos.

– En este estudio, se describen por primera vez secuencias estereotipadas, en el LBDCG, aunque en frecuencia baja. Dos de ellas, perteneciendo a subgrupos previamente reportados en la LLC, afianzando la relación existente entre ambas patologías. La tercera secuencia estereotipada, que lleva el gen *IGHV4-34*, y comparte segmentos génicos y homología con una secuencia de LBDCG pre-

viamente publicada, podría constituir un nuevo subgrupo provisional. Estos hallazgos apoyan el papel del antígeno en el LBDCG.

– En cuanto al origen de las células del LBDCG no-CGB, se han propuesto varias hipótesis. Algunos autores afirman, que estas células no pueden sufrir mutaciones somáticas adicionales. Mientras que en otra publicación se describen mutaciones "ongoing" en estas células. En nuestro estudio aportamos información sobre las características de las mutaciones de este subtipo, que cumplen los criterios canónicos, lo que no apoya la necesidad de que exista una maquinaria diferente para el proceso de mutación somática, aunque no logramos resolver la duda planteada.

– Los reordenamientos de IGH del LBDCG tienen características específicas cuando los comparamos con otros SLP-B y con células normales.

Todos estos datos necesitan validarse en series prospectivas, más grandes e independientes. Nuestro estudio, abre futuras investigaciones de la biología del LBDCG basadas en el BCR.

The American Journal of Pathology, Vol. 181, No. 5, November 2012

http://dx.doi.org/10.1016/j.ajpath.2012.07.028

Tumorigenesis and Neoplastic Progression

Molecular Characterization of Immunoglobulin Gene Rearrangements in Diffuse Large B-Cell Lymphoma

Antigen-Driven Origin and IGHV4-34 *as a Particular Subgroup of the Non-GCB Subtype*

Elena Sebastián,[*†] Miguel Alcoceba,[*†] Ana Balanzategui,[*†] Luis Marín,[*†] Santiago Montes-Moreno,[‡] Teresa Flores,[§] David González,[¶] M. Eugenia Sarasquete,[*†] M. Carmen Chillón,[*†] Noemí Puig,[*†] Rocío Corral,[*†] Emilia Pardal,[‖] Alejandro Martín,[*†‖] Eva González-Barca,[‖] M. Dolores Caballero,[*†‖] Jesús F. San Miguel,[*†**] Ramón García-Sanz,[*†**] Marcos González[*†**]

From the Departments of Hematology and Pathology,[§] University Hospital of Salamanca, Salamanca, Spain; the Institute of Biomedical Research of Salamanca (IBSAL),[†] Salamanca, Spain; the Servicio de Anatomía Patológica,[‡] University Hospital of Marqués de Valdecilla/IFIMAV, Santander, Spain; the Institute of Cancer Research,[¶] Royal Cancer Hospital, London, United Kingdom; the Spanish Lymphoma/Autologous Bone Marrow Transplant Study Group,[‖] University Hospital Marqués de Valdecilla, Santander, Spain; and the Cancer Research Center-IBMCC (USAL-CSIC),[**] Salamanca, Spain*

The pathogenesis of diffuse large B-cell lymphoma (DLBCL) remains partially unknown. The analysis of the B-cell receptor of the malignant cells could contribute to a better understanding of the DLBCL biology. We studied the molecular features of the immunoglobulin heavy chain (IGH) rearrangements in 165 patients diagnosed with DLBCL not otherwise specified. Clonal IGH rearrangements were amplified according to the BIOMED-2 protocol and PCR products were sequenced directly. We also analyzed the criteria for stereotyped patterns in all complete IGHV-IGHD-IGHJ (V-D-J) sequences. Complete V-D-J rearrangements were identified in 130 of 165 patients. Most cases (89%) were highly mutated, but 12 sequences were truly unmutated or minimally mutated. Three genes, *IGHV4-34*, *IGHV3-23*, and *IGHV4-39*, accounted for one third of the whole cohort, including an overrepresentation of *IGHV4-34* (15.5% overall). Interestingly, all *IGHV4-34* rearrangements and all unmutated sequences belonged to the nongerminal center B-cell–like (non-GCB) subtype. Overall, we found three cases following the current criteria for stereotyped heavy chain VH CDR3 sequences, two of them belonging to subsets previously described in CLL. IGHV gene repertoire is remarkably biased, implying an antigen-driven origin in DLBCL. The particular features in the sequence of the immunoglobulins suggest the existence of particular subgroups within the non-GCB subtype. ***(Am J Pathol 2012, 181:1879–1888; http://dx.doi.org/10.1016/j.ajpath.2012.07.028)***

Diffuse large B-cell lymphoma (DLBCL) is a type of aggressive lymphoma that accounts for about 40% of B-cell non-Hodgkin lymphomas in adults,[1] most of them classified as DLBCL not otherwise specified (NOS) according to World Health Organization classification (2008).[1] Genome-wide expression profile technologies have allowed DLBCL NOS to be separated into two biological subtypes: germinal center B-cell–like (GCB) DLBCL and activated B-cell–like (ABC) DLBCL,[2] based on their presumed cell of origin. Immunostaining approaches allow these entities to be classified into GCB or non-GCB subtypes[3,4] and have been proposed as feasible surrogates of genome-wide expression profiles.

Supported in part by grants from the Health Council of Castilla y León (GRS265/A/08) and the Health Research Program (PS0901382 and CM10/00078-Río Hortega, ISCIII).

Accepted for publication July 30, 2012.

Supplemental material for this article can be found at *http://ajp.amjpathol.org* or at *http://dx.doi.org/10.1016/j.ajpath.2012.07.028*.

Address reprint requests to Marcos González, M.D., Ph.D., Servicio de Hematología, Hospital Universitario de Salamanca, Paseo de San Vicente, 58-182, 37007, Salamanca, Spain. E-mail: margondi@usal.es.

DLBCL is constituted by large B cells harboring a clonal rearrangement of immunoglobulin (IG) genes. The rearrangement of the immunoglobulin heavy chain (IGH) gene occurs during B-cell ontogeny in an ordered fashion, leading to the assembly of distinct variable (IGHV), diversity (IGHD), and joining (IGHJ) genes.[5] During B-cell differentiation, rearrangements start with an IGHD gene joining to an IGHJ gene, forming a partial IGH-D-IGHJ (D-J) rearrangement, a process that can simultaneously occur in both alleles. This is followed by the joining of an IGHV gene to the partial D-J rearrangement, forming a completely functional IGHV-IGHD-IGHJ (V-D-J) rearrangement in one allele, but can also retain a partial nonfunctional rearrangement in the other allele. Once a rearrangement has become completely functional, the naive B cell can react to antigens in the germinal center (GC) of the secondary lymphoid organs, undergoing affinity maturation by the processes of somatic hypermutation (SHM) and class switch recombination.

The molecular features of IGH rearrangements may provide important information about the ontogeny of B cells. Thus, the intrinsic mutability of IGHV genes has been widely analyzed.[6–8] Since normal B cells undergo SHM as a response to antigen, classically within GC, the pattern of SHM is thought to provide clues about the mutation machinery and the role of antigen selection. Nonrandom, potentially antigen-driven SHM has been described in normal B cells, as well as in chronic lymphocytic leukemia (CLL) and classic hairy cell leukemia (HCL), which has been called canonical SHM.[9–13] However, there are reports suggesting that certain B cells may accumulate IGHV gene mutations in an alternative T-cell–independent fashion.[14–16] The similarities or differences of mutations acquired through an alternative pathway are not yet defined. Several models[17–19] have been used for quantifying antigen selection, but results may not be reliably interpretable.[8] Since precise tools are still evolving,[19] canonical mutation criteria remain as an available method for establishing whether or not a set of sequences are affected by similar processes and selections.

Moreover, IGHV gene usage in IGH rearrangements has been analyzed in normal B cells,[20,21] in autoimmune diseases,[22] and in many B-cell lymphoproliferative disorders (B-LPDs).[13,23–30] Certain IGHV genes have been closely related to some pathogens[31] or linked to autoimmunity.[32] Particular antigens have been identified as possibly being responsible for lymphomagenesis or defined as a secondary event.[30] The involvement of particular antigens in lymphomagenesis may be supported by the identification of stereotyped groups with closely similar complementarity-determining region 3 (VH CDR3).[33] The stereotyped B-cell receptors (BCRs) have been studied extensively in CLL[25,34,35] with a clinical correlation[36] and, more recently, in other lymphomas.[27–30]

Although DLBCL is the most prevalent aggressive lymphoma, few studies have characterized the BCR in this entity, and most are based on small series of patients or in selected populations.[23,23,37–43] In addition, partial D-J rearrangements, the pattern and distribution of mutations, and the existence of stereotyped sequences in DLBCL are still not well defined.[41]

Here, we report a detailed immunogenetic analysis of the complete and partial IGH rearrangements in 165 DLBCL, the largest series to date, to improve our understanding of the biology of this disease.

Materials and Methods

Patient Characteristics and Immunohistochemistry

A total of 165 patients with DLBCL were selected retrospectively on the basis of a diagnosis of *de novo* DLBCL NOS according to the 2008 World Health Organization classification.[1] The median age was 59 years, with a range from 15 to 86 years. Fifty-one percent of the patients were male, and 48% of the total series had an International Prognostic Index score ≥3.[44] To distinguish DLBCL biological subtypes, Hans' algorithm[3] was used as a feasible surrogate of genome-wide expression profiles, since it has demonstrated a reproducibility beyond 80%.[3,4] Accordingly, immunostaining with antibodies to BCL-6, CD10, and melanoma-associated antigen (mutated) 1 (MUM1/IRF4) was performed in 121 patients, which allowed them to be separated into GCB DLBCL (38%) and non-GCB (62%). A total of 159 patients received treatment with curative intent, 125 of them (81%) with rituximab-based regimens (Table 1).

The study was approved by the local ethics review committee in accordance with Spanish legislation, and informed consent was obtained from all participants.

Table 1. Clinical-Biological Characteristics of DLBCL Patients at Diagnosis (n = 165)

Variable	% of Patients
Median follow-up, months (range)	27.8 (0–198)
Age [years, median (range)]	58.7 (15–86)
≤60	54
>60	46
Sex	
Male	51
Female	49
Performance status	
ECOG <2	70
ECOG ≥2	30
β2-microglobulin >3 mg/L	36
Albumin <3.5 g/dL	47
LDH elevated	59
IPI	
IPI 0–1	24
2	28
3	24
IPI 4–5	24
Immunostaining	
GCB	38
Non-GCB	62
Treatment	
Rituximab-based	81
Others	19

ECOG, Eastern Cooperative Oncology Group; GCB, germinal center B-cell–like; IPI, International Prognostic Index; LDH, lactate dehydrogenase.

DNA Extraction

Tumor DNA was extracted from samples collected at the time of diagnosis. These involved lymph nodes (84%), bone marrow (10%), and other tumor tissues (6%). In fresh samples, high-molecular-weight DNA was isolated using the DNAzol reagent (MRC, Cincinnati, Ohio).[45] DNA was extracted from formalin-fixed paraffin-embedded (FFPE) tissue in half of the lymph nodes using RecoverAll Total Nucleic Acid Isolation Kit (Ambion/Applied Biosystems, Foster City, California) or QuickExtract FFPE DNA Extraction Kit (Epicenter Biotechnologies, Madison, Wisconsin), with no differences in terms of recovery capacity.

Identification of Clonal Rearrangements

All samples were tested for the amplification of IGH rearrangements according to the BIOMED-2 Concerted Action protocols, where we actively participated in the standardization.[46] In this strategy, complete V-D-J rearrangement amplification was performed by multiplex PCR with a set of family-specific primers of the framework region 1 (FR1) and one IGHJ consensus primer. In cases using the *IGHV4-34* gene, a second PCR was performed with primers from the IGHV leader region. For the samples with no detectable amplification from FR1, PCR was performed from FR2. Additionally, partial D-J rearrangements were amplified in two different reactions using family-specific primers for the IGHD1-IGHD6 and IGHD7 families, respectively, together with the consensus IGHJ primer. The presence of the monoclonal rearrangement was then confirmed by GeneScan with an ABI 3130 DNA Sequencer (Applied Biosystems).

Sequencing and Identification of IGH Gene Rearrangements

PCR products were sequenced directly, including both forward and reverse reads, using Big-Dye terminators (Applied Biosystems).[46]

Germline IGHV, IGHD, and IGHJ genes from complete V-D-J rearrangements were identified using the IMGT/V-QUEST database (*http://www.imgt.org*, last accessed March 21, 2012). The following information was extracted: IGHV, IGHD, and IGHJ gene usage, the percentage of IGHV identity to the closest germline gene, and the VH CDR3 length and composition. The VH CDR3 length was determined by counting the number of amino acids from the cysteine at position 104 and the tryptophan at position 118, according to the IMGT unique numbering system.[47] Partial D-J rearrangements were analyzed using BLAST database.

Gene usage in DLBCL was compared with that previously described in normal B lymphocytes and in other B-LPDs for complete[13,20,24–28] and partial[26,48–50] rearrangements.

Analysis of Somatic Hypermutation

Percentage Identity with the Closest IGHV Germline

The use of 98% as the percentage identity cutoff value has customarily been used for the clinical consideration of a patient's sequence as mutated or unmutated based on the exclusion of potential genomic polymorphisms. This cutoff has a prognostic value in CLL.[51] However, it has been shown that most sequence differences in the 98% to 99.99% group correspond to low-level SHM,[52] as applied in recent studies.[27,53]

We, therefore, classified the sequences as previously described[27]: "truly unmutated" when no deviation from the germline was observed; "minimal/borderline mutated" when there was 97.0% to 99.9% similarity with the germline; and finally, "highly mutated" when the deviation from the germinal sequence was >97%.

SHM Characteristics

Nucleotide characteristics of IGHV mutations in DLBCL were studied to establish whether the pattern of distribution was consistent with the canonical SHM process.[12] This analysis was also performed separately in both the GCB and non-GCB subtypes, as well as in DLBCL expressing particular IGHV genes.

The defined characteristics of canonical SHM include: i) a higher ratio of replacement (R) versus silent (S) mutations in the CDRs than in the FRs; ii) base change bias for transitions; and iii) more than the expected percentage of mutations in the RGYW hotspots (R = A/G, Y = C/T, and W = A/T) relative to the total number of nucleotides located within these hotspots. In addition, to determine whether these characteristics were due to selection or to mutation machinery, the mutations at redundant wobble bases were analyzed as previously reported for CLL and HCL.[12,13] Wobble bases are the third nucleotide of redundant codons. Finally, mutations at A:T and G:C pairs were studied to determine whether the mutations were targeted with approximately the same frequency as in normal cells.[54] Moreover, recurrent nucleotide and amino acid changes for the most common IGHV genes were analyzed to find evidence of driven SHM in DLBCL.

VH CDR3 Analysis and Identification of Stereotyped Rearrangements

The criteria to define a stereotyped sequence are not yet well established.[35] To perform this analysis, we used the previously described pattern-based method,[55] updated with the recently published, more stringent criteria,[35] to identify possible restrictions in the VH CDR3 amino acid composition of rearrangements. Briefly, the criteria for stereotyped patterns included: i) at least 50% amino acid identity; ii) at least 70% similarity between sequences; iii) IGHV usage of the same clan; iv) identical VH CDR3 length; and v) exact location of the sequence pattern within the VH CDR3 region. Analyses were performed using the ClustalW/ClustalX 2.0 multiple sequence alignment software.

Subsets were named as previously published,[25] those composed of only two sequences being considered "provisional." A letter of the alphabet was assigned to the new subset. These were compared with previously published stereotyped patterns[25,35] and with DLBCL sequences from public databases (EMBL, *http://www.ebi.ac.uk/embl*, last accessed March 28, 2012; NCBI, *http://www.ncbi.nlm.nih.gov*, last accessed March 28, 2012).

Statistical Analysis

The groups of patients were compared by the χ^2 or Fisher exact tests using Graph-Pad 4.0 (GraphPad Software, San Diego, California) and SPSS 15.0 (SPSS, Chicago, Illinois).

Differences were considered to be statistically significant for values of $P < 0.05$.

Results

Detection of Clonal Rearrangements

The overall detection rate of monoclonality by amplifying V-D-J and/or D-J rearrangements was 90% (149 of 165). V-D-J rearrangements were detected in 130 of 165 patients (79%), and D-J rearrangements were seen in 68 of 165 patients (41%).

Complete IGHV-IGHD-IGHJ Rearrangements

A total of 109 of 130 complete V-D-J rearrangements were successfully sequenced. Reasons for not having a sequence in the remaining 21 cases were as follows: high polyclonal background ($n = 2$); double sequences, with no possible individual reading ($n = 1$); unspecific amplification, which means that a sequence could be read but the result did not fit with any possible rearrangement ($n = 2$); low fluorescence (intensity) of the peak rendering an amplification band too weak for accurate sequencing performance ($n = 11$); and partial sequence, enough to be identified as a V-D-J rearrangement but insufficient to be completely identified ($n = 5$).

A total of 103 were productive rearrangements, whereas 6 were nonproductive. Most sequences (82%) were obtained from FR1. Within the productive setting, IGHV gene repertoire analysis revealed that IGHV3 and IGHV4 were the predominant subgroups (Table 2). A total of 27 functional IGHV genes were identified, of which *IGHV4-34* (15.5%), *IGHV3-23* (9.7%), and *IGHV4-39* (8.7%) were the most frequent. Regarding IGHD genes, a total of 25 were identified, with a predominance of *IGHD3* (Table 2), especially *IGHD3-22* and *IGHD3-10*. With respect to IGHJ gene usage, most cases used the *IGHJ4* and *IGHJ6* genes.

Partial IGHD-IGHJ Rearrangements

Sequence analysis was possible in 64 partial D-J rearrangements. The IGHD2 subgroup genes were most frequent (Table 2). In comparison with complete rearrangements, IGHD4 was significantly overrepresented in partial rearrangements (18.8% versus 5.8%, $P < 0.05$), whereas IGHD3 was underrepresented (14.1% versus 35.9%, $P < 0.01$). Overall, 20 IGHD genes were identified, of which *IGHD2-2* was the most frequent. *IGHD3-22* and *IGHD3-3*, two of the most frequent IGHD genes in V-D-J rearrangements, were completely absent in partial rearrangements, suggesting that the use of certain IGHD genes may favor the rearrangement process to obtain a final complete functional rearrangement. Finally, the distribution pattern of the IGHJ genes was similar to the complete rearrangements (Table 2).

Table 2. Distribution of Families in Productive V-D-J and Partial D-J Rearrangements

Variable	V-D-J ($n = 103$) n (%)	D-J ($n = 64$) n (%)
IGHV		
1	17 (16.5)	
2	2 (1.9)	
3	45 (43.7)	
4	34 (33.0)	
5	4 (3.9)	
6	1 (1.0)	
7	0	
IGHD		
1	11 (10.7)	6 (9.4)
2	22 (21.4)	19 (29.7)
3	37 (35.9)	**9 (14.1)**†
4	6 (5.8)	**12 (18.8)***
5	10 (9.7)	5 (7.8)
6	15 (14.6)	12 (18.8)
7	2 (1.9)	1 (1.6)
IGHJ		
1	4 (3.9)	3 (4.9)
2	5 (4.8)	5 (8.2)
3	13 (12.6)	8 (13.1)
4	46 (44.7)	32 (52.5)
5	13 (12.6)	3 (4.9)
6	22 (21.4)	10 (16.4)

Significant differences between productive V-D-J and incomplete D-J rearrangements in DLBCL are depicted in bold.
*$P < 0.05$.
†$P < 0.01$.

Somatic Hypermutation

Percentage Identity to the Closest IGHV Germline

Somatic hypermutation (<98% of identity to the closest germline IGHV gene) was present in 94 of 103 (91%) patients who had productive complete V-D-J rearrangements (median: 10.8%; range: 0% to 26%). As expected, only 9% of the sequences were unmutated, similar to previously reported results.[23,39,42,56] Overall, point mutations predominated, whereas only five cases had deletions ($n = 3$) or insertions ($n = 2$) in their sequences.

According to the criteria for SHM described previously,[27] 4, 8, and 91 cases were "truly unmutated," "borderline mutated," and "highly mutated," respectively. All truly unmutated and borderline mutated sequences were obtained from FR1.

Interestingly, all cases with the *IGHV4-34* gene but one (15 of 16) were "highly mutated." However, *IGHV1-69*,

Table 3. Characteristics of SHM in DLBCL (Global Series) and with Respect to Immunohistochemistry and the Most Common IGHV Families and Genes

	Number of sequences	Number of mutations	R/S ratio			Transition/transversion ratio			% nucleotides in RGYW	% of mutations in RGYW
			Total	CDR	FR	Total	CDR	FR		
DLBCL	96	2485	1.69	2.56	1.42	1.161	1.13	1.18	14.0	24.9
GCB	31	931	1.71	2.77	1.42	1.155	0.97	1.25	13.3	22.9
Non-GCB	46	1037	1.56	2.05	1.39	1.216	1.30	1.18	14.2	25.2
IGHV1	15	631	1.42	**1.45**	**1.42**	1.246	1.29	1.24	17.9	26.7
IGHV3	42	1127	1.81	2.71	1.49	1.159	1.16	1.16	12.0	19.3
IGHV4	32	662	1.70	2.98	1.27	1.115	1.07	1.14	15.9	31.6
IGHV1-69	4	46	0.92	**0.88**	**0.94**	1.300	1.50	1.21	19.8	32.6
IGHV3-23	10	243	1.60	2.24	1.31	1.505	1.47	1.53	13.1	26.3
IGHV4-34	15	304	1.34	1.91	1.22	1.338	1.39	1.32	14.8	30.6

Figures in bold indicate different R/S ratios from that expected.
CDR, complementarity-determining regions; FR, framework region; R, replacement mutation; S, silent mutation; RGYW, hotspots.

IGHV5-51, and *IGHV3-21* were the most frequent genes used in "truly unmutated" and "borderline mutated" cases.

As expected from the physiopathological point of view, mutations were observed in a minority of cases (12%) of partial D-J rearrangements with a low mutation frequency (median: 4.3%, range: 3.5% to 11.1%).

Nucleotide Characteristics of Somatic Hypermutation

S mutations occurred randomly, with a relatively even distribution, whereas R mutations were mainly localized in the CDR1 and CDR2 regions, at the 5′ and 3′ ends of FR2 and the 3′ end of FR3 (see Supplemental Figure S1 at *http://ajp.amjpathol.org*).

Analysis of canonical SHM was possible in 96 DLBCL sequences with <100% identity with the germline. For the 2485 mutations analyzed, the R/S mutation ratio was 1.72, being greater in the CDR (2.56) than in the FR (1.42) regions ($P < 0.01$) (Table 3). The transition-to-transversion ratio was 1.16, significantly higher than the expected ratio of 0.5 if the frequencies of the types of mutation were random ($P < 0.01$), and there was no significant difference between the CDR and FR regions. Finally, 24.9% of the DLBCL mutations were located in RGYW hotspots, a value significantly greater than the expected percentage (14.0%, $P < 0.01$). Thus, SHM in DLBCL displayed all of the established characteristics of canonical SHM. These canonical characteristics persisted in the most common IGHV genes, except for *IGHV1* (mainly *IGHV1-69*), which met these three criteria to a lesser extent (Table 3).

The bias for transitions over transversions was also reflected by the wobble base mutations (1.20, $P < 0.01$) (Figure 1), implying that this bias in DLBCL is due to the same mechanism of SHM.

Finally, G:C pairs were preferentially targeted for SHM relative to A:T (ratio: 1.86) (Figure 1). This observation was especially evident in the *IGHV4-34* gene (ratio: 2.23).

Recurrent Amino Acid Changes

Recurrent amino acid changes were mainly found among *IGHV4-34*, *IGHV4-39*, and *IGHV1-18* rearrangements. For example, a substitution at the FR3 codon 92 occurred in 13 of 16 (81%) cases of *IGHV4-34*; the substitution of S to T (6 of 16, 38%) and S to N (4 of 16, 25%) were the most frequent changes observed. This recurrent amino acid change contrasts with the adjacent codon, 93, which did not undergo any substitution in any of the 16 sequences. Although the *IGHV4-34* genes were highly mutated, all 10 available *IGHV4-34* sequences conserved the specific AVY motif at the end of FR1 codons 24 to 26, and the W at position 7 was intact in the 5

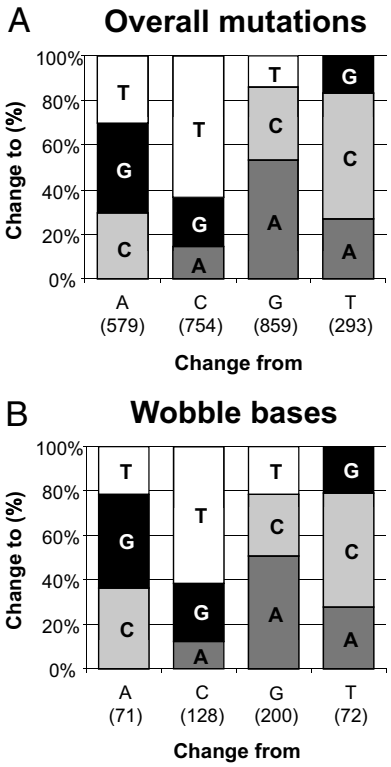


Figure 1. Nucleotides mutated in DLBCL. Relative frequency with which each nucleotide mutates to form each of the other three. The number of mutations occurring in each nucleotide is shown in parentheses. **A:** Total mutations in DLBCL. **B:** Mutations in redundant wobble bases.

Table 4. Stereotyped VH CDR3 Amino Acid Sequences

Subset	IGHV	IGHD	IGHJ	Mutational status	VH CDR3 amino acid sequences			
Subset 5	Published							
11–3436	**IGHV1-69∗13**	**IGHD3-10∗01**	**IGHJ6∗03**	**100**	**AR**	**TMVRGVI**	**NID**	**YYYYYMDV**
SE-01-0444-H1	IGHV1-69∗01	IGHD3-10∗01	IGHJ6∗03	99.6	AR	GMVRGVI	NID	YYYYYMDV
UK-01-0345-H1	IGHV1-69∗01	IGHD3-10∗01	IGHJ6∗03	100	AR	TMVRGVI	RVM	YYYYYMDV
FR-01-0286-H1	IGHV1-69∗01	IGHD3-10∗01	IGHJ6∗03	100	AR	VMVRGVI	SLD	YYYYYMDV
GR-02-0510-H1	IGHV1-69∗01	IGHD3-10∗01	IGHJ6∗03	100	AR	SMVQGVI	NAY	YYYYYMDV
NL-01-0294-H1	IGHV1-69∗01	IGHD3-10∗01	IGHJ6∗03	100	AR	TMVQGVI	RVY	YYYYYMDV
Subset 7D	Published							
11–1949	**IGHV1-69∗01**	**IGHD3-3∗01**	**IGHJ6∗02**	**100**	**AS PVGEG**	**DDFWSGYY**	**PN**	**YYDYGMDV**
CZ-01-0372-H1	IGHV1-69∗01	IGHD3-3∗01	IGHJ6∗02	100	AS SLGEN	YDFWSGYY	PN	YYYYGMDV
FR-01-0143-H1	IGHV1-69∗01	IGHD3-3∗01	IGHJ6∗02	100	AR PVQAK	YDFWSGYY	PN	YYYYGMDV
SE-01-1016-H1	IGHV1-69∗01	IGHD3-3∗01	IGHJ6∗02	99.6	AR PGGLY	YDFWSGYY	PN	YYYYGMDV
CZ-01-0484-H1	IGHV1-69∗01	IGHD3-3∗01	IGHJ6∗02	100	AR QGTLG	YDFWSGYY	PN	YYYYGMDV
UK-01-0427-H1	IGHV1-69∗13	IGHD3-3∗01	IGHJ6∗02	100	AR SGRED	YDFWSGYY	PN	YYYYGMDV
Subset A	Novel							
11–1967	**IGHV4-34∗01**	**IGHD3-22∗01**	**IGHJ2∗01**	**93.8**	**AR AL**	**DYYDNSG**	**RLP**	**GYFDL**
HQ392881.1	IGHV4-34∗01	IGHD3-22∗01	IGHJ2∗01	95.7	AR AL	DYYDSSG	RVLA	YFDL

Stereotyped rearrangements are compared with representative sequences from each subset. A complete list of sequences from subsets 5 and 7D is provided in Supplemental Table S2 at *http://ajp.amjpathol.org*. Sequences derived from the present study are remarked in bold.

IGHV4-34 cases that could be analyzed from the leader region. By contrast, the N-glycosylation site in CDR2 of *IGHV4-34* was lost in 7 of 16 (44%), and only one of the sequences that lost its natural site generated two novel sites in FR3. A summary list of recurrent amino acid substitutions for the predominant IGHV genes is given in Supplemental Table S1 (available at *http://ajp.amjpathol.org*). Some of these changes may indicate specific driven SHM in DLBCL.

VH CDR3 Composition and Stereotyped VH CDR3 Sequences

The median VH CDR3 length was 15 amino acids (range: 7 to 26). VH CDR3 was longer in the IGHV4 family than in the other families (16.8 versus 14.8, $P = 0.01$), and in "truly unmutated" and borderline cases as compared with "highly mutated" cases (18.2 versus 15.1, $P = 0.01$). This was particularly pronounced among the "truly unmutated" cases (21.5 versus 15.1, $P < 0.01$).

Three sequences among the 103 patients in our series met the criteria for stereotyped rearrangements and could be included into different subsets. Two sequences clustered within previously described CLL subsets (subsets 5 and 7D);[25,35] both contained the *IGHV1-69* gene and were "truly unmutated." The other case formed a probably realistic novel provisional subset (Table 4): subset A, sharing the V-D-J IGHV4-34-IGHD3-22-IGHJ2 conformation and with 79% identity of the amino acid sequence with a public DLBCL sequence.[42] A complete list of the sequences published in the literature from subsets 5 and 7D is shown in Supplemental Table S2 (available at *http://ajp.amjpathol.org*).

Molecular and Immunohistochemical Correlations

IGHV usage differed between the immunohistochemical subtypes. *IGHV1* (10 of 14) and *IGHV4* (15 of 21) were more frequent in the non-GCB subtype, especially for *IGHV1-69* (3 of 4) and *IGHV4-34* (12 of 12). A higher percentage of SHM was observed in GCB (13.8%, range: 2.9% to 26.0%) than in non-GCB (10.1%, range: 0.4% to 23.6%, $P < 0.01$) cases. Moreover, 6 of 7 borderline or unmutated cases with available immunostaining were non-GCB, which could explain this difference. By contrast, all *IGHV4-34* cases, despite being highly mutated, belonged to the non-GCB subtype. With respect to other molecular features, VH CDR3 length and the canonical characteristics for SHM showed no differences between the GCB and non-GCB subtypes (Table 3).

Comparative Analysis of IGH Rearrangements in DLBCLs with Other Normal B Cells and B-LPDs

To highlight the singular characteristics of DLBCLs, we compared our results with those reported for normal B cells and other B-LPDs. As shown in Table 5, the *IGHV4-34* gene was significantly overrepresented in DLBCL relative to $CD5^-/IgM^+$ normal B cells (3% to 9%, $P < 0.01$) and to other B-LPDs, such as CLL (8.7%, $P < 0.05$), HCL (6.9%, $P = 0.07$), and especially with multiple myeloma and Waldeström macroglobulinemia (2.7% and 0%, respectively, $P < 0.01$). However, it was found in a similar proportion of patients with splenic marginal zone lymphoma and mantle cell lymphoma.[13,16,20,21,24–26,57] By contrast, some genes frequently used in the rearrangements of normal B cells and other B-LPDs,[13,24–28] such as *IGHV3-30*, were completely absent from the present series ($P < 0.01$). Most complete V-D-J rearrangements were mutated, with a similar percentage of SHM (11.8%) to those previously reported in multiple myeloma (9%), but higher values than in CLL and Waldeström macroglobulinemia (2.4% to 6.6%).[24–26] Regarding SHM, G:C pairs were more often targeted for SHM than A:T pairs (ratio: 1.86) (Figure 1). This differs significantly from the distribution observed in CLL (ratio:

Table 5. Comparison of the IGHV, IGHD, and IGHJ Families and Significantly Different IGHV Genes in Our Series of DLBCL with Those Reported for B-LPDs and Normal B Cells in the Complete V-D-J Setting

Variable	DLBCL (n = 103)	MCL (n = 807)	B-CLL (n = 7596)	SMZL (n = 133)	HCL (n = 102)	MM (n = 270)	WM (n = 58)	$CD5^-/IgM^+$ (n = 206)
IGHV								
1	16.5	15.5	23.8	30.1	14.7	15.6	6.9	13.1
2	1.9	1.5	3.4	0.8	2.9	6.3	0	1.9
3	43.7	51.6	48.2	49.6	53.9	48.9	75.9	53.9
4	33.0	25.8	20.6	17.3	23.5	20.4	13.8	24.8
5	3.9	5.1	2.5	2.3	1.0	7.8	1.7	2.9
6	1.0	0.6	1.2	0	2.9	1.1	1.7	2.4
7	0	0	0.4	0	1.0	0	0	1.0
4–34	15.5	14.6	**8.9***	7.5	7.4	**0.9†**	**0†**	**3.9†**
3–30	0	3.5	**5.5***	**6.0***	**8.5†**	**10.0†**	**8.6†**	**5.8†**
IGHD								
1	10.7	10.6	8.2	5.2	12.7	7.6	27.3	6.4
2	21.4	17.4	19.6	15.5	12.7	24.2	4.5	21.2
3	35.9	34.2	40.3	46.6	41.2	25.8	18.2	37.2
4	5.8	8.5	6.3	7.8	6.9	10.6	13.6	10.9
5	9.7	8.3	8.8	9.5	6.9	13.6	9.1	10.3
6	14.6	20.3	15.9	15.5	18.6	12.1	27.3	10.3
7	1.9	0.6	0.7	0	1.0	6.1	0	3.8
IGHJ								
1	3.9	0.6	1.8	1.5	2.9	1.4	4.4	1.0
2	4.8	3.5	2.3	0	3.9	1.4	8.9	2.0
3	12.6	7.6	9.9	14.3	9.8	17.8	13.3	8.0
4	44.7	43.7	43.3	37.6	46.1	53.4	35.6	55.0
5	12.6	15.0	10.5	19.5	12.7	8.2	13.3	10.0
6	21.4	29.6	32.3	27.1	24.5	17.8	24.4	24.0

Frequencies expressed in percentages. References for comparisons are as follows: MCL,[27] CLL,[55] SMZL,[29] HCL,[13] MM,[24] WM,[26] and normal B cells.[20] Significant differences between DLBCL and other B-LPDs are depicted in bold.
*$P < 0.05$.
†$P < 0.01$.
MM, multiple myeloma; SMZL, splenic marginal zone lymphoma; WM, Waldeström macroglobulinemia.

1.56, $P < 0.01$), HCL (ratio 0.96, $P < 0.01$), and in normal B cells (ratio: 1.00, $P < 0.01$).[12,13,54]

Partial D-J rearrangements were present in 41% of the cases in our series (Table 2), with a similar IGHD usage to that of other B-LPDs (data not shown).

Discussion

Although DLBCL is the most frequent variant of non-Hodgkin lymphoma, its ontogeny is still not well understood. In this study, we analyzed complete V-D-J rearrangements in 165 untreated DLBCL, which is the largest series reported to date. Additionally, we also characterized the partial D-J rearrangements. Our results, based on the characterization of the BCR in DLBCLs, with a biased use of certain IGHV genes, evidence of canonical SHM, and the findings of shared amino acid changes and stereotyped sequences, provide new information about the ontogeny of DLBCL. Our data suggest that certain particular antigens must be involved in the development of DLBCL by stimulating the proliferation of B cells that express surface IG encoded by certain IGHV genes. Moreover, DLBCLs show clear differences from other B-LPDs with respect to IGHV usage (Table 5) and SHM.

A strong bias at the level of IGHV gene usage was observed, with the *IGHV4-34*, *IGHV3-23*, and *IGHV4-39* genes accounting for one third of the cohort. In addition, recurrent amino acid changes were found in certain IGHV genes, which may indicate specific driven SHM in DLBCL. These findings suggest that DLBCL could derive from certain B-cell populations specifically stimulated during the antigenic response. This hypothesis is even more plausible if the presence of some stereotyped sequences is taken into account. We identified three VH CDR3 stereotyped sequences in our series, which is a lower frequency than in CLL.[35] The stereotyped sequences had a rearrangement involving the *IGHV1-69* or *IGHV4-34* genes. Moreover, two of them shared their *IGHV1-69* rearrangement with subgroups previously described in CLL.[25,35] It is well known that there is a relationship between CLL and DLBCL, since some cases of DLBCL occur on a CLL background (Richter transformation). However, the cases reported here are *de novo* DLBCL, indicating "common" stereotypes between DLBCL and CLL, in contrast to other B-cell malignancies.[35] This highly restricted immunoglobulin gene repertoire with stereotyped VH CDR3 and common SHM targeting in DLBCL suggests a role for antigen selection in this entity, at least for some subsets of cases. Some of the antigens could be common to different B-LPDs but could ultimately lead to a different entity, depending on which population they stimulate. Although the results reported here support the notion that stereotyped BCRs in DLBCL do not seem to have as relevant implications as in CLL or other B-LPDs, the frequency of these stereotyped sequences and their possible biological and clinical significance in DLBCL need further analysis in larger and independent series.

The *IGHV4-34* gene was overrepresented in our series (15.5% overall), appearing at a higher frequency than in normal B cells,[16,20] in agreement with previous reports.[23,41,42] However, there are other series in which this percentage was slightly smaller, although this discrepancy could be accounted for in terms of differences in the sample size or the CGB/non-GCB distribution.[37,39,58] Interestingly, the patients showing *IGHV4-34* rearrangements in our series shared several common features: i) all but one were highly mutated (15 of 16 > 2%, median: 9.5%, range: 1.7% to 23.6%); ii) all of them (12 of 12) belonged to the non-GCB subtype; and iii) all (10 of 10) conserved the AVY motif. Moreover, the critical W at position 7 was intact in the five cases that could be analyzed. This specific FR1 motif is thought to be responsible for binding the *N*-acetyllactosamine antigenic determinant, implying that these cells may retain the ability to bind to and be activated by superantigens, despite intense SHM activity.[59] In light of these results, our findings are consistent with those from previous studies,[42,60] and we confirm, in a larger series, the strong association between DLBCL expressing the *IGHV4-34* and non-GCB subtype. However, these results require future confirmation using genome-wide expression profiles as a tool for distinguishing between the various biological subtypes of DLBCL. In addition, we describe a new stereotyped VH CDR3 pattern in DLBCL using *IGHV4-34*. The molecular features of this subgroup differ from other B-LPDs using this gene. For example, in HCL *IGHV4-34* is mutated in only 1 of 17 cases (6%)[13]; in the small proportion of CLLs with the IgG isotype,[61] the *IGHV4-34* is the predominant gene,[25] whereas all of our *IGHV4-34* DLBCL cases belonged to the non-GCB subtype, which is assumed to be IgM^+.[42] Finally, mutations in codons 64 and 80 are found in *IGHV4-34* sequences from DLBCL and from ocular adnexal extranodal marginal zone lymphoma, which is more frequent than in normal B cells or other B-LPDs.[20,25,27,57,62] Taken together, our results suggest that DLBCL expressing IGHV4-34 may constitute a separate subgroup within the non-GCB DLBCL subtype.

Several hypotheses have been proposed to explain the origin and frequency of clonal cells from non-GCB DLBCL. Lossos et al[63] reported that non-GCB DLBCL must derive from B cells that have passed through the GC with no possibility of additional SHM. However, Jardin et al[60] observed ongoing mutations in non-GCB DLBCL, including *IGHV4-34* cases, which would suggest an independent GC mechanism of SHM. Regardless of the GC dependence of these mutations, we observed that SHM of the non-GCB subtype met the canonical criteria (*IGHV4-34* included), which do not support the need to invoke a mutational machinery different from the normal process for IGHV mutations.

In conclusion, on the basis of particular features in the sequence of the immunoglobulin genes, we found that different subgroups exist within the non-GCB subtype. Our data support the idea that DLBCL is characterized by a highly distinctive IG gene repertoire and, to the best of our knowledge, for the first time, we describe stereotyped sequences in DLBCL and "common" stereotypes between CLL and other B-LPD. These results indicate a role for antigen selection in DLBCL development and also open possibilities for future investigations into the biology of DLBCL.

Acknowledgments

We thank Alicia Antón, Felicitación García, Montserrat Hernández-Ruano (Hospital Universitario de Salamanca, Spain), and Mar López-Colomer (Centro Nacional de Investigaciones Oncológicas, Madrid, Spain) for their technical support. In addition, we thank all of the members of GELTAMO and the patients who participated in this study.

References

1. Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri SA, Stein H, Thiele J, Vardiman JW (Eds); International Agency for Research on Cancer: WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues, 4th ed. Geneva Switzerland, World Health Organization Press, 2008
2. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X, Powell JI, Yang L, Marti GE, Moore T, Hudson J Jr., Lu L, Lewis DB, Tibshirani R, Sherlock G, Chan WC, Greiner TC, Weisenburger DD, Armitage JO, Warnke R, Levy R, Wilson W, Grever MR, Byrd JC, Botstein D, Brown PO, Staudt LM: Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature 2000, 403:503–511
3. Hans CP, Weisenburger DD, Greiner TC, Gascoyne RD, Delabie J, Ott G, Muller-Hermelink HK, Campo E, Braziel RM, Jaffe ES, Pan Z, Farinha P, Smith LM, Falini B, Banham AH, Rosenwald A, Staudt LM, Connors JM, Armitage JO, Chan WC: Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. Blood 2004, 103:275–282
4. Choi WW, Weisenburger DD, Greiner TC, Piris MA, Banham AH, Delabie J, Braziel RM, Geng H, Iqbal J, Lenz G, Vose JM, Hans CP, Fu K, Smith LM, Li M, Liu Z, Gascoyne RD, Rosenwald A, Ott G, Rimsza LM, Campo E, Jaffe ES, Jaye DL, Staudt LM, Chan WC: A new immunostain algorithm classifies diffuse large B-cell lymphoma into molecular subtypes with high accuracy. Clin Cancer Res 2009, 15: 5494–5502
5. Tonegawa S: Somatic generation of antibody diversity. Nature 1983, 302:575–581
6. Betz AG, Rada C, Pannell R, Milstein C, Neuberger MS: Passenger transgenes reveal intrinsic specificity of the antibody hypermutation mechanism: clustering, polarity, and specific hot spots. Proc Natl Acad Sci U S A 1993, 90:2385–2388
7. Wagner SD, Milstein C, Neuberger MS: Codon bias targets mutation. Nature 1995, 376:732
8. Dunn-Walters DK, Dogan A, Boursier L, MacDonald CM, Spencer J: Base-specific sequences that bias somatic hypermutation deduced by analysis of out-of-frame human IgVH genes. J Immunol 1998, 160:2360–2364
9. Smith DS, Creadon G, Jena PK, Portanova JP, Kotzin BL, Wysocki LJ: Di- and trinucleotide target preferences of somatic mutagenesis in normal and autoreactive B cells. J Immunol 1996, 156:2642–2652
10. Spencer J, Dunn M, Dunn-Walters DK: Characteristics of sequences around individual nucleotide substitutions in IgVH genes suggest different GC and AT mutators. J Immunol 1999, 162:6596–6601
11. Shapiro GS, Aviszus K, Murphy J, Wysocki LJ: Evolution of Ig DNA sequence to target specific base positions within codons for somatic hypermutation. J Immunol 2002, 168:2302–2306
12. Messmer BT, Albesiano E, Messmer D, Chiorazzi N: The pattern and distribution of immunoglobulin VH gene mutations in chronic lymphocytic leukemia B cells are consistent with the canonical somatic hypermutation process. Blood 2004, 103:3490–3495
13. Arons E, Roth L, Sapolsky J, Suntum T, Stetler-Stevenson M, Kreitman RJ: Evidence of canonical somatic hypermutation in hairy cell leukemia. Blood 2011, 117:4844–4851

14. Weller S, Faili A, Garcia C, Braun MC, Le Deist FF, de Saint Basile GG, Hermine O, Fischer A, Reynaud CA, Weill JC: CD40-CD40L independent Ig gene hypermutation suggests a second B cell diversification pathway in humans. Proc Natl Acad Sci U S A 2001, 98:1166–1170
15. Chiorazzi N, Ferrarini M: B cell chronic lymphocytic leukemia: lessons learned from studies of the B cell antigen receptor. Annu Rev Immunol 2003, 21:841–894
16. Berkowska MA, Driessen GJ, Bikos V, Grosserichter-Wagener C, Stamatopoulos K, Cerutti A, He B, Biermann K, Lange JF, van der Burg M, Van Dongen JJ, van Zelm MC: Human memory B cells originate from three distinct germinal center-dependent and -independent maturation pathways. Blood 2011, 118:2150–2158
17. Chang B, Casali P: The CDR1 sequences of a major proportion of human germline Ig VH genes are inherently susceptible to amino acid replacement. Immunol Today 1994, 15:367–373
18. Lossos IS, Tibshirani R, Narasimhan B, Levy R: The inference of antigen selection on Ig genes. J Immunol 2000, 165:5122–5126
19. Uduman M, Yaari G, Hershberg U, Stern JA, Shlomchik MJ, Kleinstein SH: Detecting selection in immunoglobulin sequences. Nucleic Acids Res 2011, 39:W499–W504
20. Brezinschek HP, Foster SJ, Brezinschek RI, Dorner T, Domiati-Saad R, Lipsky PE: Analysis of the human VH gene repertoire. Differential effects of selection and somatic hypermutation on human peripheral CD5(+)/IgM+ and CD5(−)/IgM+ B cells. J Clin Invest 1997, 99: 2488–2501
21. Wu YC, Kipling D, Leong HS, Martin V, Ademokun AA, Dunn-Walters DK: High-throughput immunoglobulin repertoire analysis distinguishes between human IgM memory and switched memory B-cell populations. Blood 2010, 116:1070–1078
22. Collins AM, Sewell WA, Edwards MR: Immunoglobulin gene rearrangement, repertoire diversity, and the allergic response. Pharmacol Ther 2003, 100:157–170
23. Hsu FJ, Levy R: Preferential use of the VH4 Ig gene family by diffuse large-cell lymphoma. Blood 1995, 86:3072–3082
24. Gonzalez D, van der Burg M, Garcia-Sanz R, Fenton JA, Langerak AW, Gonzalez M, Van Dongen JJ, San Miguel JF, Morgan GJ: Immunoglobulin gene rearrangements and the pathogenesis of multiple myeloma. Blood 2007, 110:3112–3121
25. Stamatopoulos K, Belessi C, Moreno C, Boudjograh M, Guida G, Smilevska T, Belhoul L, Stella S, Stavroyianni N, Crespo M, Hadzidimitriou A, Sutton L, Bosch F, Laoutaris N, Anagnostopoulos A, Montserrat E, Fassas A, Dighiero G, Caligaris-Cappio F, Merle-Beral H, Ghia P, Davi F: Over 20% of patients with chronic lymphocytic leukemia carry stereotyped receptors: pathogenetic implications and clinical correlations. Blood 2007, 109:259–270
26. Martin-Jimenez P, Garcia-Sanz R, Balanzategui A, Alcoceba M, Ocio E, Sanchez ML, Gonzalez M, San Miguel J: Molecular characterization of heavy chain immunoglobulin gene rearrangements in Waldenstrom's macroglobulinemia and IgM monoclonal gammopathy of undetermined significance. Haematologica 2007, 92:635–642
27. Hadzidimitriou A, Agathangelidis A, Darzentas N, Murray F, fau-Larue MH, Bredo PL, Navarro LA, Dagklis A, Rombout P, Beldjord K, Kolstad A, Dreyling MH, Anagnostopoulos A, Tsaftaris A, Mavragani-Tsipidou P, Rosenwald A, Ponzoni M, Groenen P, Ghia P, Sander B, Papadaki T, Campo E, Geisler C, Rosenquist R, Davi F, Pott C, Stamatopoulos K: Is there a role for antigen selection in mantle cell lymphoma? Immunogenetic support from a series of 807 cases. Blood 2011, 118:3088–3095
28. Bikos V, Darzentas N, Hadzidimitriou A, Davis Z, Hockley S, Traverse-Glehen A, Algara P, Santoro A, Gonzalez D, Mollejo M, Dagklis A, Gangemi F, Bosler DS, Bourikas G, Anagnostopoulos A, Tsaftaris A, Iannitto E, Ponzoni M, Felman P, Berger F, Belessi C, Ghia P, Papadaki T, Dogan A, Degano M, Matutes E, Piris MA, Oscier D, Stamatopoulos K: Over 30% of patients with splenic marginal zone lymphoma express the same immunoglobulin heavy variable gene: ontogenetic implications. Leukemia 2012, 26:1638–1646
29. Zibellini S, Capello D, Forconi F, Marcatili P, Rossi D, Rattotti S, Franceschetti S, Sozzi E, Cencini E, Marasca R, Baldini L, Tucci A, Bertoni F, Passamonti F, Orlandi E, Varettoni M, Merli M, Rizzi S, Gattei V, Tramontano A, Paulli M, Gaidano G, Arcaini L: Stereotyped patterns of B-cell receptor in splenic marginal zone lymphoma. Haematologica 2010, 95:1792–1796
30. Dagklis A, Ponzoni M, Govi S, Cangi MG, Pasini E, Charlotte F, Vino A, Doglioni C, Davi F, Lossos IS, Ntountas I, Papadaki T, Dolcetti R, Ferreri AJ, Stamatopoulos K, Ghia P: Immunoglobulin gene repertoire in ocular adnexal lymphomas: hints on the nature of the antigenic stimulation. Leukemia 2012, 26:814–821
31. Chan CH, Hadlock KG, Foung SK, Levy S: V(H)1–69 gene is preferentially used by hepatitis C virus-associated B cell lymphomas and by normal B cells responding to the E2 viral antigen. Blood 2001, 97:1023–1026
32. Cappione AJ, Pugh-Bernard AE, Anolik JH, Sanz I: Lupus IgG VH4.34 antibodies bind to a 220-kDa glycoform of CD45/B220 on the surface of human B lymphocytes. J Immunol 2004, 172:4298–4307
33. Messmer BT, Albesiano E, Efremov DG, Ghiotto F, Allen SL, Kolitz J, Foa R, Damle RN, Fais F, Messmer D, Rai KR, Ferrarini M, Chiorazzi N: Multiple distinct sets of stereotyped antigen receptors indicate a role for antigen in promoting chronic lymphocytic leukemia. J Exp Med 2004, 200:519–525
34. Murray F, Darzentas N, Hadzidimitriou A, Tobin G, Boudjogra M, Scielzo C, Laoutaris N, Karlsson K, Baran-Marzsak F, Tsaftaris A, Moreno C, Anagnostopoulos A, Caligaris-Cappio F, Vaur D, Ouzounis C, Belessi C, Ghia P, Davi F, Rosenquist R, Stamatopoulos K: Stereotyped patterns of somatic hypermutation in subsets of patients with chronic lymphocytic leukemia: implications for the role of antigen selection in leukemogenesis. Blood 2008, 111:1524–1533
35. Agathangelidis A, Darzentas N, Hadzidimitriou A, Brochet X, Murray F, Yan XJ, Davis Z, van Gastel-Mol EJ, Tresoldi C, Chu CC, Cahill N, Giudicelli V, Tichy B, Pedersen LB, Foroni L, Bonello L, Janus A, Smedby K, Anagnostopoulos A, Merle-Beral H, Laoutaris N, Juliusson G, di Celle PF, Pospisilova S, Jurlander J, Geisler C, Tsaftaris A, Lefranc MP, Langerak AW, Oscier DG, Chiorazzi N, Belessi C, Davi F, Rosenquist R, Ghia P, Stamatopoulos K: Stereotyped B-cell receptors in one third of chronic lymphocytic leukemia: towards a molecular classification with implications for targeted therapeutic interventions. Blood 2012, 119:4467–4475
36. Tobin G, Thunberg U, Johnson A, Eriksson I, Soderberg O, Karlsson K, Merup M, Juliusson G, Vilpo J, Enblad G, Sundstrom C, Roos G, Rosenquist R: Chronic lymphocytic leukemias utilizing the VH3-21 gene display highly restricted Vlambda2-14 gene use and homologous CDR3s: implicating recognition of a common antigen epitope. Blood 2003, 101:4952–4957
37. Kuppers R, Rajewsky K, Hansmann ML: Diffuse large cell lymphomas are derived from mature B cells carrying V region genes with a high load of somatic mutation and evidence of selection for antibody expression. Eur J Immunol 1997, 27:1398–1405
38. Rosenquist R, Lindstrom A, Holmberg D, Lindh J, Roos G: V(H) gene family utilization in different B-cell lymphoma subgroups. Eur J Haematol 1999, 62:123–128
39. Lossos IS, Okada CY, Tibshirani R, Warnke R, Vose JM, Greiner TC, Levy R: Molecular analysis of immunoglobulin genes in diffuse large B-cell lymphomas. Blood 2000, 95:1797–1803
40. Nakamura N, Kuze T, Hashimoto Y, Hara Y, Hoshi S, Sasaki Y, Shirakawa A, Sato M, Abe M: Analysis of the immunoglobulin heavy chain gene variable region of CD5-positive and -negative diffuse large B cell lymphoma. Leukemia 2001, 15:452–457
41. Capello D, Martini M, Gloghini A, Cerri M, Rasi S, Deambrogi C, Rossi D, Spina M, Tirelli U, Larocca LM, Carbone A, Gaidano G: Molecular analysis of immunoglobulin variable genes in human immunodeficiency virus-related non-Hodgkin's lymphoma reveals implications for disease pathogenesis and histogenesis. Haematologica 2008, 93:1178–1185
42. Ruminy P, Etancelin P, Couronne L, Parmentier F, Rainville V, Mareschal S, Bohers E, Burgot C, Cornic M, Bertrand P, Lenormand B, Picquenot JM, Jardin F, Tilly H, Bastard C: The isotype of the BCR as a surrogate for the GCB and ABC molecular subtypes in diffuse large B-cell lymphoma. Leukemia 2011, 25:681–688
43. Catherwood MA, Gonzalez D, Patton C, Dobbin E, Venkatraman L, Alexander HD: Improved clonality assessment in germinal centre/post-germinal centre non-Hodgkin's lymphomas with high rates of somatic hypermutation. J Clin Pathol 2007, 60:524–528
44. A clinical evaluation of the International Lymphoma Study Group classification of non-Hodgkin's lymphoma. The Non-Hodgkin's Lymphoma Classification Project. Blood 1997, 89:3909–3918
45. Gonzalez M, Gonzalez D, Lopez-Perez R, Garcia-Sanz R, Chillon MC, Balanzategui A, Mateos MV, Alaejos I, Langerak AW, Orfao A, Van Dongen JJ, San Miguel JF: Heteroduplex analysis of VDJ amplified segments from rearranged IgH genes for clonality assessments in

B-cell non-Hodgkin's lymphoma. A comparison between different strategies. Haematologica 1999, 84:779–784
46. Van Dongen JJ, Langerak AW, Bruggemann M, Evans PA, Hummel M, Lavender FL, Delabesse E, Davi F, Schuuring E, Garcia-Sanz R, van Krieken JH, Droese J, Gonzalez D, Bastard C, White HE, Spaargaren M, Gonzalez M, Parreira A, Smith JL, Morgan GJ, Kneba M, Macintyre EA: Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936. Leukemia 2003, 17: 2257–2317
47. Lefranc MP, Pommie C, Ruiz M, Giudicelli V, Foulquier E, Truong L, Thouvenin-Contet V, Lefranc G: IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains. Dev Comp Immunol 2003, 27:55–77
48. Tsakou E, Agathagelidis A, Boudjoghra M, Raff T, Dagklis A, Chatzouli M, Smilevska T, Bourikas G, Merle-Beral H, Manioudaki-Kavallieratou E, Anagnostopoulos A, Bruggemann M, Davi F, Stamatopoulos K, Belessi C: Partial versus productive IGH locus rearrangements in chronic lymphocytic leukemia: implications for B-cell receptor stereotypy. Mol Med 2012, 18:138–145
49. Gonzalez D, Gonzalez M, Balanzategui A, Sarasquete ME, Lopez-Perez R, Chillon MC, Garcia-Sanz R, San Miguel JF: Molecular characteristics and gene segment usage in IGH gene rearrangements in multiple myeloma. Haematologica 2005, 90:906–913
50. Martin-Jimenez P, Garcia-Sanz R, Gonzalez D, Balanzategui A, Perez JJ, Caballero MD, Sarasquete ME, Galende J, Orfao A, Lopez-Berges MC, San Miguel JF, Gonzalez M: Molecular characterization of complete and incomplete immunoglobulin heavy chain gene rearrangements in hairy cell leukemia. Clin Lymphoma Myeloma 2007, 7:573–579
51. Hamblin TJ, Davis Z, Gardiner A, Oscier DG, Stevenson FK: Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. Blood 1999, 94:1848–1854
52. Davis ZA, Orchard JA, Corcoran MM, Oscier DG: Divergence from the germ-line sequence in unmutated chronic lymphocytic leukemia is due to somatic mutation rather than polymorphisms. Blood 2003, 102:3075
53. Hockley SL, Giannouli S, Morilla A, Wotherspoon A, Morgan GJ, Matutes E, Gonzalez D: Insight into the molecular pathogenesis of hairy cell leukaemia, hairy cell leukaemia variant and splenic marginal zone lymphoma, provided by the analysis of their IGH rearrangements and somatic hypermutation patterns. Br J Haematol 2010, 148:666–669
54. Di Noia JM, Neuberger MS: Molecular mechanisms of antibody somatic hypermutation. Annu Rev Biochem 2007, 76:1–22
55. Darzentas N, Hadzidimitriou A, Murray F, Hatzi K, Josefsson P, Laoutaris N, Moreno C, Anagnostopoulos A, Jurlander J, Tsaftaris A, Chiorazzi N, Belessi C, Ghia P, Rosenquist R, Davi F, Stamatopoulos K: A different ontogenesis for chronic lymphocytic leukemia cases carrying stereotyped antigen receptors: molecular and computational evidence. Leukemia 2010, 24:125–132
56. Smit LA, Bende RJ, Aten J, Guikema JE, Aarts WM, van Noesel CJ: Expression of activation-induced cytidine deaminase is confined to B-cell non-Hodgkin's lymphomas of germinal-center phenotype. Cancer Res 2003, 63:3894–3898
57. Klein U, Rajewsky K, Kuppers R: Human immunoglobulin (Ig)M+IgD+ peripheral blood B cells expressing the CD27 cell surface antigen carry somatically mutated variable region genes: cD27 as a general marker for somatically mutated (memory) B cells. J Exp Med 1998, 188:1679–1689
58. Daley MD, Berinstein NL, Siminovitch KA: Immunoglobulin heavy chain variable gene utilization in human large cell and Burkitt's lymphoma cell lines. Clin Invest Med 1994, 17:522–530
59. Li Y, Spellerberg MB, Stevenson FK, Capra JD, Potter KN: The I binding specificity of human VH 4–34 (VH 4–21) encoded antibodies is determined by both VH framework region 1 and complementarity determining region 3. J Mol Biol 1996, 256:577–589
60. Jardin F, Sahota SS, Ruminy P, Parmentier F, Picquenot JM, Rainville V, Buchonnet G, Lepretre S, Tilly H, Bastard C: Novel Ig V gene features of t(14;18) and t(3;14) de novo diffuse large B-cell lymphoma displaying germinal center-B cell like and non-germinal center-B cell like markers. Leukemia 2006, 20:2070–2074
61. Potter KN, Mockridge CI, Neville L, Wheatley I, Schenk M, Orchard J, Duncombe AS, Packham G, Stevenson FK: Structural and functional features of the B-cell receptor in IgG-positive chronic lymphocytic leukemia. Clin Cancer Res 2006, 12:1672–1679
62. Zhu D, Lossos C, Chapman-Fredricks JR, Matthews JM, Ikpatt OF, Ruiz P, Lossos IS: Biased use of the IGHV4 family and evidence for antigen selection in Chlamydophila psittaci-negative ocular adnexal extranodal marginal zone lymphomas. PLoS One 2011, 6:e29114
63. Lossos IS, Alizadeh AA, Eisen MB, Chan WC, Brown PO, Botstein D, Staudt LM, Levy R: Ongoing immunoglobulin somatic mutation in germinal center B cell-like but not in activated B cell-like diffuse large cell lymphomas. Proc Natl Acad Sci U S A 2000, 97:10209–10213

Supplemental Table S1. Summary of recurrent mutations in complete V-D-J rearrangements of the predominant IGHV genes in DLBCL

IGHV gene	Total No of rearrangements	Region	Codon	No of rearrang. changing *(to any AA)*	AA from *(Germline)*	AA to* *(new AA)*	Frequency of new AA
IGHV1-18	4	FR3	71	4	L	F	75.0
			89	4	M	L	50.0
			93	4	S	R	50.0
IGHV1-69	4	CDR2	57	2	I	T	50.0
			59	2	I	M	50.0
		FR3	85	2	S	†	50.0
			87	2	A	V	50.0
IGHV3-21	5	CDR1	38	3	S	N	40.0
		FR2	40	3	N	T	60.0
		CDR2	59	3	S	†	60.0
IGHV3-23	8	CDR1	29	5	T	S	37.5
			36	6	S	N	50.0
	10	FR2	40	5	S	N	37.5
			55	10	A	T	30.0
						S	20.0
						G	20.0
						N	20.0
		CDR2	63	7	G	D	40.0
			64	7	S	N	30.0
IGHV4-34	13	FR2	40	8	S	T	30.8
						N	15.4
	16	CDR2	64	8	S	N	18.8
						R	12.5
						I	12.5
		FR3	80	8	V	I	18.8
						A	12.5

Supplemental Table S1. (continued)

	Total No of rearrangements	Region	Codon	No of rearrang. changing *(to any AA)*	AA from *(Germline)*	AA to* *(new AA)*	Frequency of new AA
IGHV4-34 (continued)	16	FR3	90	8	K	R	25.0
						T	12.5
			92	13	S	T	37.5
						N	25.0
IGHV4-39	7	CDR1	28	5	G	D	42.9
			34	7	S	R	42.9
						T	28.6
	8	FR2	45	5	P	S	50.0
			55	7	S	N	37.5
		CDR2	57	5	Y	S	25.0
						F	25.0
			64	6	S	N	25.0
		FR3	76	6	V	L	25.0
						I	25.0
						A	25.0
			80	6	V	I	62.5
			90	6	K	R	25.0
						Q	25.0
			92	7	S	T	37.5
			103	5	Y	F	62.5
IGHV4-59	5	CDR2	64	4	S	T	40.0
		FR3	80	5	V	L	40.0
						I	40.0
			90	4	K	R	40.0

* Only the most frequent are shown, † There was no preference as regards the AA substituting the original one

Supplemental Table S2. Stereotyped VH CDR3 amino acid sequences. Sequences derived from the present study are marked in bold.

Subset / Seq. ID	IGHV	IGHD	IGHJ	Mutational status	VH CDR3 amino acid sequences
SUBSET 5	Published				
11-3436	**IGHV1-69*13**	**IGHD3-10*01**	**IGHJ6*03**	**100**	**ARTMVRGVINIDYYYYYMDV**
CZ-01-0060-H1	IGHV1-69*01	IGHD3-3*01	IGHJ6*02	100	ARSEIFGVVELYYYYYGMDV
CZ-01-0131-H1	IGHV1-69*01	IGHD3-10*01	IGHJ6*02	100	ARAMVQGVITYYYYYGMDV
CZ-01-0385-H1	IGHV1-69*06	IGHD3-10*01	IGHJ6*02	99.31	ASGVVRGVIPIYYYYYGMDV
FR-01-0138-H1	IGHV1-69*01	IGHD6-13*01	IGHJ6*03	100	ARDWGSSWYPAYYYYYYMDV
FR-01-0286-H1	IGHV1-69*01	IGHD3-10*01	IGHJ6*03	100	ARVMVRGVISLDYYYYYMDV
FR-01-0490-H1	IGHV1-69*01	IGHD2-2*01	IGHJ6*03	100	ARVRTEGVVPAFYYYYYMDV
FR-01-0524-H1	IGHV1-69*01	IGHD3-10*01	IGHJ6*02	100	ARDSVRGVISYYYYYYGMDV
FR-01-0565-H1	IGHV1-69*01	IGHD3-10*01	IGHJ6*02	100	ARWEVRGVISYYYYYYGMDV
FR-01-0568-H1	IGHV1-69*01	IGHD3-3*01	IGHJ6*02	100	ASITIFGVVNVAYYYYGMDV
FR-01-0736-H2	IGHV1-69*01	IGHD7-27*01	IGHJ6*02	100	ARDQGTLTNWGPYYYYGMDV
FR-01-0802-H1	IGHV1-69*01	IGHD3-3*01	IGHJ6*03	100	ARVTIFGVVSYYYYYYYMDV
FR-01-0823-H1	IGHV1-69*11	IGHD3-3*01	IGHJ6*03	99.65	ARGQIFGVVSILYYYYYMDV
FR-01-0933-H1	IGHV1-69*01	IGHD3-3*01	IGHJ6*02	99.65	AGITIFGVVNIYYYYYGMDV
GR-01-0079-H1	IGHV1-2*04	IGHD2-2*01	IGHJ6*02	100	ARSSLVVPAAIGYYYYGMDV
GR-01-0175-H1	IGHV1-69*01	IGHD3-10*01	IGHJ6*02	99.64	AVLLWFGDWSVYYYYYGMDV
GR-01-0186-H1	IGHV1-69*06	IGHD3-10*01	IGHJ6*01	100	ARDAVRGVIGVYYYYYGMDV
GR-01-0264-H1	IGHV1-69*01	IGHD2-21*02	IGHJ6*02	99.31	ARANGVVTALGYYYYYGMDV
GR-02-0510-H1	IGHV1-69*01	IGHD3-10*01	IGHJ6*03	100	ARSMVQGVINAYYYYYYMDV
IT-01-0165-H1	IGHV1-69*01	IGHD5-12*01	IGHJ6*03	99.65	ARKDIVATITTYYYYYYMDV
IT-01-0167-H1	IGHV1-69*01	IGHD3-10*01	IGHJ6*02	100	ARDMVRGVIPVYYYYYGMDV
IT-01-0185-H1	IGHV1-69*01	IGHD3-10*01	IGHJ6*02	100	ARSKVRGVIPIYYYYYGMDV
IT-01-0186-H1	IGHV1-69*01	IGHD3-10*01	IGHJ6*02	100	AREQVRGVISILVYYYGMDV
IT-01-0187-H1	IGHV1-69*01	IGHD3-10*01	IGHJ6*03	100	ARVMVQGVIALSYYYYYMDV

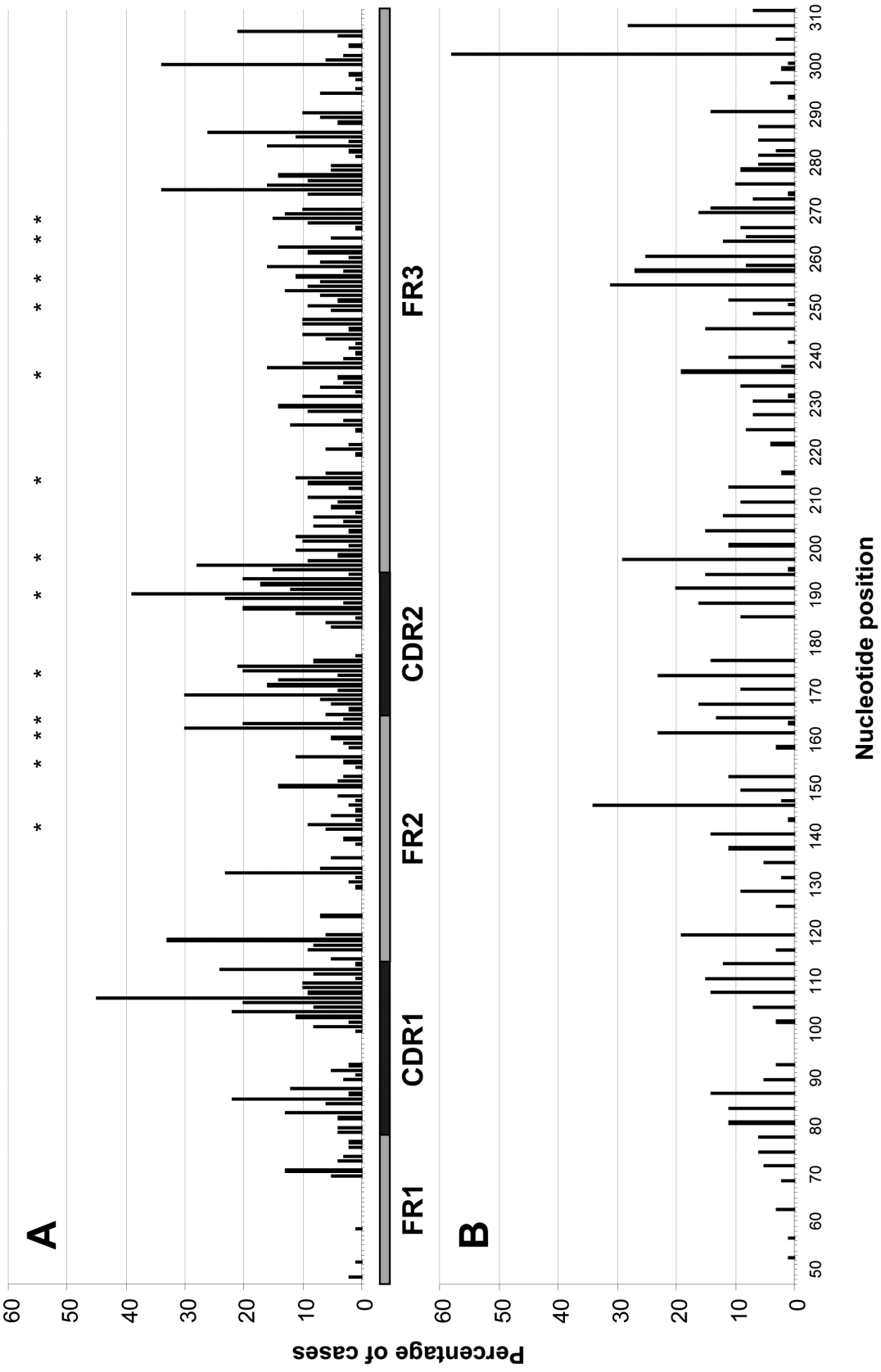
Supplemental Table S2. (continued)

Subset / Seq ID	IGHV	IGHD	IGHJ	Mutational	VH CDR3 amino acid sequences
IT-01-0188-H1	IGHV1-69*01	IGHD3-10*01	IGHJ6*02	100	ALTMVRGVIIFTYYYYGMDV
IT-01-0189-H1	IGHV1-69*01	IGHD3-3*01	IGHJ6*02	100	ARDRISGVVNIYYYYYGMDV
IT-02-0158-H1	IGHV1-69*01	IGHD3-10*01	IGHJ6*03	100	ARSTVRGVIWTLLDYYYMDV
IT-02-0159-H1	IGHV1-69*01	IGHD3-10*01	IGHJ6*03	100	ARARVQGVIGAYYYYYYMDV
IT-02-0194-H1	IGHV1-2*02	IGHD2-15*01	IGHJ6*02	100	AREEDVVVVAASYYYYGMDV
IT-02-0308-H1	IGHV1-69*01	IGHD3-10*01	IGHJ6*04	100	ARDEVQGVIDYYYYYYGMDV
IT-02-0387-H1	IGHV1-2*02	IGHD5-5*01	IGHJ6*02	100	AREEGVDTAMAYYYYYGMDV
IT-02-0459-H1	IGHV1-69*01	IGHD2-2*01	IGHJ6*02	100	ARDQPIVVVPVNYYYYGMDV
NL-01-0156-H1	IGHV1-69*06	IGHD6-13*01	IGHJ6*03	100	ARAGSSSWYEAYYYYYYMDV
NL-01-0246-H1	IGHV1-69*01	IGHD3-3*01	IGHJ6*02	100	ARGGVFGVVIISLNYYGMDV
NL-01-0294-H1	IGHV1-69*01	IGHD3-10*01	IGHJ6*03	100	ARTMVQGVIRVYYYYYYMDV
NY-01-0030-H1	IGHV1-69*01	IGHD3-3*01	IGHJ6*03	100	ARVEIFGVVNLNYYYYYMDV
NY-01-0070-H1	IGHV1-69*01	IGHD3-3*01	IGHJ6*03	100	ARVEIFGVVGLSYYYYYMDV
NY-01-0108-H1	IGHV1-69*01	IGHD1-26*01	IGHJ6*02	100	AREGVSGSYQVYYYYYGMDV
NY-01-0394-H1	IGHV1-69*01	IGHD3-10*01	IGHJ6*02	100	ARGPVRGVINVYYYYYGMDV
NY-01-0509-H1	IGHV1-69*01	IGHD3-10*01	IGHJ6*02	99.65	ARGMVQGVIPVEVYYYGMDV
NY-01-0641-H1	IGHV1-69*01	IGHD3-10*01	IGHJ6*03	100	ARATVRGVINVYYYYYYMDV
NY-01-0899-H1	IGHV1-69*13	IGHD3-16*02	IGHJ6*03	100	ARAMVQGVIQTYYYYYYMDV
SE-01-0069-H1	IGHV1-69*01	IGHD3-10*01	IGHJ6*01	100	ARAMVRGVINYYYYYYGMDV
SE-01-0390-H1	IGHV1-69*01	IGHD3-10*01	IGHJ6*03	98.16	ARGMVQGVINFSYYYYYMDV
SE-01-0393-H1	IGHV1-69*01	IGHD3-10*01	IGHJ6*01	100	ARDPVRGVISVYYYYYGMDV
SE-01-0444-H1	IGHV1-69*01	IGHD3-10*01	IGHJ6*03	99.6	ARGMVRGVINIDYYYYYMDV
SE-01-0782-H1	IGHV1-69*01	IGHD3-3*01	IGHJ6*02	100	AREGGLITIFGNYYYYGMDV
UK-01-0229-H1	IGHV1-69*01	IGHD3-3*01	IGHJ6*02	100	ARDPREGVVNLYYYYYGMDV
UK-01-0345-H1	IGHV1-69*01	IGHD3-10*01	IGHJ6*03	100	ARTMVRGVIRVMYYYYYMDV
UK-01-0387-H1	IGHV1-69*01	IGHD3-10*01	IGHJ6*03	100	ARAMVRGVITVVYYYYYMDV
UK-01-0437-H1	IGHV1-69*01	IGHD3-3*01	IGHJ6*03	100	ARVVLAGVVSIYYYYYYMDV

Supplemental Table S2. (continued)

Subset / Seq ID	IGHV	IGHD	IGHJ	Mutational	VH CDR3 amino acid sequences
UK-01-0539-H1	IGHV1-69*01	IGHD3-10*01	IGHJ6*03	100	ARAMVRGVIQVNYYYYYMDV
UK-01-0733-H1	IGHV1-69*01	IGHD3-10*01	IGHJ6*02	99.65	ARSMVRGVIHVAYYYYGMDV
UK-02-0288-H1	IGHV1-69*01	IGHD3-22*01	IGHJ6*03	100	ARVGLRGIVVVNDYYYYMDV
SUBSET 7D	Published				
11-1949	**IGHV1-69*01**	**IGHD3-3*01**	**IGHJ6*02**	**100**	**ASPVGEGDDFWSGYYPNYYDYGMDV**
CZ-01-0372-H1	IGHV1-69*01	IGHD3-3*01	IGHJ6*02	100	ASSLGENYDFWSGYYPNYYYYGMDV
CZ-01-0484-H1	IGHV1-69*01	IGHD3-3*01	IGHJ6*02	100	ARQGTLGYDFWSGYYPNYYYYGMDV
FR-01-0143-H1	IGHV1-69*01	IGHD3-3*01	IGHJ6*02	100	ARPVQAKYDFWSGYYPNYYYYGMDV
FR-01-0213-H1	IGHV1-69*06	IGHD3-3*01	IGHJ6*02	95.09	ARFAPNYDFWSGYQIARYYYYGMDV
GR-01-0505-H1	IGHV1-69*01	IGHD3-3*01	IGHJ6*02	100	ARGSPNYDFWSGYSPGYYYYYGMDV
GR-02-0303-H1	IGHV1-69*01	IGHD3-3*01	IGHJ6*02	99.65	ARATFSKGDFWSGYYPNYYYYGMDV
IT-02-0357-H1	IGHV1-69*01	IGHD3-3*01	IGHJ6*02	100	ARNTPNYDFWSGYTPHYYYYYGMDV
SE-01-0670-H1	IGHV1-69*01	IGHD3-3*01	IGHJ6*01	100	ARAVKNYDFWSALSVGYYYYYGMDV
SE-01-0848-H1	IGHV1-69*12	IGHD3-3*01	IGHJ6*02	100	ARGTPNYDFWSGYAPGYYYYYGMDV
SE-01-1016-H1	IGHV1-69*01	IGHD3-3*01	IGHJ6*02	99.65	ARPGGLYYDFWSGYYPNYYYYGMDV
UK-01-0427-H1	IGHV1-69*13	IGHD3-3*01	IGHJ6*02	100	ARSGREDYDFWSGYYPNYYYYGMDV
UK-01-0447-H1	IGHV1-69*01	IGHD3-3*01	IGHJ6*02	100	ARGTPNYDFWSGYSYAYYYYYGMDV
UK-01-0581-H1	IGHV1-69*06	IGHD3-3*01	IGHJ6*02	100	ARSCPGDDFWSGYYNGYYYYYGMDV
UK-01-0740-H1	IGHV1-69*01	IGHD3-3*01	IGHJ6*02	100	ARSTPNYDFWSGYTVGYYYYYGMDV
UK-02-0096-H1	IGHV1-69*06	IGHD3-3*01	IGHJ6*04	99.59	ARGQGNYDFWSGWSLANYYYYGMDV
SUBSET A	Novel				
11-1967	**IGHV4-34*01**	**IGHD3-22*01**	**IGHJ2*01**	**93.8**	**ARALDYYDNSGRLPGYFDL**
HQ392881.1	IGHV4-34*01	IGHD3-22*01	IGHJ2*01	95.7	ARALDYYDSSGRVLAYFDL

FIGURE S1


A
B
FR1
CDR1
FR2
CDR2
FR3
Percentage of cases
Nucleotide position
0
10
20
30
40
50
60
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
210
220
230
240
250
260
270
280
290
300
310

2. SEGUNDO TRABAJO: ANÁLISIS DE LAS ALTERACIONES EN EL NÚMERO DE COPIAS POR ALTA RESOLUCIÓN DE MUESTRAS PAREADAS NORMAL-TUMORAL DE LINFOMA B DIFUSO DE CÉLULA GRANDE

Introducción y Objetivos

Estudios previos de las alteraciones genéticas en el número de copias (*copy number alteration,* CNA) en linfoma B difuso de célula grande (LBDCG) se han llevado a cabo con diferentes tecnologías. En la mayoría de las series sólo se analiza la muestra tumoral. Sin embargo, el estudio comparativo de ADN de muestra tumoral con la pareada no tumoral del mismo paciente es importante para distinguir las pérdidas de heterocigosidad con número de copias normales (CNN-LOH) germinales, de las adquiridas por el tumor.

En este trabajo se analizan las CNAs de 60 pacientes con LBDCG, 40 de ellos con muestra normal pareada.

El objetivo principal es la correcta identificación de las CNAs y CNN-LOH germinales y de las lesiones adquiridas por las células tumorales, basándonos en estos datos:

– construir un modelo predictor con CNAs específicas, que nos pueda ayudar a clasificar los pacientes con LBDCG en los dos subtipos histológicos, centro germinal B (CGB) y no centro germinal (no-CGB).

– analizar la relación entre las CNAs y las características clínicas y biológicas de nuestra serie de pacientes con LBDCG, identificando alteraciones que podrían estar asociadas con el pronóstico.

Material y métodos

Se incluyeron 60 pacientes con un diagnóstico de LBDCG *de novo*, de acuerdo con la Clasificación Mundial de la Salud 2008. En 40 de ellos estaba disponible la muestra normal pareada. Las muestras fueron clasificadas en no-CGB *vs* CGB, según el algoritmo de Hans. La mayoría de los pacientes habían sido tratados con regímenes basados en rituximab.

El DNA tumoral se obtuvo de las muestras del diagnóstico. Se comprobó la existencia de un suficiente número de células tumorales para garantizar la fiabilidad de los estudios. El DNA no tumoral se obtuvo en la mayoría de los casos de muestra de sangre periférica no infiltrada al diagnóstico.

En 38 de 60 casos disponíamos de DNA de tejido fijado con formol e incluido en parafina y se hibridaron con MIP (*molecular inversion probe*) arrays usando el Oncoscan FFPE Express custom service (Affymetrix, Santa Clara, CA, USA). En los 22 casos restantes en los que disponíamos de tejido fresco o congelado, y en las 40 muestras de DNA no tumoral, se utilizó el SNP 6.0 array. Las ganancias y pérdidas fueron evaluadas con inspección visual de 2 observadores usando el software Nexus Biodiscovery version 7 (Biodiscovery, El Segundo, CA, USA). Se consideraron las CNA con un tamaño de 100 kb y las pérdidas de heterocigosidad (LOH) con un tamaño de 5 Mb. Las alteraciones del tumor se compararon con el DNA constitucional de cada muestra, y en los casos en los que el DNA normal no estaba disponible, las alteraciones fueron comprobadas con las bases de datos de *copy number variation*, CNV. Los hallazgos en la muestra no tumoral de los pacientes con LBDCG fueron comparados con una serie de 500 controles sanos españoles.

Se analizaron también las mutaciones del gen *TP53*, y se realizó el tipaje HLA de clase I y II de los pacientes.

Para el análisis estadístico se usó el programa IBM SPSS v.23.0

Resultados

El análisis de CN se realizó de forma exitosa en los 60 pacientes. No se encontraron diferencias relevantes entre las dos plataformas, por lo que los datos se analizaron de forma conjunta en una sola base de datos. Las principales comparaciones se hicieron en base a los dos grupos clasificados con el algoritmo de Hans (no-CGB, n=32 vs CGB, n=23).

Todos los casos, excepto uno, presentaron CNAs, identificamos 63 CNAs recurrentes y 9 CNN-LOH. Encontramos 59 regiones de CNN-LOH germinales, pero no observamos CNAs germinales.

Se observaron deleciones homocigotas de genes supresores de tumores como *CDKN2A* y *B*, en 3 casos y otras deleciones homocigotas en casos individuales, así como

6p21 (HLA clase II), 15q21 (B2M), entre otras. Además, detectamos amplificaciones que implican genes con un importante papel en la linfomagénesis, siendo la más frecuente en 2p15-p16.1, implicando a los genes *REL y BCL11A*, en el 13% de los casos.

Al evaluar las CNAs de acuerdo con los grupos de la clasificación inmunohistoquímica (CGB *vs* no-CGB) vimos que algunas alteraciones estaban significativamente asociadas con uno de los dos subtipos. El modelo que mejor discriminó entre CGB y no-CGB incluía las siguientes alteraciones: las pérdidas de 1p36.32 y 15q21.1 y las ganancias del cromosoma 7 como predictores del subtipo CGB, mientras que las pérdidas de 8p23.3 y las ganancias de 11q24.3 fueron más específicas del subtipo no-CGB.

Con respecto a las pérdidas de heterocigosidad, la CNN-LOH más frecuente en las muestras tumorales de nuestra serie de LBDCG se encontró en la región 6p21.3, en el 20% de los casos.

Al analizar las muestras no tumorales, se detectaron CNN-LOH germinales en 16 de los 40 casos. Las únicas recurrentes fueron 6p21 y 9p24.1 (n=2). Estas CNN-LOH en células no tumorales de pacientes con LBDCG se compararon con las CNN-LOH de una serie de 500 individuos sanos. Algunas de las regiones de CNN-LOH, se encontraron más frecuentemente en los pacientes con LBDCG que en los sujetos sanos.

Otros estudios genéticos

Un 28% de los casos presentaron anomalías en el gen *TP53*: deleción en 17%, CNN-LOH en 5% y mutación en el 10%. Dos de los casos presentaron simultáneamente mutación y CNN-LOH. Además, estos dos casos, eran los genéticamente más complejos, ya que mostraron un mayor número de alteraciones, comparado con el resto de los casos.

Correlaciones clínicas

La mediana de seguimiento de la serie fue de 56 meses (rango 20-211), con una supervivencia libre de progresión (PFS) y supervivencia global (OS), ambas a 5 años de 67 y 74% respectivamente para la serie completa.

Algunas CNAs tuvieron un valor pronóstico en nuestra serie. Así, los pacientes con pérdida de 18p11.32 tuvieron una menor SLP (71% vs 33%, p=0.027). Por el contrario, las ganancias de 11q se relacionaron con un mejor pronóstico en esta serie.

Los pacientes con inactivación bialélica de la región 17p13.1 que implica el gen *TP53*, tuvieron una SLP y SG muy baja con respecto al resto de la serie (p=0.0007 y p=0.014 respectivamente).

Conclusiones

Nuestro estudio presenta un análisis detallado de 60 LBDCG mediante CN-*arrays* de alta resolución. Una aportación importante de nuestro trabajo es el estudio simultáneo de la muestra no tumoral pareada, lo que permite una distinción precisa entre las alteraciones en el número de copias y pérdidas de heterocigosidad germinales de las adquiridas por el tumor.

– Nuestro estudio presenta buena concordancia con la frecuencia de las alteraciones más importantes descritas en estudios previos.

– Describimos alteraciones en genes importantes para la linfomagénesis y el sistema inmune: así la amplificación más frecuente en nuestra serie fue 2p16, que incluye los genes *REL* y *BCL11A*. Y detectamos pérdidas homocigotas en 6p21.32 *(HLA II)* y 15q21 (*B2M*).

– Encontramos algunas CNAs específicas en los dos subtipos inmunohistoquímicos. Un modelo con algunas CNAs podría ayudar a predecir el subtipo de LBDCG, estando las pérdidas de 1p36.32 y 10q23.31 restringidas al subtipo CGB, y por el contrario, las ganancias de 11q24.3 y las pérdidas de 8p23.3 serían altamente características del subtipo no-CGB. Hasta donde sabemos, este es el primer modelo derivado de alteraciones de CNA, para predecir el subtipo de LBDCG.

– Identificamos CNN-LOH de 6p21 en el 20% de los casos, que implica a la región del HLA, corroborando su importante papel en la patogenia del LBDCG.

– Además observamos CNN-LOH en las células no tumorales de pacientes con LBDCG, que no habían sido reportadas previamente, lo que podría indicar que ciertos genes podrían estar implicados en la aparición del LBDCG.

– Con respecto a la evolución clínica, a inactivación bialélica de *TP53* y la pérdida de 8p11.32 se asociaron con un peor pronóstico, mientras que los casos con ganancias de 11q24.3 parecen tener un mejor pronóstico.

En resumen, las CNAs identifican subtipos de LBDCG específicos en cuanto a sus características biológicas y pronósticas. Además, aportamos que ciertos genes podrían estar alterados en las células no tumorales de los pacientes con LBDCG, lo que implicaría un papel en su patogenia.

Ann Hematol
DOI 10.1007/s00277-015-2552-3

ORIGINAL ARTICLE

High-resolution copy number analysis of paired normal-tumor samples from diffuse large B cell lymphoma

Elena Sebastián[1,2,3] · Miguel Alcoceba[1,2,3] · David Martín-García[4] · Óscar Blanco[5] · Mercedes Sanchez-Barba[6] · Ana Balanzategui[1,2] · Luis Marín[1,2] · Santiago Montes-Moreno[3,7] · Eva González-Barca[3] · Emilia Pardal[3] · Cristina Jiménez[1] · María García-Álvarez[1] · Guillem Clot[4] · Ángel Carracedo[8,9] · Norma C. Gutiérrez[1,2] · M. Eugenia Sarasquete[1,2] · Carmen Chillón[1,2] · Rocío Corral[1,2] · M. Isabel Prieto-Conde[1] · M. Dolores Caballero[1,2,3] · Itziar Salaverria[4] · Ramón García-Sanz[1,2,3,10] · Marcos González[1,2,10]

Received: 13 August 2015 / Accepted: 6 November 2015

Abstract Copy number analysis can be useful for assessing prognosis in diffuse large B cell lymphoma (DLBCL). We analyzed copy number data from tumor samples of 60 patients diagnosed with DLBCL de novo and their matched normal samples. We detected 63 recurrent copy number alterations (CNAs), including 33 gains, 30 losses, and nine recurrent acquired copy number neutral loss of heterozygosity (CNN-LOH). Interestingly, 20 % of cases acquired CNN-LOH of 6p21 locus, which involves the HLA region. In normal cells, there were no CNAs but we observed CNN-LOH involving some key lymphoma regions such as 6p21 and 9p24.1 (5 %) and 17p13.1 (2.5 %) in DLBCL patients. Furthermore, a model with some specific CNA was able to predict the subtype of DLBCL, 1p36.32 and 10q23.31 losses being restricted to germinal center B cell-like (GCB) DLBCL. In contrast, 8p23.3 losses and 11q24.3 gains were strongly associated with the non-GCB subtype. A poor prognosis was associated with biallelic inactivation of *TP53* or 18p11.32 losses, while prognosis was better in cases carrying 11q24.3 gains. In summary, CNA abnormalities identify specific DLBCL groups, and we describe CNN-LOH in germline cells from DLBCL patients that are associated with genes that probably play a key role in DLBCL development.

Keywords Diffuse large B cell lymphoma · CNA · CNN-LOH · GC · Non-GC · Paired samples

Elena Sebastián, Miguel Alcoceba, David Martín-García, Itziar Salaverria, Ramón García-Sanz and Marcos González contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s00277-015-2552-3) contains supplementary material, which is available to authorized users.

✉ Ramón García-Sanz
rgarcias@usal.es

[1] Molecular Biology & Histocompatibility Unit, Department of Hematology, IBSAL – University Hospital of Salamanca, Paseo de San Vicente, 58-182, 37007 Salamanca, Spain

[2] Institute of Biomedical Research of Salamanca (IBSAL), Salamanca, Spain

[3] Spanish Lymphoma/Autologous Bone Marrow Transplant Study Group (GELTAMO), Salamanca, Spain

[4] Hematopathology Unit, Hospital Clínic, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain

[5] Department of Pathology, University Hospital of Salamanca, Salamanca, Spain

[6] Department of Statistics, University of Salamanca, Salamanca, Spain

[7] Department of Pathology, University Hospital of Marqués de Valdecilla/IFIMAV, Santander, Spain

[8] Fundación Pública Galega de Medicina Xenómica, IDIS, SERGAS, Santiago de Compostela, Spain

[9] Grupo de Medicina Xenómica, CIBERER, Universidad de Santiago de Compostela, Santiago de Compostela, Spain

[10] Center for Cancer Research (CIC, IBMCC-USAL-CSIC), Salamanca, Spain

Published online: 14 November 2015

Introduction

Diffuse large B cell lymphoma (DLBCL) is a clinically and biologically heterogeneous disease defined as a diffuse proliferation of large neoplastic mature B cells [1]. Gene expression profiling (GEP) has enabled the characterization of two main biological subgroups of DLBCL, the germinal center B cell-like (GCB), which usually has a favorable prognosis, in contrast to the activated B cell-like (ABC), which is more frequently associated with a poor outcome [2, 3]. However, GEP is not used in standard practice due to its complexity and cost. Immunostaining approaches have been proposed as feasible surrogates for GEP because they allow DLBCL to be classified into GC or non-GC subtypes [4, 5], although discrepant results have been noted [6]. The prognostic value of the immunostaining classification is even more doubtful in the rituximab era because the addition of this drug to conventional chemotherapy has modified the prognostic capacity of already established prognostic factors [7–10].

Genetic abnormalities are other potentially applicable prognostic indicators. Traditional methods have frequently identified *BCL2*, *BCL6*, and *MYC* gene abnormalities in DLBCL, but only the latter has been consistently associated with poor prognosis in these patients [11–13]. This seems to be especially relevant in 20 % of cases with *MYC* breaks that have concurrent *BCL2* and/or *BCL6* abnormalities, the so-called double-hit lymphomas [14, 15].

Next-generation sequencing has identified many mutated genes in DLBCL that are involved in chromatin remodeling (*MLL2* and *EZH2*), immunoresponse (*B2M*), BCR activation (*CD79B*, *MYD88*), B cell differentiation (*BCL6*, *BLIMP1*), and cell survival (*TP53*, *TNFRSF14*) [16–19]. The distribution of these mutations differs between DLBCL subtypes. *TP53* mutations occur in 18–30 % of DLBCLs [17, 19, 20] and have been associated with poor prognosis in DLBCL by several groups [21, 22], particularly in the small fraction of GC lymphomas that harbor such mutations [23]. Other studies, however, have failed to demonstrate any association between *TP53* mutations and prognosis [24, 25]. Furthermore, the *MYD88* L265P mutation, which is found in 12 % of DLBCLs [19, 26], is preferentially associated with the ABC (29 %) subtype rather than the GC subtype (6 %) [17, 27, 28]. However, these frequencies do not help the final distinction between ABC and GC types, so alternatives need to be found.

In recent years, copy number alterations (CNAs) detected by comparative genomic hybridization or high-density SNP arrays have been examined in lymphoproliferative disorders. CNA analysis of DLBCL tumor samples has revealed a number of genomic alterations with frequencies differing between DLBCL subgroups [29, 30]. According to these studies, most ABC-DLBCLs have *BCL2* amplifications as well as *CDKN2A* and 6q21–q22 deletions, whereas GCB-DLBCLs frequently harbor *REL* locus amplifications, *PTEN* deletions, apart from the t(14;18), which is present in around 20 % of the cases. Moreover, ABC-DLBCLs frequently present abnormalities that are characteristically associated with NF-κB pathway activation, similar to primary mediastinal B cell lymphomas [31].

All these findings are challenged because it is difficult to distinguish somatically acquired tumor changes from germinally inherited copy number variations or homozygosity segments. Paired studies with matched tumor and normal DNA samples are superior to those with multiple unpaired samples since they avoid the false discoveries of SNP array analysis [32], including aneuploidy and copy number neutral-loss of heterozygosity (CNN-LOH) that have been detected in cancer-free individuals [33]. High-resolution CN studies of DLBCL have recently been published [30, 34–36]. Unfortunately, studies analyzing paired samples (neoplastic and germinal) are scarce and involve a small number of patients. Moreover, molecular inversion probe (MIP) technology has recently been optimized for copy number analysis of highly degraded formalin-fixed paraffin-embedded (FFPE) samples. The assay performance has been extensively validated with archived FFPE samples [37].

Here, we report a detailed copy number (CN) analysis of paired samples of DLBCL that allowed the proper identification of patient-specific germinal CNAs and CNN-LOH and the somatically acquired tumor lesions. Data were obtained from FFPE and frozen tissue-derived DNA. Based on these findings, we derived a GC/non-GC DLBCL predictor and investigated the relationship between CNAs and clinical features, identifying several alterations that could be associated with poor prognosis patients.

Materials and methods

Patient characteristics

Sixty patients with a diagnosis of de novo DLBCL NOS, according to the 2008 World Health Organization Classification [1], were included in the study. All patients had available diagnostic tumor DNA containing more than 40 % of tumor cells. Paired constitutional DNA was available from 40 of these patients. The median age was 59 years (range 25–85); 57 % were male and 40 % had an International Prognostic Index (IPI) score of ≥3. Immunostaining for BCL-6, CD10, and MUM1 enabled them to be classified into GC DLBCL (42 %) and non-GC (58 %) according to the Hans algorithm [4].

Information about patients' age, gender, and lactate dehydrogenase (LDH), beta-2-microglobulin (B2M), and C-reactive protein (CRP) serum levels, extra nodal involvement, International Prognostic Index (IPI), and age-adjusted IPI (aaIPI) was collected. Most patients were treated with

rituximab-based regimens, and risk-adapted therapies were used (Table 1).

The CN and CNN-LOH changes in the germline of DLBCL patients and in a series of 500 healthy Spanish controls were selected for comparison. The median age of these controls was 67 years (range 33–95) and the male to female ratio was 1.98. These controls were recruited from the database of the Hospital Clínico Universitario de Santiago de Compostela, which has registered healthy individuals who have been examined using the SNP Array 6.0 (Affymetrix, Inc.).

Informed consent was provided by all participants. The study was performed in accordance with the Declaration of Helsinki and Spanish legislation. The study was approved by the local Ethics Review Committee.

Table 1 Clinical and biological characteristics of 60 DLBCL patients at diagnosis

Variable	% of patients
Median follow-up, months (range)	56 (20–211)
Immunostaining	
GC	41.8
Non-GC	58.2
Age (years)	
≤60	55
>60	45
Sex	
Male	56.7
Female	43.3
Performance status	
ECOG <2	69
ECOG ≥2	31
β2-microglobulin >3 mg/L	20
LDH elevated	55.2
PCR elevated	71.4
Extranodal involvement ≥2	23.3
IPI	
0–1	31.6
2	28.1
3	24.6
4–5	15.8
aaIPI	
aaIPI 0–1	49.1
aaIPI 2–3	50.9
Treatment	
Rituximab-based	81.7
Other	18.3

DNA extraction

Tumor DNA was extracted from diagnostic samples. Sufficient tumor cells were tested by morphological and immunophenotypic studies. High molecular weight DNA from fresh samples was isolated using the DNAzol reagent (MRC, Cincinnati, OH, USA). DNA was extracted from FFPE tissue with the RecoverAll™ Total Nucleic Acid Isolation Kit (Ambion Inc., Applied Biosystems, Foster City, CA, USA) or the QuickExtract™ FFPE DNA Extraction Kit (Epicenter Biotechnologies, Madison, WI, USA). DNA from non-tumoral matched samples was isolated from non-infiltrated peripheral blood and non-infiltrated bone marrow in 80 and 20 % of the cases, respectively, using DNAzol or the MagNA Pure LC system (Roche Diagnostics GmbH, Mannheim, Germany).

Detection of significant lymphoma genetic abnormalities

Interphase fluorescence in situ hybridization (FISH) analysis was carried out to detect t(14;18)/*IGH-BCL2* and t(8;14)/*IGH-MYC* translocations, and *BCL6* breaks using specific probes supplied by Abbott Molecular (Des Plaines, IL, USA) following the manufacturer's specifications. The presence of the *IGH-BCL2* rearrangement was also analyzed by polymerase chain reaction (PCR) using standard protocols [38].

Copy number analysis

Sixty DLBCL cases and paired constitutional DNA from 40 of these patients were successfully analyzed using CN arrays. Thirty-eight cases with available FFPE DNA were hybridized on the MIP assay using the Oncoscan FFPE Express custom service *2.0* (Affymetrix, Santa Clara, CA, USA), this version does not interrogate somatic mutations. Twenty-two cases with HMW DNA from frozen tissue and 40 normal DNAs were analyzed using the SNP 6.0 array. CNA and presence of CNN-LOH were evaluated by visual inspection by two different observers using Nexus Biodiscovery version 7 software (Biodiscovery, El Segundo, CA, USA). SNP-FASST2 segmentation algorithm, which is based on hidden Markov model algorithm, was used to mark out each segment. The human reference genome was GRCh37/hg19. Only gains and losses with a minimum size of 100 kb and CNN-LOH more than 5 Mb were considered. CNA and CNN-LOH located in sex chromosomes were not considered. Tumor alterations were considered germinal or acquired by comparing with the corresponding constitutional DNA of each sample. In cases in which normal DNA was not available, alterations were crosschecked with copy number variation databases (http://dgv.tcag.ca/dgv/app/). Physiological deletions of the immunoglobulin heavy and light chain loci (*IGH*@14q32, *IGK*@2p11, and *IGL*@22q11.22) were excluded from the analysis. Results from both arrays were integrated in a single database.

Other studies

TP53 mutations in exons 4–9 were analyzed by standard procedures. PCR products were directly sequenced by the Sanger method with BigDye terminators v3.1 (Applied Biosystems, UK) from both sides [39]. Mutations were crosschecked with the IARC *TP53* reference database (http://p53.iarc.fr/). The *MYD88* L265P mutation was analyzed using real-time allele-specific oligonucleotide PCR (ASO-PCR) as previously described [40].

IGH rearrangements were amplified, identified, and analyzed using the BIOMED-2 protocols [38, 41], HLA class I (-A, -B, and -C) and class II (-DRB1 and -DQB1) typing at low resolution (two digits) was carried out using standardized methods [42]. Phenotype frequencies and HLA polymorphisms were categorized into previously defined supertypes [43, 44].

Statistical analysis

IBM SPSS v.23.0 was used to examine associations between clinical and biological variables by means of Fisher's exact test or non-parametric tests, as appropriate. Survival curves were analyzed by the Kaplan-Meier method and compared using the log-rank test. All statistical tests were two-sided and statistical significance was concluded for values of $p<0.05$. Student's independent samples t test was used to compare the means of the groups. Non-parametric tests were used to examine continuous variables that were not normally distributed. Appropriate cut-off points for CN alterations were calculated using maximally selected rank statistics (*maxstat* package, R, version 3.0.1). Survival analysis was performed on all patients using a two-sided log-rank test to determine the association between variables. Differences were considered to be statistically significant for values of $p<0.05$.

In order to identify those variables independently related to the GC and non-GC groups, a backward stepwise logistic regression, with Akaike information criterion (AIC) as the stop criterion, was performed (*glm* function, R, version 3.0.1).

Results

Genetic abnormalities

The t(14;18)/*IGH-BCL2* translocation was present in 25 % of the cases studied, most of which were from GC DLBCL (47.4 % of GC versus 7.4 % of non-GC DLBCL). There were also *MYC* translocations in 4 % and *BCL6* breaks in 29 % of cases.

Copy number analysis was successful in all 60 cases. No relevant differences were observed between the two CN platforms used (Supplemental Fig. 1 and Supplemental Tables 1 and 2), so the results were subsequently analyzed together. Major comparisons were based on the Hans classifier (32 non-GC versus 23 GC DLBCLs) (Fig. 1, Supplemental Fig. 2 and Supplemental Table 2).

A total of 867 CNAs were identified (mean 14.5±9.6 SD, median of 13 range 0–45). All cases except one had CNAs. There were 412 deletions and 455 gains, but only 63 were recurrent CNAs, including 33 gains and 30 losses. There were also nine recurrent regions of acquired CNN-LOH. We found 59 germinal regions of CNN-LOH. No germinal CNAs were observed. The most frequent minimal regions alterations were gains on 12q15 ($n=26$), 1q32 (n=22), 6p25.3 ($n=17$), 6p21 ($n=16$), and 18q21.3 ($n=16$) and losses on 6q15 ($n=19$), 6q21 ($n=19$), 6q23.3 ($n=18$), 8p23.2 ($n=18$), 1p36.32 ($n=14$), 15q21-22.2 ($n=14$), and 17p13.1 ($n=9$) (Fig. 2, Supplemental Tables 3 and 4). Interestingly, homozygous deletions of 9p21.3 containing *CDKN2A* and *CDKN2B* were detected in three cases, and other homozygous deletions were noted in single cases, such as 6p21 including *HLA class II*, 6q21 (*PRMD1*), 15q21 (*B2m*), 3p14.2 (*FHIT*) and in the region 2q22.3. In addition, 48 amplifications were detected in these series. The most common amplification peak, seen in eight cases (13 %), occurred at 2p15-p16.1 (*REL* and *BCL11A*), and 1p36.22 amplification (*MTHFR*) was present in two cases. Other single amplifications that involve regions of relevance in lymphomagenesis were 12q15 (*MDM2*), 1q32 (*MDM4*), 1q23.3 (*FCRL*), 6p25.3 (*IRF4*), 11q22.3 (*ATM*), 11q24.3 (*ETS1* and *FLI1*), 9p24.1, 5q35.3 (*CANX*), 3q27.3 (*BCL6*), 11q25, 19q13.42, and 9p21.3 (*CDKN2A* and *CDKN2B*).

There were two cases with a copy number pattern suggesting chromothripsis [45–47], one of which carried the previously described aberrant 11q pattern [48] and double *TP53* inactivation (data not shown).

Specific copy number profiles of DLBCL subtypes

CNAs were evaluated according to their GC and non-GC DLBCL classification. Levels of genetic complexity, as measured by the total number of chromosomal imbalances, did not differ significantly between GC-DLBCL (mean 12±7 SD, $n=23$), and non-GC (mean 16±10 SD, $n=32$). (Fig. 1, Supplemental Table 2). Losses of 6q21 and 9p21 were shared features of the two subtypes in our series.

Several chromosomal alterations were significantly associated with one of the DLBCL subgroups. A detailed description of such abnormalities is presented in Table 2, Fig. 1, and Supplemental Fig. 2B.

The logistic regression model predicting the best discrimination between GC and non-GC DLBCL included losses of 1p36.32 and 15q21.1 and gains of chromosome 7 as predictors for the GC subtype, whereas losses

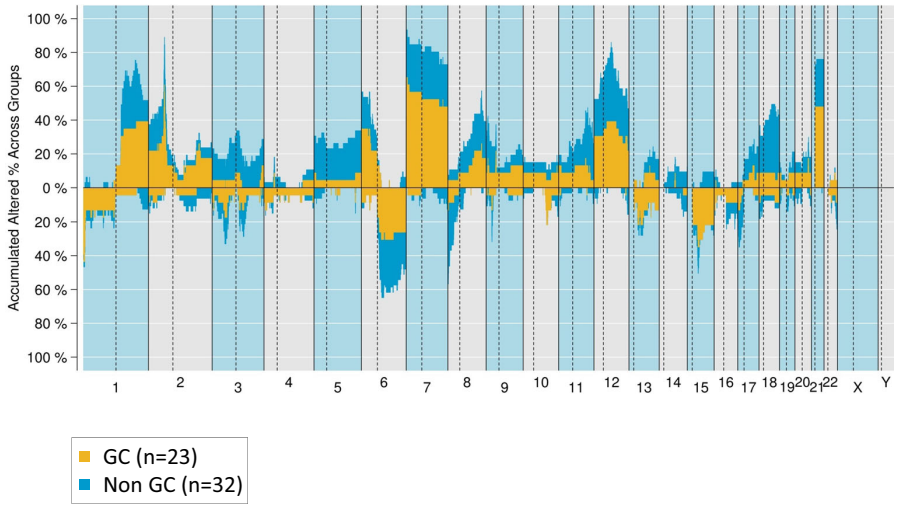


Fig. 1 Comparative CNA plot between GC and Non-GC DLBCL. *X-axis* shows chromosome positions and its respective centromeres marked with dotted lines. *Y-axis* depicts the frequency alterations per each position and per each DLBCL group. GC DLBCL alteration frequencies are colored in *yellow* and non-GC DLBCL in *blue*

of 8p23.3 and gains of 11q24.3 were more specific to the non-GC DLBCL subtype (Supplemental Table 5).

Copy neutral loss of heterozygosity

We found 94 regions with an allelic pattern of copy number neutral loss of heterozygosity (CNN-LOH) in the tumor samples (mean 1.6±1.8 SD; median 1, range 0–8) (Supplemental Fig. 3). The CNN-LOH regions were similarly distributed in DNA samples derived from frozen and FFPE materials. The 6p21.3 region was the most frequent CNN-LOH region in tumors in our series 8/40 (20 %). We also found CNN-LOH regions involving genes of importance in lymphoma, such as *TNFRSF14* and *TP53*. The exact frequencies of the recurrent regions of CNN-LOH are provided in Supplemental Table 6, and the comparison of the subgroups is illustrated in Supplemental Fig. 3C.

The analysis of 40 paired normal-tumor DLBCL samples detected 59 regions of germinal CNN-LOH (mean 1.5±2.8 SD; median 0, range 0–13). Sixteen out of 40

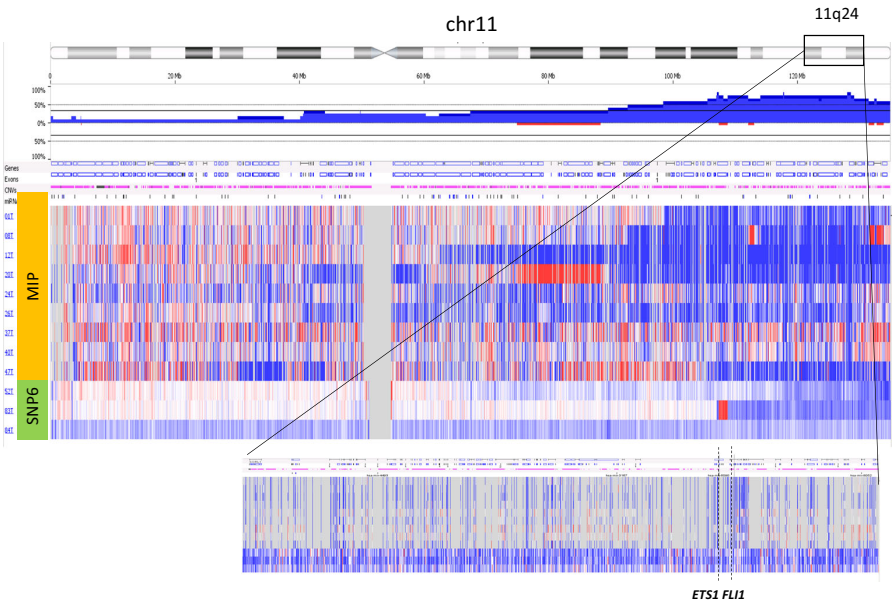


Fig. 2 Cases with gains of 11q24.3 including *ETS1* and FLI1 genes. MIP-labeled cases were samples with DNA from formalin-fixed paraffin-embedded tissue hybridized with Oncoscan. SNP6-labeled cases were DNA samples derived from frozen tissue hybridized by SNP6

Table 2 Differences in CNAs between non-GC and GC DLBCL

Type	Chr	Region	% All DLBCL (n=60)	% GC (n=23)	% non-GC (n=32)	p	Target gene	Publication
Loss	1	1p	6.7	17.4*	0	0.026		
Loss	1	1p36.32	23.3	47.8**	3.1	0	*TNFRSF14*	
Gain	1	1q	21.7	34.8*	12.5	0.048		
Gain	2	2p16.1	36.7	56.5*	28.1	0.034	*REL* and *BCL11A*	[30, 49]
Gain	7	Chr7	31.7	52.2*	21.9	0.025		
Loss	8	8p23.3	30	13	43.8*	0.019	*ERICH1*	
Loss	10	10q23.31	8.3	17.4*	0	0.026	*PTEN*	[49]
Loss	10	10q23.31	11.7	21.7**	0	0.01	*FAS*	[30]
Gain	11	11q24.3	20	4.3	34.4**	0.009	*ETS1* and *FLI1*	
Loss	15	15q	11.7	26.1**	0	0.003		
Loss	15	15q21	23.3	39.1**	2	0.005	*B2M*	
Gain	18	18q21.3	26.7	4.3	40.6**	0.002	*BCL2*	[30, 49]

*p<0.05, **p<0.01

cases had germinal CNN-LOH, although the only recurrent regions were found at 6p21 and 9p24.1 (n=2). Germinal CNN-LOH at 17p13.1 including the *TP53* gene was seen only in one case.

The incidence of the 59 CNN-LOH regions found in 16 normal samples of DLBCL cases was compared with the CNN-LOH regions found in 152 of the 500 normal controls. Thirty CNN-LOH regions were more frequently seen in DLBCL than in healthy controls (p<0.05) (Supplemental Table 7).

Other genetic studies

Seventeen cases (28 %) showed abnormalities at the *TP53* gene, distributed as follows: deletions in ten (17 %), CNN-LOH in three (5 %), and mutations in six (10 %) patients (Supplemental Table 8). There were no cases with concomitant *TP53* deletion and mutation, although there were two cases with concurrent CNN-LOH and *TP53* mutations. Interestingly, these two cases with double *TP53* inactivation were the most genetically complex cases, exhibiting more CNAs than the other tumors (mean 26±19 SD, p<0.01). In contrast, cases with the *TP53* mutation only (n=4) showed no major differences from wild-type cases and were less genetically complex.

The *MYD88* L265P mutations were detected in ten cases (17 %). Mutations were present in both DLBCL subtypes (13 % of GC and 20 % of non-GC DLBCL, p=ns) but cases with mutated *MYD*88 were associated with gains of the *BCL2* locus (p=0.013).

Thirty-five of the 38 cases analyzed carried mutations of the *IGHV* genes, with a cut-off established at 98 %. The most common *IGHV* rearrangements were *IGHV3-23* and *IGHV4-34*, which affected one third of the series. All patients carrying *IGHV4-34* (n=7) were of the non-GC subtype (p=0.017) and had more frequent gains of 8q24, including *MYC* (p=0.02). The three cases with unmutated *IGHV* genes were classified as non-GC.

No associations were seen between HLA specificities and genetic abnormalities in this series.

Clinical correlations

The median follow-up of the series was 56 months (range 20–211), yielding a 5-year progression-free survival (PFS) and overall survival (OS) of 67 and 74 %, respectively, for the whole series. Response to treatment was not associated with any single clinical parameter, including IPI and aaIPI. In addition, PFS and OS were not associated with any clinical or biological parameter, including LDH, β2m, RCP, age, sex, or extra nodal sites. Similarly, we could not detect clinical or biological differences between GC and non-GC immunohistochemical subtypes, with the exception of lung involvement (n=6), which was skewed towards the non-GC DLBCL subtype (p=0.032). Likewise, survival was similar in these two subgroups (5-year PFS 70 versus 71 %, 5-year OS 78 versus 78 %; p>0.05).

As previously reported by our group, patients with the B44 supertype had a worse prognosis as other patients, although the difference was not statistically significant in this series (5-year PFS 50 versus 77 %, 5-year OS 69 versus 83 %; p>0.05). There were no differences in B44 supertype frequencies between the two immunohistochemical subtypes, although the B44 supertype was particularly related to a worse prognosis in the GC subtype (5-year PFS 29 versus 85 %; p=0.016).

Clinical correlations of genetic abnormalities

Some CNAs were found to have a prognostic value in our series (Supplemental Table 9). Patients with 18p11.32 loss

had a lower 5-year PFS (71 versus 33 %, p=0.027). By contrast, 11q gains were related to a better prognosis in this series (5-year PFS 92 versus 61 %, p=0.08; 5-year OS 100 versus 68 %; p=0.04).

Patients with biallelic inactivation of 17p13.1 involving the *TP53* gene demonstrated a very low 5-year PFS (0 %) compared with single allelic inactivation and wild-type cases (79 and 65 %, respectively; p=0.007), which translated into similar numbers for 5-year OS (0 versus 86 versus 73 %, respectively; p=0.014) (Supplemental Fig. 4). Remarkably, no patients with biallelic *TP53* inactivation achieved complete response. Similarly, the two patients with a chromothripsis-like copy number pattern had a very poor prognosis, achieving a median PFS and OS of only 3 and 9 months, respectively.

Losses of 10q23.3 (*FAS*), which are characteristic of the GC group, conferred a worse prognosis on this group (5-year OS=40 versus 89 %, p=0.01), whereas gains of 11q24.3, targeting *ETS1* and *FLI1* genes, and which are more frequently found in non-GC DLBCLs, were associated with a better prognosis within this subtype (5-year OS 100 versus 67 %, p=0.042)

Discussion

DLBCL is the most frequent variant of non-Hodgkin lymphoma, and its biological features have been extensively studied. One of the most important findings was the definition of two subtypes on the basis of the cell of origin by GEP [2]. Moreover, genetic alterations, such as translocations, have been thoroughly analyzed. The copy number profile of DLBCLs was first described using comparative genomic hybridization, but more recently this has been assessed using high-resolution arrays. High-resolution SNP and MIP-assay arrays detect unbalanced gains and losses of genetic material and provide information about CNN-LOH status in frozen and FFPE tissue-derived materials. Our study presents an in-depth CNA analysis of 60 DLBCLs with high-resolution CN arrays. Other studies have been published on this subject [30, 34–36], but most of them did not compare the tumor cells and their normal counterparts in enough patients. Our study provides the first analysis of 40 paired tumor-normal DLBCL cases mainly treated using rituximab. This strategy allows an accurate distinction between germinal and tumor-acquired CNA and CNN-LOH regions.

The CNAs identified in our study show quite good concordance with the high-frequency aberrations described in other studies [30, 36, 50]. Consistent with previously published data, the most frequently observed amplification in our series targeted the 2p16 region, including *REL* and *BCL11A* genes [29, 51]. In addition, other amplifications involving regions of importance in lymphomagenesis were seen in single cases, such as *MDM2*, *IRF4*, *ETS1*, and *FLI1*. In parallel, we also describe homozygous losses in relevant genes such as 6p21.32 (*HLA II*) and 15q21 (*B2M*), which are required for immune response. Losses and mutations of these two genes have been described in DLBCLs of immunoprivileged sites [52, 53]; more recently, inactivation of the *B2M* gene has been reported in DLBCL [17]. In the current study, biallelic inactivation of *TP53* was associated with genomic complexity, but monoallelic *TP53* mutations were not.

When we studied the CN profiles of the two immunohistochemical subtypes [4] we found some specific aberrations in both subgroups. Indeed, we conclude that a few CNAs could reliably stratify patients. Specifically, the predictive model identified losses of 1p36.32 containing *TNFRSF14* and 15q21.1 including *B2M*, as well as gains of chromosome 7, as the most powerful CNAs for classification in the GC subtype. By contrast, 8p23.3 losses including *ERICH1* and 11q24.3gains, containing *ETS1* and *FLI1*, were more specific to the non-GC subtype. To our knowledge, this is the first time that a predictor of DLBCL subgroups has been derived from CN data. This analysis and classification based on CN arrays in combination with an immunohistochemical classifier could have advantages over RNA-based GEP [2], mainly because the latter method requires very high-quality material that is not currently available from FFPE samples.

We also describe here germinal CNN-LOH in patients diagnosed with DLBCL. This has not been reported before, probably because other methods are not sufficiently sensitive to detect them. The most frequent region of germinal CN-LOH was 6p21 (HLA region), which was seen in two cases. We also found one case with germinal CNN-LOH of 17p including *TP53*; this CNN-LOH would have been considered to have been acquired in the tumor if we had not analyzed the normal paired sample. The comparison of our patients with the normal Spanish population demonstrated that 6p22.3-21.33 was significantly more frequent in DLBCL than in control cases. These results strongly suggest that the 6p22.3-21.33 region plays a role in DLBCL development. This is consistent with the idea that the HLA system, which is essential in immunological surveillance, is related to lymphoma susceptibility. Our group found that some HLA polymorphisms seem to influence the development and outcome of DLBCL [42]. In this updated series, the presence of any allele of the HLA-B44 supertype also had a significant influence on PFS and OS. We examined whether patients carrying the HLA-B44 supertype had a higher frequency of CN-LOH or loss in the 6p21 region, but found no association. An alternative explanation for these differences in prognosis could be that cases with the HLA-B44 supertype escape immune surveillance, resulting in poor disease control.

The prognostic analysis of our series did not identify a clinical or immunohistochemical variable that reliably identified high-risk patients. This may have been due to the small sample size of our series, because most patients were treated

with rituximab, and because risk-adapted therapies were used [54, 55]. However, a small subset of very high-risk patients was identified with a few genetic alterations such as biallelic *TP53* inactivation and 18p11.32 losses including the *METTL4* gene. In addition, 11q24.3 gains including the *ETS1* and *FLI1* genes identified a good prognostic subgroup of DLBCL patients, suggesting a tumor suppressor role for these genes. The pathogenic function of *ETS1* in mature B cell lymphomas has not yet been fully elucidated, although some studies suggests that one exists [48, 56]. Finally, losses of *FAS* identified patients with worse prognosis. *FAS* is frequently downregulated in DLBCL, usually due to DNA losses, which probably favors the inhibition of tumor cell apoptosis and contributes to immune escape [57].

In summary, SNP/MIP-assay array analysis could be used as a complementary method to immunohistochemistry for predicting the GC/non-GC subtype of DLBCL and identifying high-risk DLBCL patients. We identify double-allele inactivation of *TP53* and loss of *METTL4* that heavily influence PFS and OS and germline CN-LOH as potential lymphoma susceptibility factors. These findings support the use of this method for evaluating genetic abnormalities associated with DLBCL patients whose ultimate clinical value will be validated in large and independent series.

Acknowledgments We thank Alicia Antón, Montserrat Hernández-Ruano, Rebeca Maldonado, Alejandra Martín, and Ana Díaz (University Hospital of Salamanca, Salamanca, Spain); Eva García-García (Center of Cancer Research, Salamanca, Spain); and Mark Catherwood (Belfast City Hospital, Belfast, UK) for their technical support. We are grateful to José Ramón González-Porras, María Victoria Mateos, and Jesús Martín Sánchez for their intellectual and statistical support. We also thank all members of GELTAMO and the patients who participated in this study.

This work was supported by research funding from the Health Council of Castilla y León (GRS265/A/08), the Health Research Program (PS09/01382), and the Red Temática de Investigación Cooperativa en Cáncer (RTICC) grant RD12/0036 (groups 0069, 0029, 0036, 0058, and 0060) included in the National Plan I+D+I supported by the Instituto Carlos III and the Fondo Europeo de Desarrollo Regional (FEDER), the Spanish Ministry of Economy and Competitiveness, and the European Regional Development Fund (ERDF) "Una manera de hacer Europa" (Innocampus; CEI-2010-1-0010).

ES was supported by CM10/00078-Río Hortega, an ISCIII contract, FEHH grant 2013–2014 and JR14/00025-Juan Rodés, an ISCIII contract. IS was supported by the Subprograma Juan de la Cierva (JCI-2011-10232) and a Miguel Servet contract (CP13/00159).

Authors' contributions ES, MA, LM, MDC, RGS, and MG made the conception and design of the research. ES obtained molecular results, assisted by MA, AB, CJ, MG, NCG, and AA. MES, IP, and RC provided support for the molecular results. ES, IS, and DMG analyzed copy number data. MA provided support for data analysis. MSB performed the statistical analyses, assisted by DMG and ES. GP, LM, and MCC provided statistical support. EGB, EP, and MDC provided patient samples and/or clinical data. AC provided biological data. OB and SMM provided samples and performed the pathology review. ES wrote the first draft of the manuscript. IS reviewed the first version and rewrote it. ES, IS, DMG, MA, LM, RGS, and MG reviewed the final manuscript. MG, RGS, LM, and ES obtained the financial support for the study. MG was the head of the group. IS, RGS, and MG produced the final revision of the manuscript. All the authors read the manuscript and gave the final approval for publication.

Compliance with ethical standards Informed consent was provided by all participants. The study was performed in accordance with the Declaration of Helsinki and Spanish legislation. The study was approved by the local Ethics Review Committee.

Conflict of interests The authors declare that they have no conflict of interest.

References

1. Swerdlow SH, Cancer IA for R on (2008) WHO classification of tumours of haematopoietic and lymphoid tissues: vol. 2: International Agency for Research on Cancer, Edición: 4th ed., 2008. WHO
2. Alizadeh AA, Eisen MB, Davis RE et al (2000) Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature 403:503–511. doi:10.1038/35000501
3. Lenz G, Wright G, Dave SS et al (2008) Stromal gene signatures in large-B-cell lymphomas. N Engl J Med 359:2313–2323. doi:10.1056/NEJMoa0802885
4. Hans CP, Weisenburger DD, Greiner TC et al (2004) Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. Blood 103:275–282. doi:10.1182/blood-2003-05-1545
5. Choi WWL, Weisenburger DD, Greiner TC et al (2009) A new immunostain algorithm classifies diffuse large B-cell lymphoma into molecular subtypes with high accuracy. Clin Cancer Res Off J Am Assoc Cancer Res 15:5494–5502. doi:10.1158/1078-0432.CCR-09-0113
6. Meyer PN, Fu K, Greiner TC et al (2011) Immunohistochemical methods for predicting cell of origin and survival in patients with diffuse large B-cell lymphoma treated with rituximab. J Clin Oncol Off J Am Soc Clin Oncol 29:200–207. doi:10.1200/JCO.2010.30.0368
7. Visco C, Li Y, Xu-Monette ZY et al (2012) Comprehensive gene expression profiling and immunohistochemical studies support application of immunophenotypic algorithm for molecular subtype classification in diffuse large B-cell lymphoma: a report from the International DLBCL Rituximab-CHOP Consortium Program Study. Leukemia 26:2103–2113. doi:10.1038/leu.2012.83
8. Gutiérrez-García G, Cardesa-Salzmann T, Climent F et al (2011) Gene-expression profiling and not immunophenotypic algorithms predicts prognosis in patients with diffuse large B-cell lymphoma treated with immunochemotherapy. Blood 117:4836–4843. doi:10.1182/blood-2010-12-322362
9. Nyman H, Adde M, Karjalainen-Lindsberg M-L et al (2007) Prognostic impact of immunohistochemically defined germinal center phenotype in diffuse large B-cell lymphoma patients treated with immunochemotherapy. Blood 109:4930–4935. doi:10.1182/blood-2006-09-047068
10. Fu K, Weisenburger DD, Choi WWL et al (2008) Addition of rituximab to standard chemotherapy improves the survival of both the germinal center B-cell-like and non-germinal center B-cell-like subtypes of diffuse large B-cell lymphoma. J Clin Oncol Off J Am Soc Clin Oncol 26:4587–4594. doi:10.1200/JCO.2007.15.9277
11. Savage KJ, Johnson NA, Ben-Neriah S et al (2009) MYC gene rearrangements are associated with a poor prognosis in diffuse large

B-cell lymphoma patients treated with R-CHOP chemotherapy. Blood 114:3533–3537. doi:10.1182/blood-2009-05-220095
12. Barrans S, Crouch S, Smith A et al (2010) Rearrangement of MYC is associated with poor prognosis in patients with diffuse large B-cell lymphoma treated in the era of rituximab. J Clin Oncol Off J Am Soc Clin Oncol 28:3360–3365. doi:10.1200/JCO.2009.26.3947
13. Akyurek N, Uner A, Benekli M, Barista I (2012) Prognostic significance of MYC, BCL2, and BCL6 rearrangements in patients with diffuse large B-cell lymphoma treated with cyclophosphamide, doxorubicin, vincristine, and prednisone plus rituximab. Cancer 118:4173–4183. doi:10.1002/cncr.27396
14. Ueda C, Nishikori M, Kitawaki T et al (2004) Coexistent rearrangements of c-MYC, BCL2, and BCL6 genes in a diffuse large B-cell lymphoma. Int J Hematol 79:52–54
15. Hummel M, Bentink S, Berger H et al (2006) A biologic definition of Burkitt's lymphoma from transcriptional and genomic profiling. N Engl J Med 354:2419–2430. doi:10.1056/NEJMoa055351
16. Pasqualucci L, Dominguez-Sola D, Chiarenza A et al (2011) Inactivating mutations of acetyltransferase genes in B-cell lymphoma. Nature 471:189–195. doi:10.1038/nature09730
17. Pasqualucci L, Trifonov V, Fabbri G et al (2011) Analysis of the coding genome of diffuse large B-cell lymphoma. Nat Genet 43:830–837. doi:10.1038/ng.892
18. Lohr JG, Stojanov P, Lawrence MS et al (2012) Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. Proc Natl Acad Sci U S A 109:3879–3884. doi:10.1073/pnas.1121343109
19. Morin RD, Mungall K, Pleasance E et al (2013) Mutational and structural analysis of diffuse large B-cell lymphoma using whole-genome sequencing. Blood 122:1256–1265. doi:10.1182/blood-2013-02-483727
20. Lossos IS, Morgensztern D (2006) Prognostic biomarkers in diffuse large B-cell lymphoma. J Clin Oncol Off J Am Soc Clin Oncol 24:995–1007. doi:10.1200/JCO.2005.02.4786
21. Ichikawa A, Kinoshita T, Watanabe T et al (1997) Mutations of the p53 gene as a prognostic factor in aggressive B-cell lymphoma. N Engl J Med 337:529–534. doi:10.1056/NEJM199708213370804
22. Young KH, Weisenburger DD, Dave BJ et al (2007) Mutations in the DNA-binding codons of TP53, which are associated with decreased expression of TRAILreceptor-2, predict for poor survival in diffuse large B-cell lymphoma. Blood 110:4396–4405. doi:10.1182/blood-2007-02-072082
23. Zainuddin N, Berglund M, Wanders A et al (2009) TP53 mutations predict for poor survival in de novo diffuse large B-cell lymphoma of germinal center subtype. Leuk Res 33:60–66. doi:10.1016/j.leukres.2008.06.022
24. Osada M, Ishioka C, Ichinohasama R et al (1999) Influence of p53 mutation on pathological grade, but not prognosis of non-Hodgkin's lymphoma. Anticancer Drug Des 14:107–114
25. Barrans SL, Carter I, Owen RG et al (2002) Germinal center phenotype and bcl-2 expression combined with the International Prognostic Index improves patient risk stratification in diffuse large B-cell lymphoma. Blood 99:1136–1143
26. Jiménez C, Sebastián E, Chillón MC et al (2013) MYD88 L265P is a marker highly characteristic of, but not restricted to, Waldenström's macroglobulinemia. Leukemia 27:1722–1728. doi:10.1038/leu.2013.62
27. Ngo VN, Young RM, Schmitz R et al (2011) Oncogenically active MYD88 mutations in human lymphoma. Nature 470:115–119. doi:10.1038/nature09671
28. Zhang J, Grubor V, Love CL et al (2013) Genetic heterogeneity of diffuse large B-cell lymphoma. Proc Natl Acad Sci U S A 110:1398–1403. doi:10.1073/pnas.1205299110
29. Bea S, Zettl A, Wright G et al (2005) Diffuse large B-cell lymphoma subgroups have distinct genetic profiles that influence tumor biology and improve gene-expression-based survival prediction. Blood 106:3183–3190. doi:10.1182/blood-2005-04-1399
30. Scholtysik R, Kreuz M, Hummel M et al (2015) Characterization of genomic imbalances in diffuse large B-cell lymphoma by detailed SNP-chip analysis. Int J Cancer 136:1033–1042. doi:10.1002/ijc.29072
31. Lenz G, Staudt LM (2010) Aggressive lymphomas. N Engl J Med 362:1417–1429. doi:10.1056/NEJMra0807082
32. Heinrichs S, Li C, Look AT (2010) SNP array analysis in hematologic malignancies: avoiding false discoveries. Blood 115:4157–4161. doi:10.1182/blood-2009-11-203182
33. Jacobs KB, Yeager M, Zhou W et al (2012) Detectable clonal mosaicism and its relationship to aging and cancer. Nat Genet 44:651–658. doi:10.1038/ng.2270
34. Scandurra M, Mian M, Greiner TC et al (2010) Genomic lesions associated with a different clinical outcome in diffuse large B-cell lymphoma treated with R-CHOP-21. Br J Haematol 151:221–231. doi:10.1111/j.1365-2141.2010.08326.x
35. Green MR, Aya-Bonilla C, Gandhi MK et al (2011) Integrative genomic profiling reveals conserved genetic mechanisms for tumorigenesis in common entities of non-Hodgkin's lymphoma. Genes Chromosomes Cancer 50:313–326. doi:10.1002/gcc.20856
36. Monti S, Chapuy B, Takeyama K et al (2012) Integrative analysis reveals an outcome-associated and targetable pattern of p53 and cell cycle deregulation in diffuse large B cell lymphoma. Cancer Cell 22:359–372. doi:10.1016/j.ccr.2012.07.014
37. Wang Y, Carlton VEH, Karlin-Neumann G et al (2009) High quality copy number and genotype data from FFPE samples using molecular inversion probe (MIP) microarrays. BMC Med Genomics 2:8. doi:10.1186/1755-8794-2-8
38. van Dongen JJM, Langerak AW, Brüggemann M et al (2003) Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936. Leukemia 17:2257–2317. doi:10.1038/sj.leu.2403202
39. Gonzalez D, Martinez P, Wade R et al (2011) Mutational status of the TP53 gene as a predictor of response and survival in patients with chronic lymphocytic leukemia: results from the LRF CLL4 trial. J Clin Oncol Off J Am Soc Clin Oncol 29:2223–2229. doi:10.1200/JCO.2010.32.0838
40. Jiménez C, Chillón MDC, Balanzategui A et al (2014) Detection of MYD88 L265P mutation by real-time allele-specific oligonucleotide polymerase chain reaction. Appl Immunohistochem Mol Morphol AIMM Off Publ Soc Appl Immunohistochem 22:768–773. doi:10.1097/PAI.0000000000000020
41. Sebastián E, Alcoceba M, Balanzategui A et al (2012) Molecular characterization of immunoglobulin gene rearrangements in diffuse large B-cell lymphoma: antigen-driven origin and IGHV4-34 as a particular subgroup of the non-GCB subtype. Am J Pathol 181:1879–1888. doi:10.1016/j.ajpath.2012.07.028
42. Alcoceba M, Sebastián E, Marín L et al (2013) HLA specificities are related to development and prognosis of diffuse large B-cell lymphoma. Blood 122:1448–1454. doi:10.1182/blood-2013-02-483420
43. Sidney J, Peters B, Frahm N et al (2008) HLA class I supertypes: a revised and updated classification. BMC Immunol 9:1. doi:10.1186/1471-2172-9-1
44. Greenbaum J, Sidney J, Chung J et al (2011) Functional classification of class II human leukocyte antigen (HLA) molecules reveals seven different supertypes and a surprising degree of repertoire sharing across supertypes. Immunogenetics 63:325–335. doi:10.1007/s00251-011-0513-0
45. Rausch T, Jones DTW, Zapatka M et al (2012) Genome sequencing of pediatric medulloblastoma links catastrophic DNA

rearrangements with TP53 mutations. Cell 148:59–71. doi:10.1016/j.cell.2011.12.013
46. Edelmann J, Holzmann K, Miller F et al (2012) High-resolution genomic profiling of chronic lymphocytic leukemia reveals new recurrent genomic alterations. Blood 120:4783–4794. doi:10.1182/blood-2012-04-423517
47. Zhang C-Z, Leibowitz ML, Pellman D (2013) Chromothripsis and beyond: rapid genome evolution from complex chromosomal rearrangements. Genes Dev 27:2513–2530. doi:10.1101/gad.229559.113
48. Salaverria I, Martin-Guerrero I, Wagener R et al (2014) A recurrent 11q aberration pattern characterizes a subset of MYC-negative high-grade B-cell lymphomas resembling Burkitt lymphoma. Blood 123:1187–1198. doi:10.1182/blood-2013-06-507996
49. Lenz G, Wright GW, Emre NCT et al (2008) Molecular subtypes of diffuse large B-cell lymphoma arise by distinct genetic pathways. Proc Natl Acad Sci U S A 105:13520–13525. doi:10.1073/pnas.0804295105
50. Kreisel F, Kulkarni S, Kerns RT et al (2011) High resolution array comparative genomic hybridization identifies copy number alterations in diffuse large B-cell lymphoma that predict response to immuno-chemotherapy. Cancer Genet 204:129–137. doi:10.1016/j.cancergen.2010.12.010
51. Beà S, Colomo L, López-Guillermo A et al (2004) Clinicopathologic significance and prognostic value of chromosomal imbalances in diffuse large B-cell lymphomas. J Clin Oncol Off J Am Soc Clin Oncol 22:3498–3506. doi:10.1200/JCO.2004.11.025
52. Jordanova ES, Riemersma SA, Philippo K et al (2003) Beta2-microglobulin aberrations in diffuse large B-cell lymphoma of the testis and the central nervous system. Int J Cancer 103:393–398. doi:10.1002/ijc.10824
53. Booman M, Szuhai K, Rosenwald A et al (2008) Genomic alterations and gene expression in primary diffuse large B-cell lymphomas of immune-privileged sites: the importance of apoptosis and immunomodulatory pathways. J Pathol 216:209–217. doi:10.1002/path.2399
54. Sehn LH (2009) Early detection of patients with poor risk diffuse large B-cell lymphoma. Leuk Lymphoma 50:1744–1747. doi:10.3109/10428190903308064
55. Montes-Moreno S, Batlle A, de Villambrosía SG et al (2014) Risk adapted high-dose and dose-dense therapies modulate the impact of biological classification in diffuse large B-cell lymphoma prognosis. Haematologica 99:e138–141. doi:10.3324/haematol.2014.104976
56. Testoni M, Chung EYL, Priebe V, Bertoni F (2015) The transcription factor ETS1 in lymphomas: friend or foe? Leuk Lymphoma 56:1975–1980. doi:10.3109/10428194.2014.981670
57. Siegel RM, Chan FK, Chun HJ, Lenardo MJ (2000) The multifaceted role of Fas signaling in immune cell homeostasis and autoimmunity. Nat Immunol 1:469–474. doi:10.1038/82712

Supporting Information

Supplementary methods

Statistical analysis

For clinical data collection end-points were assessed on the date of the last patient contact; the most recent follow-up was in June 2014. Progression free survival (PFS) was calculated from diagnosis until lymphoma relapse or progression or death for any cause, according to International Working Group criteria [1]. Overall survival (OS) was calculated from diagnosis until death from any cause. For PFS and OS, surviving patients were censored at last follow-up. The following variables were included in the survival analysis: treatment, sex, B2-microglobulin, CRP, aaIPI and IPI factors, performance status Eastern Cooperative Oncology Group [ECOG] ≥2, high LDH levels, Ann Arbor status, age >60 years, ≥2 extranodal sites, cell of origin according to the Hans classification, CNAs, and other biological variables such as *TP53* and *MYD88* mutation, as well as t(14;18), and *MYC* and *BCL6* translocations, HLA polymorphisms and CNAs.

Supplementary Results

Clinical correlation in rituximab-treated patients

Considering only the rituximab-treated patients (n=49), the PFS and OS were 75.8% and 73.6%, respectively, with non-attained median follow-up. Patients with gain of 1q, even when treated with rituximab, had a higher probability relapse or progression, although this was not statistically significant (*p*=0.089). Losses of 18p11.32 and double inactivation of 17p13.1 involving the

TP53 gene could have influenced the clinical outcome of DLBCL in the rituximab-treated subgroup. However, even then it is statistically significant, there were too few patients to allow firm conclusions to be drawn (Supplemental Figure 7).

Regarding the prognostic value of *TP53* and *MYD88* mutations, in cases treated with rituximab, *MYD88* mutated cases had a higher rate of CR, although this pattern was not statistically significant (p=0.054). The *TP53* mutations also had no clinically prognostic value in our series. However, when a mutation and a CN-LOH (a mechanism by which point mutations in a homozygous state can be identified) are detected, the gene undergoes double inactivation, leading to a poor prognosis in our series.

The HLA 44 supertype clearly distinguished two prognostic subgroups, particularly when we considered only rituximab-treated patients (5-year PFS=92.6% *vs.* 54.5%, p=0.003, median not reached; 5-year OS=100% *vs.* 48.5%, p=0.002). In the rituximab subgroup, the HLA 44 supertype also separated two prognostic groups within the GC subtype (5-year PFS=90.9% *vs*. 33.3%, p=0.015; median PFS=11.6 months *vs.* not reached; 5-year OS=100% *vs*. 50%, p=0.006).

Supplementary Tables

Table S1. Copy number and CNN-LOH fragments DLBCL. (Provided asSuppTableSI.XLS)

Table S2. Numbers of CN and CNN-LOH alterations in 60 DLBCL.

Group	Number of cases	Aberration	Number	Mean	SD	Median	Range
Whole series	60	CNAs	867	14.45	9.596	13	0-45
		Gains	455	7.58	4.928	7	0-19
		Losses	412	6.87	5.682	6	0-29
		CNN-LOH*	94	1.57	1.769	1	0-8
FFPE cases (MIP-assay)	38	CNAs	485	12.76	8.861	11	0-30
		Gains	288	7.58	5.315	6.5	0-18
		Losses	197	5.18	4.311	5	0-17
		CNN-LOH*	53	1.39	1.424	1	0-5
FT cases (SNP-array)	22	CNAs	382	17.36	10.312	15	4-45
		Gains	167	7.59	4.295	7	2-19
		Losses	215	9.77	6.633	8	2-29
		CNN-LOH*	41	1.86	2.253	1	0-8
Non-GC	32	CNAs	511	15.97	9.836	14	2-43
		Gains	279	8.72	5.024	8	2-19
		Losses	232	7.25	5.747	6	0-24
		CNN-LOH*	50	1.56	1.848	1	0-8
GC	23	CNAs	278	12.09	7.19	12	0-28
		Gains	147	6.39	4.39	6	0-18
		Losses	131	5.69	3.92	5	0-14
		CNN-LOH*	29	1.26	1.251	1	0-4

FFPE: DNA derived from formalin fixed paraffin embedded tissue; FT: DNA derived from frozen tissue. CNA: copy number alteration; CNN-LOH: copy number neutral loss of heterozigosity.* CNN-LOH acquired plus non-paired.

Table S3. Most frequent gains in 60 DLBCL.

Alteration	Minimal common region	Genes	%(n=60)	Publication
gain 12q14.3-q15	chr12: 65,877,845-69,293,086	*MDM2*	43.3	[2, 3]
gain 1q32.1	chr1:204,455,722-204,855,305	*MDM4*	36.7	
gain 2p16.1-p15	chr2: 60,632,192-61,349,559	*REL_BCL11A*	36.7	[2–6]
gain 1q23.3	chr1:160,738,006-163,223,426	*FCRL*	33.3	[3]
gain 6p25.3	chr6:0-411,185	*IRF4*	28.3	
gain 18q21.1-q22.1	chr18:47,847,057-65,947,826	*BCL2*	26.7	
gain 6p22.1-p21.31	chr6: 29,651,880-33,569,91830,118,816-30,631,837	*HLA class I and II*	23.3	(Monti et al, 2012)
gain 8q24.21-q24.3	chr8:129,642,808-146,364,022	*MYC*	20	[3]
gain 11q24.3	chr11:128,057,161-128,671,910	*ETS1 and FLI1*	20	
gain 5q33.3-q35.3	chr5:158,899,658-180,915,260	*CANX*	16.7	[3]
gain 9p24.1-p23	chr9:5,738,921-10,597,797		16.7	
gain 5p15.33-p13.2	chr5:0-37,788,389		15	[3]
gain 17q25.1-q25.3	chr17:72,771,242-81,195,210		13.3	[3]
gain 13q31.3	chr13:90,501,517-92,418,413		11.7	[3]
gain 19q13.33-q13.41	chr19:49,435,161-52,460,792	*SPIB*	11.7	
gain 3q27.3	chr3: 186,200,290-187,442,914	*BCL6*	11.7	
gain 3q27.3 -q29	chr3: 187,463,271-196,707,233	*BCL6*	10	
gain 19q13.42-q13.43	chr19:54,844,671-59,128,983	*U2AF2*	10	[3]
gain 11p13-p12	chr11:33,770,644-37,316,013	*LMO2*	8.3	
gain 19p13.3	chr19:0-2,433,319		8.3	[3]
gain 19q13.12-q13.41	chr19:35,988,888-52,933,177	*CD79A*	6.7	
gain 9p24.3-p13.2	chr9:0-36,714,803	*CDKN2A/ B*	6.7	
gain 1p36.33-p36.21	chr1:-11,135,154-12,770,113	*MTHFR*	3.3	
gain 16p13.3-p13.12	chr16:0-14,168,840	*ABCA3*	3.3	[5, 7]

Table S4. Most frequent losses in 60 DLBCL.

Alteration	Region	Minimal common region	Genes	% (n=60)	Publication
loss 6q15	6q	chr6: 90,966,690-104,831,210	*BACH2*	31.7	[6]
loss 6q21	6q	chr6:104,831,211-110,798,487	*PRDM1*	31.7	[3]
loss 6q23.3	6q	chr6:135,438,944-138,754,816	*TNFAPI3, MAP3K5*	30	[3]
loss 8p23.2	8p	chr8:0-1,204,899	*ERICH1*	30	[8]
loss 1p36.32	1p	chr1:2,406,793-5,438,062	*TNFRSF14*	23.3	[3]
loss 15q21-22.2	15q	chr15:41,464,786-45,543,580	*B2m*	23.3	[3]
loss 17p13.1	17p	chr17:4,884,528-8,188,137	*TP53*	15	[3, 5]
loss 3p21.31	3p	chr3:48,758,416-52,656,022	*Several genes (MAPKAPK3)*	15	[3, 7]
loss 22q13.33	22q	chr22:48,546,238-51,304,566	*Several genes (EP300)*	15	[9]
loss 9p21.3	9p	chr9:21,985,066-22,073,157	*CDKN2 B* (only)	15	[3, 6, 7]
loss 1p36.11_13	1p	chr1:12,602,954-28,245,921	*CROCC, RUNX3*	13.3	[3, 7]
loss 1p13.1	1p	chr1:116,365,780-117,273,829	*CD58*	13.3	[3]
loss 9p21.3	9p	chr9:21,914,611-22,258,766	*CDKN2A/B*	13.3	[3, 6, 7]
loss 13q14.2	13q	chr13:48,941,286-49,659,689	*RB1*	13.3	[3]
loss 16q23.1	16q	chr16:78,820,426-79,020,700	*WWOX*	13.3	[10, 11]*
loss 10q23.31	10q	chr10:90,483,623-93,389,828	*FAS (only)*	11.7	[3, 6]
loss 16q12.2	16q	chr16:46,584,098-54,894,720	*RBL2*	11.7	[3]
loss 16q24.1	16q	chr16:79,247,268-90,354,753	*IRF8*	11.7	
loss 3p14.2	3p	chr3:60,206,018-60,833,942	*FHIT*	10	[11]
loss 16p13.3	16p	chr16:3,479,542-3,851,823	*CREBBP*	10	[12]
loss 2q22.3	2q	chr2:144,090,029-152,392,299		8.3	[3]
loss 9p24.1	9p	chr9: 5,568,980-6,630,411		8.3	
loss 10q23.31	10q	chr10:89,648,496-93,389,828	*FAS and PTEN*	8.3	
loss 18p11.32	18p	chr18:0-5,035,713	*METTL4*	8.3	[3]
loss 19p13.3	19p	chr19:6,392,849-6,809,215	*TNFSF7 and TNFSF9*	6.7	[6]
loss 3q29	3q	chr3:196,707,233-198,022,430		6.7	
loss 18q21	18q	chr18:60,793,021-78,077,248	*BCL2*	6.7	
loss 6p21	6p	chr6:31,621,730-32,343,157	*HLA-II*	3.3	
loss 7q32-qter	7q	chr7:103,174,075-148,021,370		3.3	

*Described in AIDS-related lymphomas and primary effusion lymphomas

Table S5. CNA used in backward stepwise logistic regression model

Alteration	Potential target genes	Estimate	Std. Error	z-value	*p*
Loss 1p36.32	*TNFRSF14*	-3.9183	1.5879	2.4680	0.0136*
Gain Chr7		-1.7500	1.0477	1.6700	0.0949
loss 8p23.3	*ERICH1*	4.3834	1.8812	2.3300	0.0198*
Gain 11q24.3	*ETS, FLI1*	3.3032	1.4542	2.2720	0.0231*
loss 15q21.1	*B2M*	-2.2096	1.3848	1.5960	0.1106

Akaike information criteria: 49.217; Significant *p*-value*

Table S6. Recurrent regions of copy number neutral loss of heterozygosity (CNN-LOH)

Aberration	Frequency (n=60)	% germinal (n=40)	% acquired-unpaired (n=60)	% acquired (n=40)
CNN-LOH 6p *HLA*	16/60 (26.6%)	2 (5%)	14 (23.3%)	8/40 (20%)
CNN-LOH 1p36.3 *TNFRSF14*	7/60 (11.6%)	1 (2.5%)	6 (10%)	4/40 (10%)
CNN-LOH 9p24.1	6/60 (10%)	2 (5%)	4 (6.7%)	2/40 (2.5%)
CNN-LOH 17p *TP53*	4/60 (6.7%)	1 (2.5%)	3 (5%)	1/40 (2.5%)
CNN-LOH 16p13.3 *ABCA3*	3/60 (5%)	0 (0%)	3(5%)	2/40 (5%)
CNN-LOH 19p13.3	3/60 (5%)	0 (0%)	3(5%)	0/40 (0%)
CNN-LOH Chr 9	2/60 (3.3%)	0 (0%)	2(3.3%)	0/40 (0%)

Table 7. Comparative regions of germline copy number neutral loss of heterozygosity (CNN-LOH) between DLBCL patients and normal controls

CNN-LOH region	% DLBCL (n=40)	% Normal controls n=500	band
chr6:31391597-33399184	5	0	6p21.33-32
chr7:46576053-47024064	5	0	7p12.33
chr9:6628123-7791959	5	0	9p24.1
chr9:77840983-80192585	5	0	9q21.13-2
chr9:81932948-85435791	5	0	9q21.21-32
chr20:51124769-53414434	5	0	20q13.2
chr20:8653875-10288593	7.5	0.8	20p12.3-2
chr3:53861563-54196953	5	0.2	3p21.1
chr3:182765827-183709054	5	0.2	3q27.1
chr6:23309277-24121584	5	0.2	6p22.3
chr6:31115931-31391596	5	0.2	6p21.33
chr7:45184259-46576052	5	0.2	7p13-12.3
chr8:20302082-21351303	5	0.2	8p21.3
chr8:105413965-111047905	5	0.2	8q22.3-23.2
chr9:5304970-6628122	5	0.2	9p24.1
chr9:7791960-9430480	5	0.2	9p24.1-p23
chr9:75952008-77840982	5	0.2	9q21.13
chr20:50110486-51124768	5	0.2	20q13.2
chr3:53600054-538615623p	5	0.4	3p21.1
chr3:178914084-182765826	5	0.4	3q26.32-27.1
chr3:183709055-184408376	5	0.4	3q27.1
chr6:23298955-23309276	5	0.4	6p22.3
chr6:24121585-31115930	5	0.4	6p22.3-21.33
chr8:12762578-20302081	5	0.4	8p22-21.3
chr9:5038402-5304969	5	0.4	9p24.1
chr9:73758599-75952007	5	0.4	9q21.12-13
chr18:54409269-54860273	5	0.4	18q21.31
chr18:55000635-55182771	5	0.4	18q21.31
chr20:15511612-15570945	5	0.4	20p12.1
chr20:48511053-50110485	5	0.4	20q13.13-2

Table S8. *TP53* mutations detected in 59 cases of DLBCL

Case	MutationType	Exon	CDsmutation	AA mutation	17p13 status
12T	Substitution: missense	7	c.742C>T	p.R248W	wt
23T	Substitution: missense	8	c.817C>T	p.R273C	wt
31T	Substitution: missense	8	c.815T>A	p.V272E	wt
58T	Substitution: missense	8	c.818G>A	p.R273H	CNN-LOH
66T	Substitution: nonsense	6	c.637C>T	p.R213*	CNN-LOH
78T	Substitution: missense	6	c.638G>A	p.R213Q	wt

Table S9. Copy number alterations with prognostic value in DLBCL.

			Global Series (n=60)		Rituximab treated (n=49)		
Alteration	**Genes**	**Total**	**PFS**	**OS**	**PFS**	**OS**	**Comments**
loss 9p24.1	*JAK2*	5/60(8.3%)	**0.041***	0.207	*0.063*	0.375	Bad prognosis
gain 11q24.3	*ETS1 and FLI1*	12/60(20%)	*0.08*	**0.04***	0.311	0.15	Good prognosis
loss 15q21	*B2M*	14/60(23.3%)	0.74	0.366	0.283	**0.04***	Bad prognosis
gain 16p13.3	*ABCA3*	2/60(3.3%)	**0.008****	**0.007****	*0.072*	**0.006****	Bad prognosis
loss 18p11.32	*METTL4*	5/60(8.3%)	**0.003****	**0.037***	**0.000****	**0.002****	Bad prognosis
loss 19p13.3		5/60(8.3%)	*0.065*	**0.029***	0.596	**0.375**	Bad prognosis

**$P < 0.01$
*$P < 0.05$
Underlined and italic numbers: $P < 0.1$

Supporting Figures

Fig S1. Frequency of Copy Number Aberrations (CNAs) plot of FFPE and FT cases. Frequency plots representing in x-axis chromosome position and in y-axis chromosome aberration frequencies colored in blue gains and red losses. (A). Frequency aberration plot of DNA derived from formalin fixed paraffin embedded tissue samples (FFPE). (B). Frequency aberration plot of DNA derived from frozen tissue (FT).

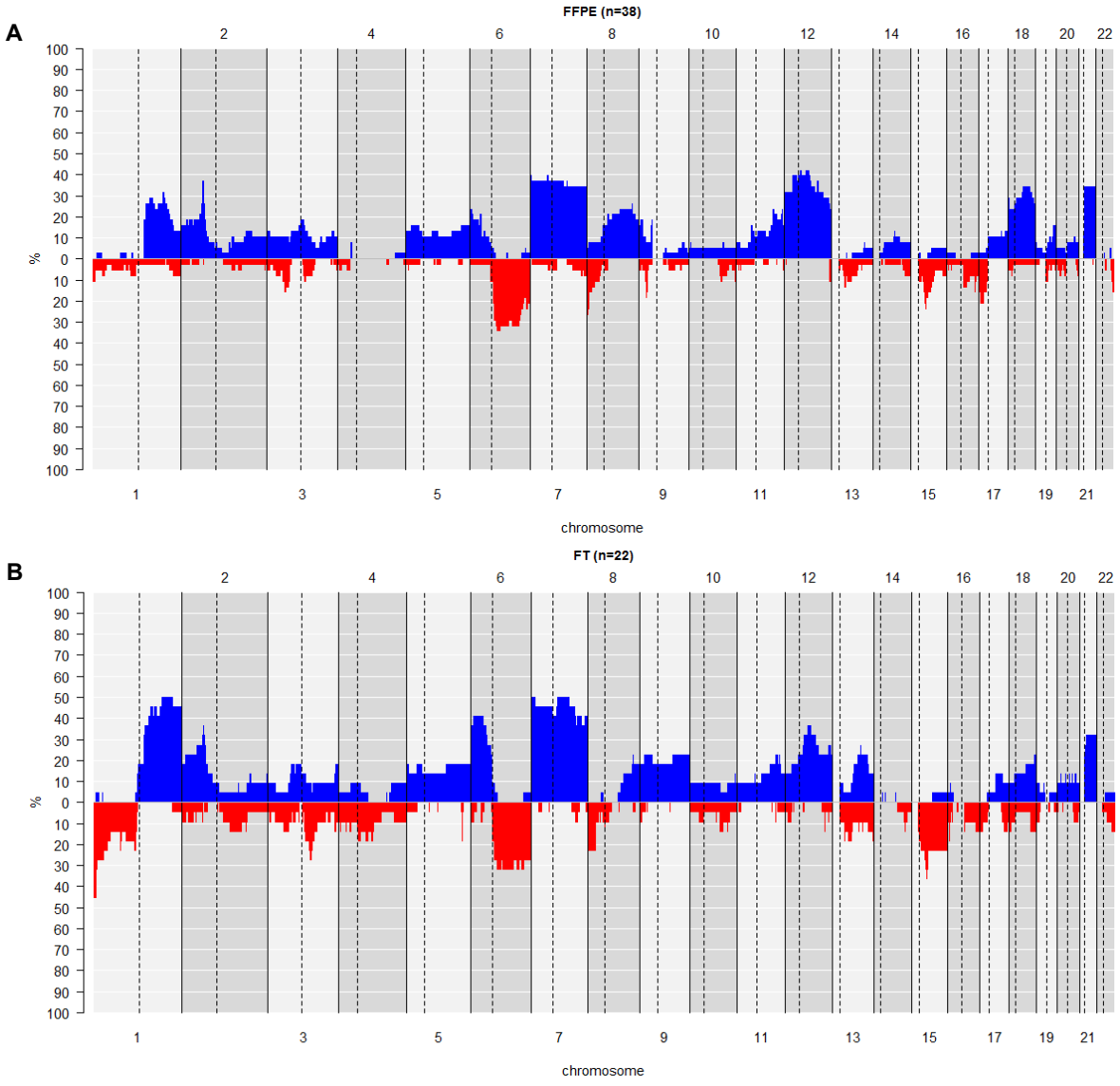


Fig S2. Frequency of Copy Number aberrations (CNAs) of DLBCL. (A) Frequency plots of all DLBCL (n=60). The x-axis depicts position, with dotted lines marking the position of centromeres. On the y-axis, the frequency of aberrations is plotted, colored in blue for gains and red for losses. Frequency plot of all DLBCL cases (n=60). (B) Comparative CNA plot between GC and Non-GC DLBCL (n=55). X-axis depicts chromosome positions with dotted lines pointing centromeres. Y-axis shows the p-value of each chromosome position comparing GC vs non-GC. Cut-off p-value is depicted with red lines (*P*-value < 0.01)

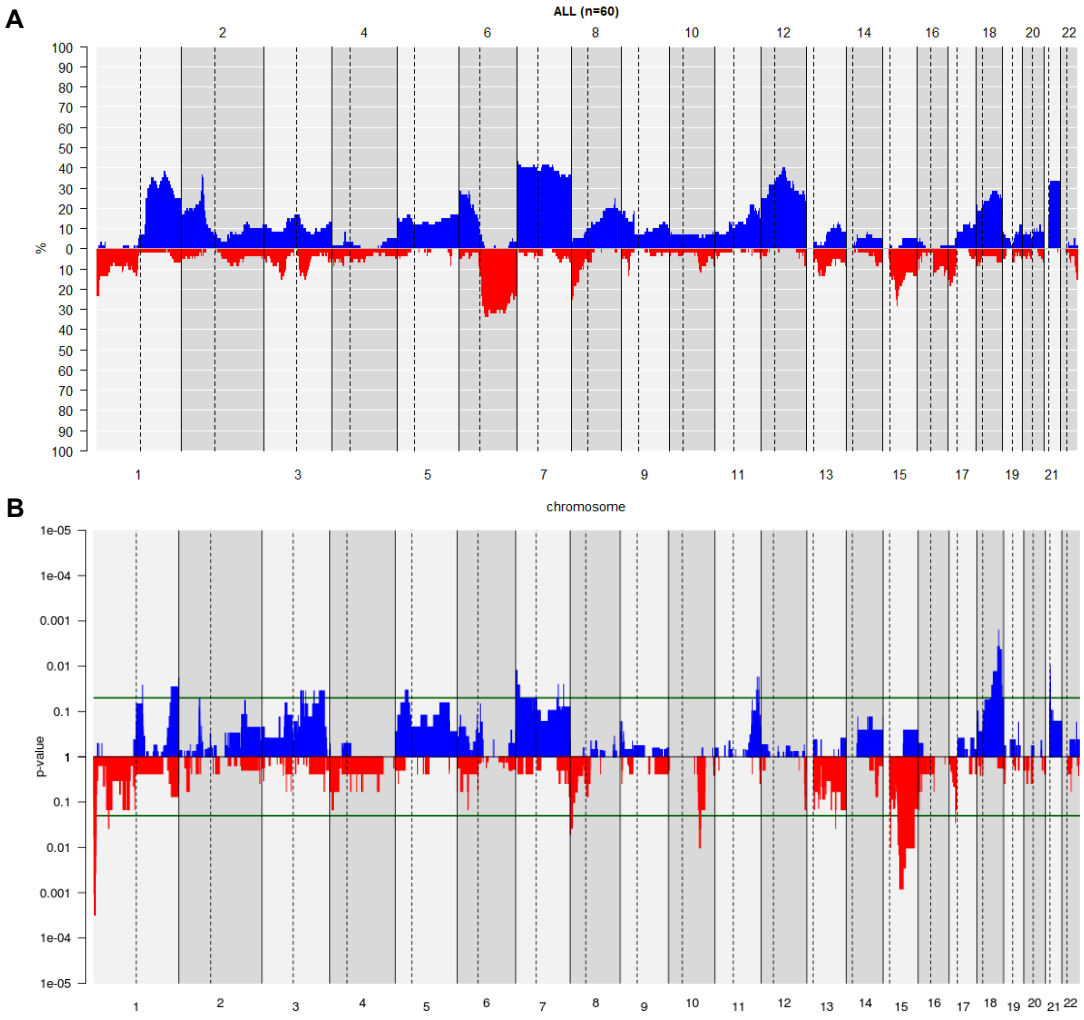


Fig S3. Frequency plot of CNN-LOH. (A) X-axis shows chromosome positions and centromeres are marked with dotted lines. Y-axis plots the frequency of CNN-LOH per each chromosome position. This plot shows the frequency of CNN-LOH regions found in all DLBCL cases (n=60). (B) The plot shows the frequency of CNN-LOH regions found in all DLBCL paired cases (n=40). (C) X-axis depicts chromosome positions where centromere is marked with dotted lines. Y-axis shows frequency alterations per each DLBCL group. GC group is colored in yellow (n=23) and non-GC group colored in blue (n=32)

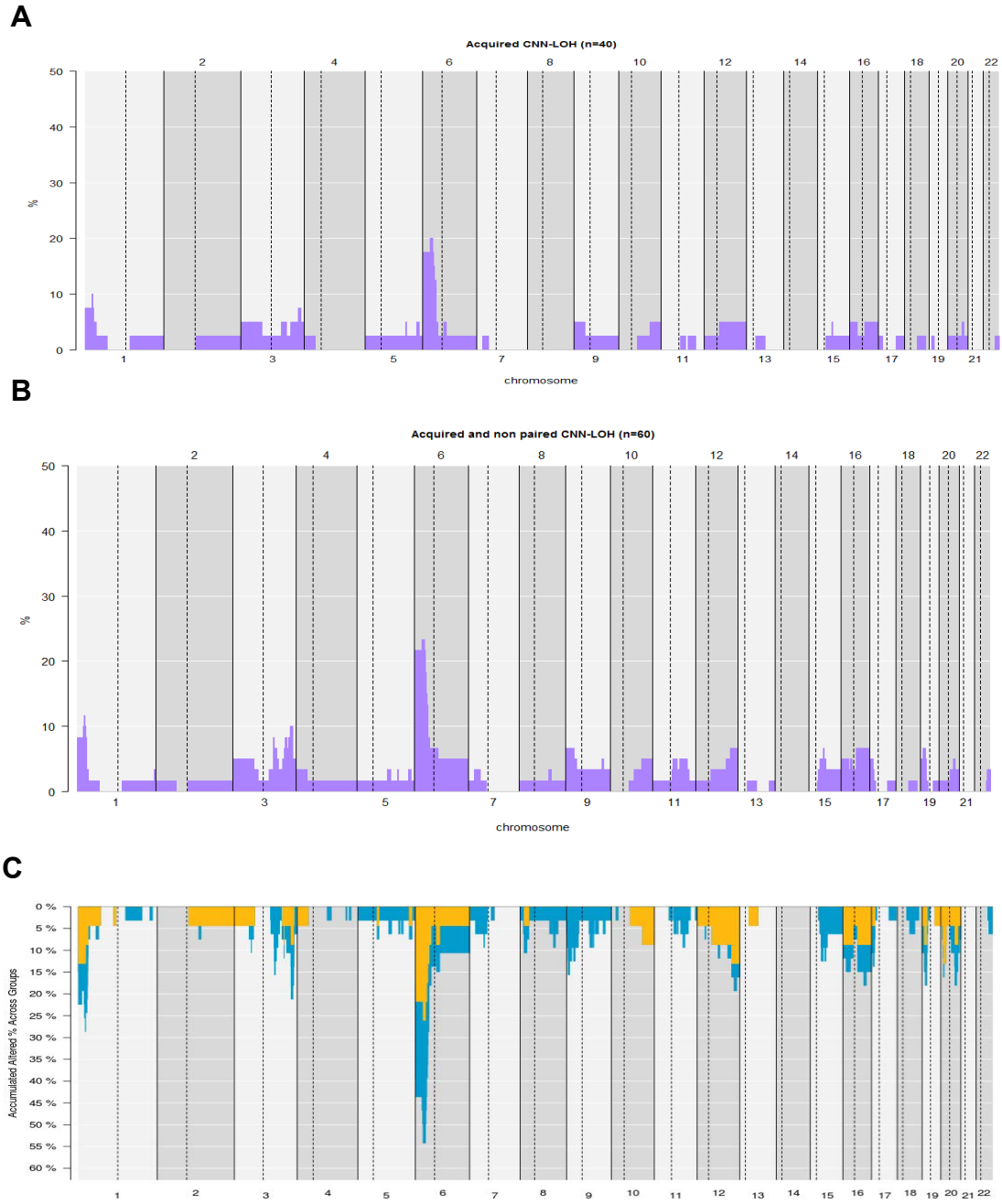
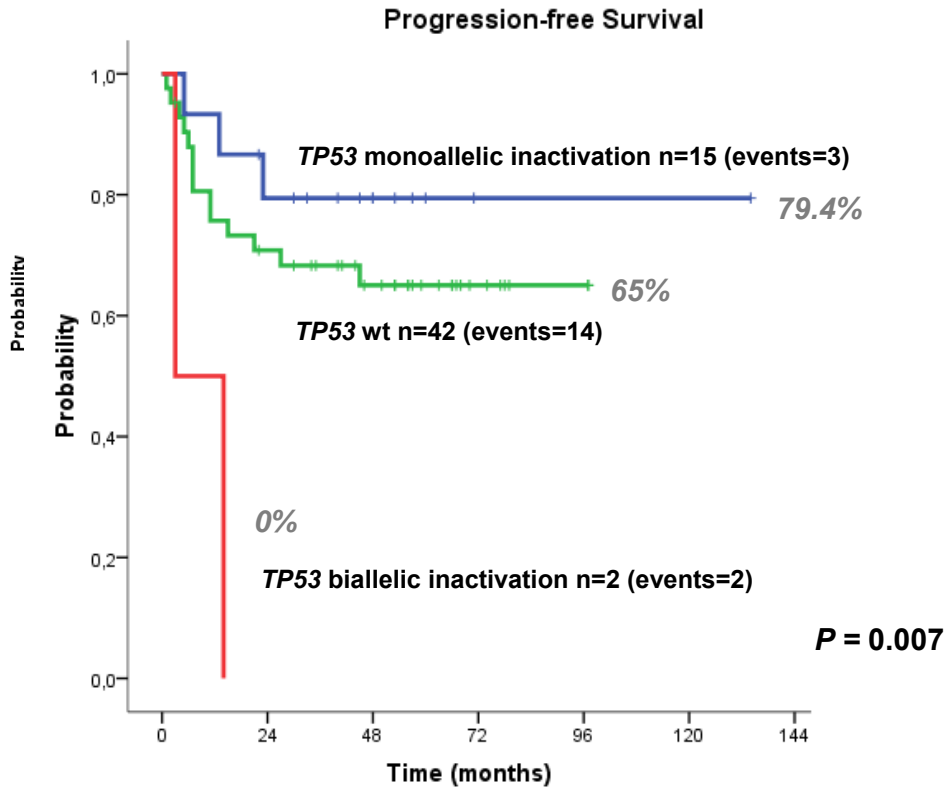
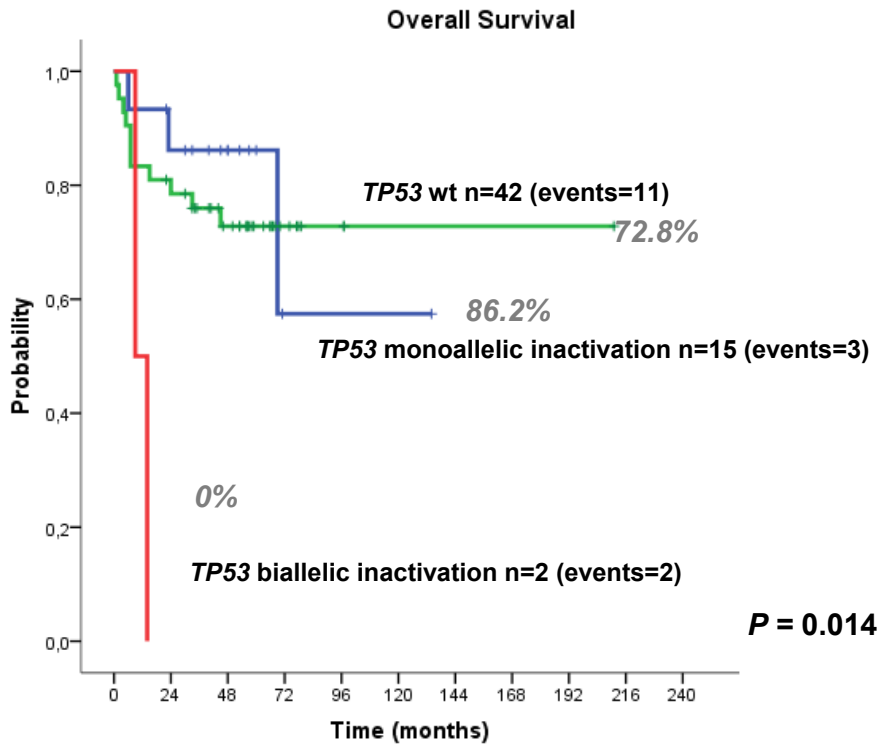


Fig S4. Progression-free survival **(A)** and Overall survival **(B)** in the whole series according to the status of *TP53.*

A





Supporting References

1. Cheson BD, Pfistner B, Juweid ME, et al. (2007) Revised response criteria for malignant lymphoma. J Clin Oncol Off J Am Soc Clin Oncol 25:579–586. doi: 10.1200/JCO.2006.09.2403

2. Bea S, Zettl A, Wright G, et al. (2005) Diffuse large B-cell lymphoma subgroups have distinct genetic profiles that influence tumor biology and improve gene-expression-based survival prediction. Blood 106:3183–3190. doi: 10.1182/blood-2005-04-1399

3. Monti S, Chapuy B, Takeyama K, et al. (2012) Integrative analysis reveals an outcome-associated and targetable pattern of p53 and cell cycle deregulation in diffuse large B cell lymphoma. Cancer Cell 22:359–372. doi: 10.1016/j.ccr.2012.07.014

4. Beà S, Colomo L, López-Guillermo A, et al. (2004) Clinicopathologic significance and prognostic value of chromosomal imbalances in diffuse large B-cell lymphomas. J Clin Oncol Off J Am Soc Clin Oncol 22:3498–3506. doi: 10.1200/JCO.2004.11.025

5. Green MR, Aya-Bonilla C, Gandhi MK, et al. (2011) Integrative genomic profiling reveals conserved genetic mechanisms for tumorigenesis in common entities of non-Hodgkin's lymphoma. Genes Chromosomes Cancer 50:313–326. doi: 10.1002/gcc.20856

6. Scholtysik R, Kreuz M, Hummel M, et al. (2015) Characterization of genomic imbalances in diffuse large B-cell lymphoma by detailed SNP-chip analysis. Int J Cancer J Int Cancer 136:1033–1042. doi: 10.1002/ijc.29072

7. Kreisel F, Kulkarni S, Kerns RT, et al. (2011) High resolution array comparative genomic hybridization identifies copy number alterations in diffuse large B-cell lymphoma that predict response to immuno-chemotherapy. Cancer Genet 204:129–137. doi: 10.1016/j.cancergen.2010.12.010

8. Scandurra M, Mian M, Greiner TC, et al. (2010) Genomic lesions associated with a different clinical outcome in diffuse large B-Cell lymphoma treated with R-CHOP-21. Br J Haematol 151:221–231. doi: 10.1111/j.1365-2141.2010.08326.x

9. Basso K, Dalla-Favera R (2015) Germinal centres and B cell lymphomagenesis. Nat Rev Immunol 15:172–184. doi: 10.1038/nri3814

10. Deffenbacher KE, Iqbal J, Liu Z, et al. (2010) Recurrent chromosomal alterations in molecularly classified AIDS-related lymphomas: an integrated analysis of DNA copy number and gene expression. J Acquir Immune Defic Syndr 1999 54:18–26. doi: 10.1097/QAI.0b013e3181d3d9eb

11. Roy D, Sin S-H, Damania B, Dittmer DP (2011) Tumor suppressor genes FHIT and WWOX are deleted in primary effusion lymphoma (PEL) cell lines. Blood 118:e32–39. doi: 10.1182/blood-2010-12-323659

12. Morin RD, Mungall K, Pleasance E, et al. (2013) Mutational and structural analysis of diffuse large B-cell lymphoma using whole-genome sequencing. Blood 122:1256–1265. doi: 10.1182/blood-2013-02-483727

3. TERCER TRABAJO: LAS ESPECIFICIDADES DEL SISTEMA LEUCOCITARIO HUMANO, HLA, HUMAN LEUKOCYTE ANTIGEN, ESTÁN RELACIONADAS CON EL DESARROLLO Y PRONÓSTICO DEL LINFOMA B DIFUSO DE CÉLULA GRANDE

Introducción y Objetivos

Algunos polimorfismos genéticos se han asociado con el desarrollo y pronóstico de varios tipos de linfoma. Estudios recientes de asociación del genoma completo, han identificado la región 6p21.3, como región de riesgo de susceptibilidad para diferentes linfomas. El sistema leucocitario humano (*human leukocyte antigen*, HLA), localizado en esta región, tiene un importante papel en la respuesta inmune. Sin embargo, hay pocos estudios en el linfoma B difuso de célula grande (LBDCG), y especialmente en la era del rituximab, sobre la relación entre los polimorfismos del HLA y el desarrollo y pronóstico de este tipo de linfoma. El objetivo de este trabajo es evaluar si los polimorfismos del HLA clase I (A, B y C) y clase II (DRB1 y DQB1), están asociados con la incidencia o evolución del LBDCG, focalizándonos en los pacientes tratados con rituximab.

Pacientes y Métodos

Analizamos 250 pacientes diagnosticados de LBDCG, según la Clasificación de la Organización de la Salud 2008. El 68% de los pacientes recibió tratamiento con rituximab, y el 48% estaban incluidos en protocolos del Grupo Español de Linfoma y Trasplante de Médula Ósea (GELTAMO). La población control estaba formada por 1940 donantes sanos de nuestra región geográfica analizados para los loci HLA-A, HLA-B, y HLA-DRB1, y 200 para los loci HLA-C y HLA DQB1.

El DNA fue extraído de muestras de sangre periférica. El HLA de clase I y II fue tipado por baja resolución (2 dígitos), mediante la reacción en cadena de la polimerasa de secuencias específicas con oligonucleótidos reversa (*polymerase chain reaction, reverse sequence-specific oligonucleotide*, PCR-SSOr) y con la tecnología Luminex XYP (Tepnel Lifecodes Corporation, Stamford, CT). Los polimorfismos del HLA fueron clasificados dentro de los supertipos previamente definidos.

Las frecuencias alélicas y el equilibrio de Hardy Weinberg fueron estimados usando el software Arlequín (versión 3.5.1.2). La comparación de las frecuencias alélicas y fenotípicas entre poblaciones se realizó con el test exacto de Fisher usando el programa GraphPad Prism 4.0. Los *P* valores se corrigieron con el test de Bonferroni (*P*c).

El efecto de las variables clínicas y biológicas en la supervivencia libre de progresión (SLP) y en la supervivencia global (SG) a los 5 años, se analizaron por métodos univariantes (prueba bilateral de log-rank) y multivariantes (análisis multivariante de Cox). Las diferencias se consideraron estadísticamente significativas cuando $P<0.05$. Estos análisis se centraron en los pacientes tratados con rituximab, dado que los regímenes con Rituximab son el tratamiento estándar de primera línea en el LBDCG.

Resultados

Asociación del HLA con el desarrollo de LBDCG

Las frecuencias de los polimorfismos de los controles y pacientes fueron consistentes con el equilibrio de Hardy-Weinberg en todos los casos.

Las frecuencias fenotípicas mostraron diferencias significativas en los pacientes con LBDCG comparado con la población control. La frecuencia de DRB1*01 fue significativamente más alta en los pacientes con LBDCG que en la población control (29% vs 19.5%, $P=0.008$, $Pc=0.104$, OR=1.69, 95% CI=1.25-2.28). Además, se observó una menor incidencia del fenotipo HLA-C*03 en los pacientes con LBDCG comparado con los controles (6.4% vs 17.9%, $P=0.005$, Pc=0.007, OR=0.31, 95% CI=0.16-0.62).

HLA y evolución clínica en la serie completa

Con una mediana de seguimiento de 81 meses (rango 1-255 meses), la SLP y la SG, ambas a 5 años, fueron 61% y 78% respectivamente.

Con respecto a los polimorfismos HLA, los pacientes con HLA-B*18 mostraron una menor SLP (35% vs 63% a 5 años, $P=0.002$) y SG (53% vs 82% a 5 años, $P=0.0005$) comparado con el resto de pacientes. La SG en los pacientes con HLA-B*44 fue menor (69% vs 81% a 5 años, $P=0.08$), aunque no estadísticamente significativo.

Así, nuestros datos revelan que el HLA-B*18 y HLA-B*44 parecen influenciar la evolución del LBDCG. De forma interesante, estas dos especificidades pertenecen al supertipo HLA-44. Por ello quisimos analizar el papel del supertipo HLA-B44 en el pronóstico del LBDCG. Los pacientes que llevaban cualquier alelo del supertipo tuvieron una menor SG que aquellos que no llevaban el supertipo (69% vs 86% a 5 años, *P*=0.004).

HLA y evolución clínica en el grupo tratado con rituximab

Con una mediana de seguimiento de 56 meses (rango, 11-136 meses), la SLP y SG a 5 años, de los pacientes tratados con rituximab fueron 65 y 83% respectivamente.

Al igual que en la serie total los pacientes que llevaban el polimorfismo HLA-B*18, tuvo un peor pronóstico con respecto a los pacientes sin esta especificidad.

Cuando se analizó la supervivencia de los pacientes con supertipo HLA-B44 se vio que los pacientes que llevaban algún alelo perteneciente al supertipo presentaban peor SLP y SG. Para determinar si los resultados obtenidos para le HLA-B44 dependían directamente del polimorfismo HLA-B*18, valoramos el efecto del HLA en 3 grupos: grupo a, pacientes que llevaban HLA-B*18 (n=15); grupo b, pacientes que llevaban cualquiera de los alelos del supertipo HLA-B44, excepto el HLA-B*18 (n=51), y el grupo c, pacientes que llevaban “otros alelos”. Los 3 grupos presentaron una diferente SLP (40%, 59% y 71% respectivamente, *P*=0.018) y SG a los 5 años (60%, 74% y 92% respectivamente).

En el análisis multivariante, la presencia de HLA-B*18, la presencia de cualquier alelo del supertipo HLA-B44 y el aaIPI≥2, fueron las únicas 3 variables independientes con una influencia significativa en PFS y SG.

Conclusiones

En este trabajo, analizamos la influencia del HLA, B, C, DRB1, y DQB1 en el desarrollo y pronóstico de 250 pacientes con LBDCG, la serie más larga hasta la fecha. Nuestros resultados muestran que la presencia del alelo HLA-DRB*01 y la ausencia del polimorfismo HLA-C*03, están asociados con la susceptibilidad a desarrollar un LBDCG. Además, describimos la influencia del supertipo HLA-B44, y especialmente del polimor-

fismo HLA-B*18 en la evolución de los pacientes con LBDCG, asociándose, con una peor SLP y SG. Esta influencia, es especialmente evidente en los pacientes tratados con rituximab y es independiente del aaIPI. Así, la presencia del supertipo HLA-B44 caracteriza un grupo de pacientes que no responde a rituximab. Por otra parte, la ausencia del supertipo HLA-B44 es capaz de identifica un subgrupo de muy buen pronóstico.

Aunque, se desconoce el mecanismo por el que este supertipo confiere un pero pronóstico al LBDCG, podemos hipotetizar que las células tumorales con estos polimorfismos escapan más fácilmente al reconocimiento del sistema inmune.

En conclusión, nuestros datos apoyan el papel del HLA, en la región 6p21, en la susceptibilidad y pronóstico del LBDCG. Nuevos ensayos clínicos del GELTAMO en los que se validarán de forma prospectiva estas conclusiones, están en marcha.

blood
2013 122: 1448-1454
doi:10.1182/blood-2013-02-483420 originally published online July 10, 2013

HLA specificities are related to development and prognosis of diffuse large B-cell lymphoma

Miguel Alcoceba, Elena Sebastián, Luis Marín, Ana Balanzategui, M. Eugenia Sarasquete, M. Carmen Chillón, Cristina Jiménez, Noemí Puig, Rocío Corral, Emilia Pardal, Carlos Grande, José Luis Bello, Carmen Albo, Fátima de la Cruz, Carlos Panizo, Alejandro Martín, Eva González-Barca, M. Dolores Caballero, Jesús F. San Miguel, Ramón García-Sanz and Marcos González

Blood (print ISSN 0006-4971, online ISSN 1528-0020), is published weekly by the American Society of Hematology, 2021 L St, NW, Suite 900, Washington DC 20036.

Regular Article

LYMPHOID NEOPLASIA

HLA specificities are related to development and prognosis of diffuse large B-cell lymphoma

Miguel Alcoceba,[1,2] Elena Sebastián,[1,2,3] Luis Marín,[1,2] Ana Balanzategui,[1,2] M. Eugenia Sarasquete,[1,2] M. Carmen Chillón,[1,2] Cristina Jiménez,[1,2] Noemí Puig,[1,2] Rocío Corral,[1,2] Emilia Pardal,[3] Carlos Grande,[3] José Luis Bello,[3] Carmen Albo,[3] Fátima de la Cruz,[3] Carlos Panizo,[3] Alejandro Martín,[1,2,3] Eva González-Barca,[3] M. Dolores Caballero,[1,2,3] Jesús F. San Miguel,[1,2,4] Ramón García-Sanz,[1,2,4] and Marcos González[1,2,4]

[1]Department of Hematology, University Hospital of Salamanca, Salamanca, Spain; [2]Institute of Biomedical Research of Salamanca, Salamanca, Spain; [3]Spanish Lymphoma/Autologous Bone Marrow Transplant Study Group (GELTAMO); and [4]Center for Cancer Research (CIC, USAL/CSIC), Salamanca, Spain

Key Points

- DLBCL patients carrying the HLA-B44 supertype have a worse progression-free and overall survival after R-CHOP-like treatment.
- The HLA-DRB1*01 allele increases the risk of DLBCL development.

Diffuse large B-cell lymphoma (DLBCL) is an aggressive disease influenced by genetic and environmental factors. The role of the HLA system in tumor antigen presentation could be involved in susceptibility and disease control. We analyzed the phenotypic frequencies of HLA-A, HLA-B, HLA-C, HLA-DRB1, and HLA-DQB1 in 250 DLBCLs, comparing them with 1940 healthy individuals. We also evaluated the influence of HLA polymorphisms on survival in those patients treated with curative intention using cyclophosphamide, doxorubicin, vincristine, and prednisolone (CHOP)-like regimen without (n = 64, 26%) or with (n = 153, 61%) rituximab. DLBCL patients have a higher phenotypic frequency of HLA-DRB1*01 (29% vs 19.5%, *P* = .0008, *Pc* = .0104) and a lower frequency of HLA-C*03 (6.4% vs 17.9%, *P* < .0005, *Pc* = .007) compared with healthy individuals. Irrespective of the age-adjusted International Prognostic Index, those patients receiving a CHOP-like plus rituximab regimen and carrying the HLA-B44 supertype had worse 5-year progression-free (54% vs 71%, *P* = .019) and 5-year overall (71% vs 92%, *P* = .001) survival compared with patients without this supertype. Our data suggest that some HLA polymorphisms influence the development and outcome of DLBCL, allowing the identification of an extremely good-risk prognostic subgroup. However, these results are preliminary and need to be validated in order to exclude a possible population effect. (*Blood*. 2013;122(8):1448-1454)

Introduction

Several genetic polymorphisms have been associated with susceptibility or prognosis in various B-cell non-Hodgkin lymphoma (B-NHL) subtypes.[1,2] In recent years, genome-wide association studies have identified 6p21.3 as a risk region for susceptibility of different lymphomas, such as follicular lymphoma[3-5] or Hodgkin lymphoma.[6,7] The HLA system, located in this region, plays a key role in antitumor immune responses and lymphoma-cell apoptosis[8] and so may be essential for neoplasia control. Previous studies have shown a relationship between HLA polymorphisms and susceptibility to certain hematologic malignancies such as chronic lymphocytic leukemia, multiple myeloma, and acute lymphoblastic leukemia.[9-11] However, there is little information about the relationship between HLA polymorphisms and susceptibility to developing B-NHL or their outcomes.[12-16] Focusing on diffuse large B-cell lymphoma (DLBCL), some associations between HLA specificities and this B-NHL subtype have been described, as well as with other genetic polymorphisms. Shorter progression-free survival (PFS) and overall survival (OS) have been observed in DLBCL patients lacking the HLA-DR2 or carrying TNF_{-308A}.[13] In addition, OS is shorter in those patients carrying the C*07-B*08-LTA+$_{252G}$-TNF_{-308A} haplotype[14] or the HLA-C*07:01.[16] However, these studies considered patients treated before the introduction of the anti-CD20 monoclonal antibody, rituximab. Because, at present, almost all patients with DLBCL receive rituximab combinations as first-line therapy, the current influence of HLA polymorphisms on DLBCL susceptibility or outcome remains to be elucidated.

In the present study, we have evaluated whether HLA class I (A, B, and C) and class II (DRB1 and DQB1) polymorphisms are associated with lymphoma incidence or outcome in a large series of DLBCL patients, including 153 cases treated with cyclophosphamide, doxorubicin, vincristine, and prednisolone (CHOP)-like regimens plus rituximab as first-line therapy. We report new findings that provide further evidence of this association, supporting the hypothesis of an important role for the HLA system in DLBCL risk and prognosis, especially when patients are treated in the rituximab era.

Submitted February 7, 2013; accepted July 1, 2013. Prepublished online as *Blood* First Edition paper, July 10, 2013; DOI 10.1182/blood-2013-02-483420.

M.A. and E.S. contributed equally to this study.

The online version of this article contains a data supplement.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

Patients and methods

Patient characteristics and immunohistochemistry

We analyzed 250 consecutive patients diagnosed with de novo DLBCL according to the 2008 World Health Organization classification.[17] The clinical characteristics of the patients are described in Table 1. Their median age was 56 years, ranging from 13 to 87 years. Fifty-two percent of the patients were male. The age-adjusted International Prognostic Index (aaIPI)[18] score was ⩾2 in 58% of the patients. In 116 samples, immunostaining with antibodies to BCL-6, CD10, and MUM1 could be performed, which allowed patients to be classified into 2 biological subtypes according to the Hans criteria[19]: germinal center B-cell–like (45%) and nongerminal center B-cell–like (55%) DLBCL. Complete clinical and therapeutic information data were available for 233 patients, 224 of whom received a treatment with curative intention. Most of the patients received a CHOP or CHOP-like regimen as first-line treatment (97%), 153 of them (68%) with rituximab (see Table 1). A total of 107 of these patients (48%) were included in Spanish Lymphoma/Autologous Bone Marrow Transplant Study Group (GELTAMO) protocols for de novo DLBCL.[20-22]

The control population for HLA-A, HLA-B, and HLA-DRB1 loci consisted of 1940 healthy donor individuals from our geographical region,[23] matched with the cases by gender, age, and ethnicity. In addition, HLA typing was extended to HLA-C and HLA-DQB1 in 200 of these donors.

This study was approved by the local ethics review committee in accordance with Spanish law. Prior informed consent was obtained from all participants in accordance with the Declaration of Helsinki.

Table 1. Clinical-biological characteristics of DLBCL patients at diagnosis (n = 250)

Variable	n (%) of patients
Age, y	
Median (range)	56 (13-87)
≤60	152 (62)
>60	93 (38)
Median (range) follow-up, mo	89 (18-273)
Sex	
Male	129 (52)
Female	120 (48)
Performance status ECOG ≥ 2	76 (34)
β2 microglobulin > 3 mg/L	76 (42)
Albumin < 3.5 g/dL	73 (39)
Elevated LDH	130 (59)
Elevated C-reactive protein	79 (59)
IPI	
0-1	53 (24)
2	55 (25)
3	75 (34)
4-5	39 (17)
Age-adjusted IPI	
0-1	84 (38)
2-3	137 (62)
Immunostaining	
GCB	52 (45)
Non-GCB	64 (55)
Treatment (n = 224)	
Rituximab based	153 (68)
R-CHOP/*R*-COMP	85 (37.5)
R-MegaCHOP	67 (30)
Others with R	1 (0.5)
Without rituximab	71 (32)
CHOP	18 (8)
MegaCHOP	46 (21)
Others	7 (3)

COMP, cyclophosphamide, vincristine, nonpegylated liposomal doxorubicin, and prednisolone; ECOG, Eastern Cooperative Oncology Group; GCB, germinal center B-cell–like; LDH, lactate dehydrogenase.

DNA extraction and HLA typing

DNA from peripheral blood samples was isolated using the DNAzol reagent (MRC, Cincinnati, OH)[24] or MagNA Pure LC system (Roche Diagnostics GmbH, Mannheim, Germany). HLA class I (A, B, and C) and class II (DRB1 and DQB1) typing at low-resolution level (2 digits) was carried out using the polymerase chain reaction reverse sequence-specific oligonucleotide and Luminex XYP technology (Tepnel Lifecodes Corporation, Stamford, CT), according to the standards of the European Federation of Immunogenetics (http://www.efiweb.org). Polymerase chain reaction sequence-specific primer methods (Dynal Biotech, Oslo, Norway) were also used as necessary for HLA-A, HLA-B, HLA-C, HLA-DRB1, and HLA-DQB1 loci. Homozygosis was confirmed by familial studies or by using 2 alternative methods. Many HLA alleles have overlapping structural and functional features through shared peptide-binding pockets and thus can be clustered into distinct functional groups, known as supertypes. In addition to phenotype frequencies, the HLA polymorphisms were categorized into previously defined supertypes.[25,26]

Definitions and statistical analyses

Allele frequencies were estimated by an expectation-maximization algorithm using the Arlequin software package (version 3.5.1.2).[27] The Hardy-Weinberg equilibrium was tested by applying a modified hidden Markov chain with the 100 000 step-length approach and 10 000 dememorization steps, as implemented in the Arlequin program. Allele and phenotype frequencies between populations were compared with the 2-sided Fisher's exact test using GraphPad Prism 4.0 (GraphPad Software, San Diego, CA). The strength of associations was estimated by odds ratios (ORs) and their 95% confidence intervals (CIs) were calculated by the Cornfield methods (values of $P < .05$ were considered statistically significant). P values were Bonferroni corrected (Pc) to take into account multiple testing.

End points were assessed on the date of the last patient contact; the most recent follow-up was in November 2012. The PFS was calculated from diagnosis until lymphoma relapse or progression or death for any cause, according to International Working Group criteria.[28] OS was calculated from diagnosis until death from any cause. For both PFS and OS, surviving patients were censored at last follow-up.

The following variables were included in the survival analysis: treatment, sex, β2 microglobulin, C-reactive protein, aaIPI and International Prognostic Index (IPI) scores and their respective individual factors (performance status Eastern Cooperative Oncology Group ⩾2, high lactate dehydrogenase levels, Ann Arbor status, age >60 years, >1 extranodal sites), cell of origin according Hans' classification, and presence of the various HLA polymorphisms. For this purpose, the 2-sided log-rank test was used to test the univariate association between variables. Survival curves were plotted according to the Kaplan-Meier method using SPSS (version 15.0; SPSS, Chicago, IL). Subsequently, all variables for which there was some indication of a significant association with DLBCL in univariate tests ($P < .1$) were examined in the multivariate analysis using the stepwise Cox regression (enter) model to estimate hazard ratios (HRs) and 95% CIs. Differences were considered to be statistically significant for values of $P < .05$. Survival analysis was performed only in the patients for whom clinical data were available and who were treated with curative intention (n = 224). Because rituximab combinations are the current standard first-line treatment of DLBCL patients, univariate and multivariate analyses were focused on rituximab-treated patients (n = 153).

Results

HLA association with DLBCL development

HLA polymorphism frequencies in DLBCL patients appeared to be in Hardy-Weinberg equilibrium in all cases. The HLA-A, B, and DRB1 allele frequencies of the control population from our geographic region have been described elsewhere.[23] The HLA-C and DQB1

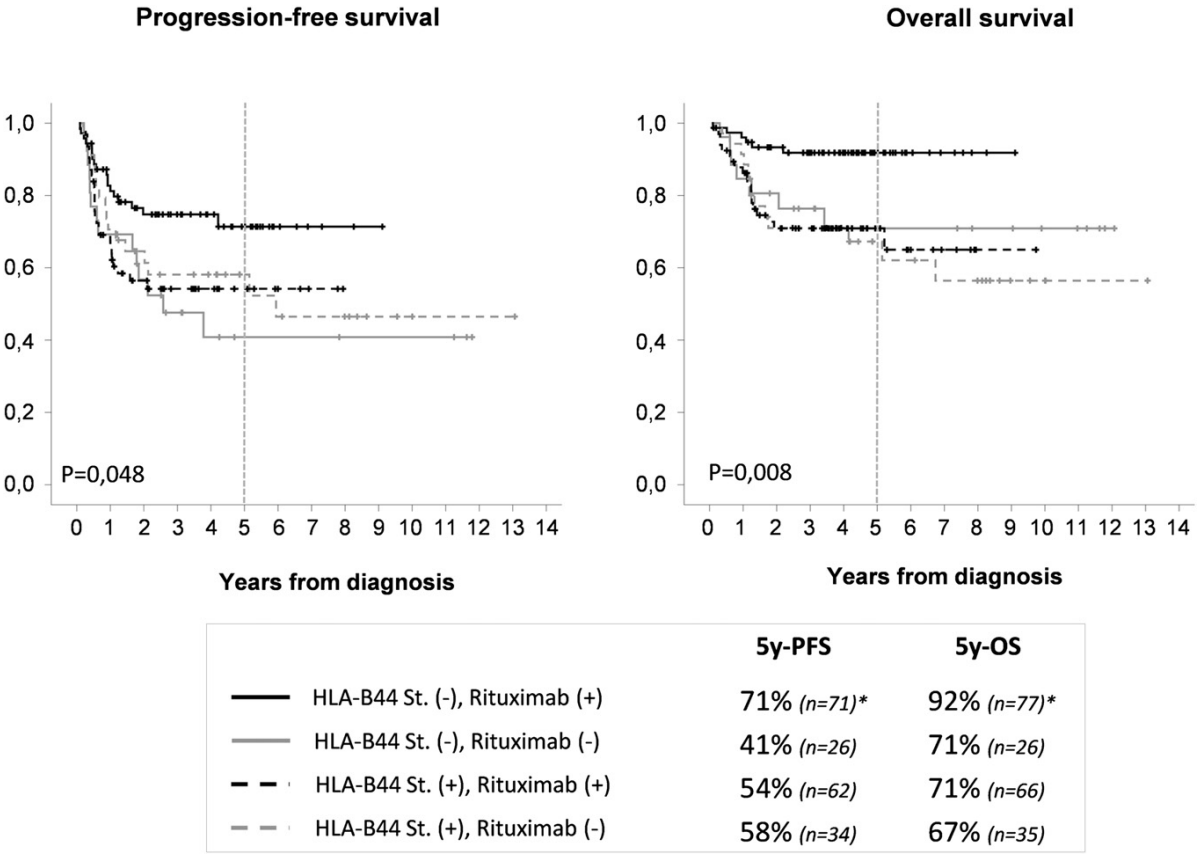


Figure 1. Progression-free survival and OS in the whole series by treatment and HLA-B44 supertype. Asterisk indicates a significant difference compared with other combinations ($P < .05$ for PFS at 5 years and $P < .05$ for OS at 5 years). St, supertype.

polymorphism frequencies were similar to those of other Iberian populations (data not shown).

Phenotypic frequencies differed significantly between DLBCL patients and healthy control individuals (supplemental Table 1). The phenotypic frequency of HLA-DRB1*01 was significantly higher in DLBCL patients than in the control population (29% vs 19.5%, $P = .0008$, OR = 1.69, 95% CI = 1.25-2.28). This difference remained statistically significant even when adjusted for multiple testing ($Pc = .0104$). In addition, we observed a lower incidence of the HLA-C*03 phenotype in DLBCL patients compared with controls (6.4% vs 17.9%, $P = .0005$, $Pc = .007$, OR = 0.31, 95% CI = 0.16-0.62). Finally, HLA-DRB1*15 (13.7% vs 19.8%, $P = .03$) and HLA-B*51 (20.9% vs 15.9%, $P = .07$) polymorphisms slightly differed between DLBCL patients and controls, although statistical significances were lost after Bonferroni correction. No other phenotypic frequencies significantly differed between the patient and control groups.

HLA and clinical outcome in the whole series

With a median follow-up of 81 months (range, 11-255 months), the actuarial 5-year PFS and 5-year OS in the whole series were 61% and 78%, respectively.

With respect to HLA polymorphisms, patients with HLA-B*18 specificity showed a shorter PFS (35% vs 63% at 5 years, $P = .002$) and OS (53% vs 82% at 5 years, $P = .0005$) compared with the other DLBCL patients (supplemental Figure 1). Differences in OS observed for HLA-B*44 (69% vs 81% at 5 years, $P = .08$) and HLA-DRB1*04 (88% vs 75% at 5 years, $P = .07$) were not statistically significant.

Our data reveal that HLA-B*18 and HLA-B*44 seem to influence the clinical outcome of DLBCL. Interestingly, these 2 specificities are members of the HLA-B44 supertype, which includes the HLA-B*18, HLA-B*37, HLA-B*40, HLA-B*41, HLA-B*44, HLA-B*45, and HLA-B*50 alleles.[25,26] Consequently, we wanted to analyze the role of the HLA-B44 supertype in DLBCL prognosis. Patients carrying any HLA allele of the HLA-B44 supertype showed a shorter OS (69% vs 86% at 5 years, $P = .004$) than those without the supertype (supplemental Figure 1).

We analyzed the effect of the treatment considering the HLA phenotype. First, we analyzed the observed frequencies of the HLA phenotypes by treatment or by the inclusion in a clinical trial, and no significant differences between groups were observed (data not shown). We then analyzed the effect of the treatment by HLA-B44 supertype. Patients carrying the HLA-B44 supertype, independent of the treatment with or without rituximab, did not show significant differences in PFS (54% vs 58% at 5 years, $P > .05$) and OS (71% vs 67% at 5 years, $P > .05$; Figure 1). By contrast, those patients carrying other HLA alleles who did not receive rituximab-based regimens showed a shorter PFS (41% vs 71% at 5 years, $P = .013$) and OS (71% vs 92% at 5 years, $P = .012$) than those receiving rituximab (Figure 1).

We checked the potential interaction of HLA phenotypes in PFS and OS considering dose-effect (heterozygosity/homozygosity) of

HLA-B*18 or HLA-B*44 as well as the simultaneous presence of 2 specificities of the HLA-B44 supertype. Our results do not seem to detect any interaction between these specificities in DLBCL outcome.

HLA and clinical outcome in the rituximab group

With a median follow-up of 56 months (range, 11-136 months), the actuarial 5-year PFS and 5-year OS of rituximab-treated patients were 65% and 83%, respectively.

As in the whole series, patients with the HLA-B*18 specificity had a poor outcome: a PFS of 40% vs 66% at 5 years ($P = .018$) and an OS of 60% vs 85% at 5 years ($P = .008$) for patients with and without this specificity, respectively (Figure 2; Table 2). Similarly, the HLA-B*44 specificity was associated with shorter PFS (52% vs 67% at 5 years, $P = .06$) and OS (69% vs 87% at 5 years, $P = .02$). Finally, differences in HLA-DRB1*04 (93% vs 79% at 5 years, $P = .09$) were not statistically significant. We have analyzed different biological and clinical variables in the subgroup of patients bearing the HLA-B*18 allele and the HLA-B44 supertype, including cell of origin and β2 microglobulin. No major differences were found (see supplemental Table 2).

When we analyzed the outcome of the patients by HLA-B44 supertype we found a shorter PFS (54% vs 71% at 5 years, $P = .019$) and OS (71% vs 92% at 5 years, $P = .001$) in those patients carrying any allele of the HLA-B44 supertype (Figure 2; Table 2).

To determine whether the results obtained with the HLA-B44 supertype directly depended on the HLA-B*18 polymorphism, we assessed the effect of HLA in 3 groups: group a, patients bearing the HLA-B*18 polymorphism (n = 15); group b, patients bearing any allele of the HLA-B44 supertype except HLA-B*18 (n = 51); and group c, patients bearing "other HLA alleles" (n = 77). The 3 groups had different 5-year PFS (40%, 59%, and 71%, respectively, $P = .018$) and OS at 5 years (60%, 74%, and 92%, respectively, $P = .0017$), as shown in Figure 2.

We then evaluated the prognostic value of the IPI in our series. Patients with high IPI showed a worse PFS (56% vs 74%, $P = .08$) and OS (88% vs 82%, $P = .2$), although the differences were not statistically significant. Since most of the patients in our series are <60 years, we analyzed the aaIPI. We found that those patients with aaIPI ⩾ 2 showed a worse PFS (60% vs 78%, $P = .018$) and OS (81% vs 91%, $P = .09$).

In the multivariate analysis, the presence of HLA-B*18 ($P = .002$, OR = 4.0; 95% CI = 1.67-9.57), the presence of any allele of the HLA-B44 supertype except HLA-B*18 ($P = .052$, OR = 2.0; 95% CI = 0.99-4.03), and aaIPI ⩾ 2 ($P = .006$, OR = 2.7; 95% CI = 1.32-5.52) were the only independent variables with a significant influence on PFS. The same 3 variables were of independent prognostic value for OS ($P = .001$, OR = 7.6; 95% CI = 2.18-26.66; $P = .017$, OR = 3.7; 95% CI = 1.27-10.97; $P = .049$, OR = 2.8; 95% CI = 1.01-7.74, respectively), as shown in Table 2.

Discussion

The HLA system plays an essential role in immunologic surveillance.[29] The different abilities of certain alleles to efficiently present tumor antigens or the deregulation of allelic expression could result in the tumor cells escaping T-cell and natural killer cell recognition.[8,30-35]

In the present study, we analyzed the influence of the HLA-A, B, C, DRB1, and DQB1 polymorphisms on the development of 250 DLBCL patients, the largest series reported to date. In addition, we describe for the first time the role of HLA polymorphisms in the prognosis of DLBCL patients receiving rituximab-based regimens as first-line treatment. Our results show that the presence of the HLA-DRB1*01 and the absence of HLA-C*03 polymorphism are associated with susceptibility to DLBCL, while the HLA-B44 supertype, especially the presence of HLA-B*18 and HLA-B*44 polymorphisms, have an independent influence on survival.

The HLA system has been related to susceptibility to several hematologic and nonhematologic diseases,[9-11,36] including a variety of HLA polymorphisms associated with B-NHL development and outcome.[12-16] However, this feature has not been extensively studied specifically in DLBCL, although some data suggest an association between HLA subtypes and disease susceptibility (ie, a higher incidence of HLA-DRB1*04:01 allele in this lymphoma subgroup). Moreover, previous studies including survival analyses were limited and performed before the rituximab era, and therefore their conclusions cannot be applied to the current standard of care.[14,16]

Our findings indicate a higher HLA-DRB1*01 frequency in DLBCL patients. Interestingly, a higher incidence of HLA-DRB1*01 has been described in other B-cell lymphoproliferative disorders such as follicular lymphoma[4,15] and smoldering multiple myeloma,[10] indicating a possible role for HLA-DRB1*01 in controlling the tumor B-cell clone. In this way, a higher incidence of this specificity was also found in some autoimmune disorders such as rheumatoid arthritis[37,38] and Crohn disease,[37,39] which have also been linked to higher susceptibility to DLBCL.[40]

We also found a lower frequency of HLA-C*03 in DLBCL. The HLA-C*03 belongs to the HLA-C group 1 (HLA-C1) and is a ligand of the activating natural killer immunoglobulin-like receptor (KIR) KIR2DS2 gene and the inhibitory KIR2DL2 and KIR2DL3 genes.[41] Some studies have related the combination of HLA and KIR genes to susceptibility to or outcome of various infectious and autoimmune disorders.[42-45] For example, homozygosity for both HLA-C1 and KIR2DL3 is related to the spontaneous clearance of hepatitis C virus infection.[45] Considering these data, we hypothesize that those patients carrying HLA-C*03 could develop more efficient antitumor activity by natural killer cells, probably influenced by their combination with KIR genes. Further analysis of the KIR genotypes combined with HLA typing of DLBCL patients should be done to address this question.

One of the important outcomes of our study is the description of the influence of the HLA-B44 supertype (and especially the HLA-B*18 polymorphism) on the prognosis of DLBCL patients. This influence is particularly evident in those patients receiving rituximab-based regimens and is independent of aaIPI, which is an important strength of this study. Although IPI is the standard prognostic index, we used aaIPI because most of the patients in our study are young and most of the elderly patients received an intensified treatment. Furthermore, age and number of extranodal sites (included in the IPI but not in the aaIPI) were not statistically significant in PFS and OS in our series. Moreover, several studies as in the present have previously shown the utility of aaIPI in patients >60 years in the rituximab era.[46,47] It is interesting to note that the presence of the HLA-B44 supertype characterizes a group of patients who do not respond to rituximab in which novel strategies (ie, novel monoclonal antibodies) should be tested. On the other hand, the absence of the HLA-B44 supertype is capable of identifying an extremely good-risk prognostic subgroup that has a cure rate of >90% with current therapeutic strategies. This should be taken into account in the design of future trials with new drugs, because the goal with these patients should be

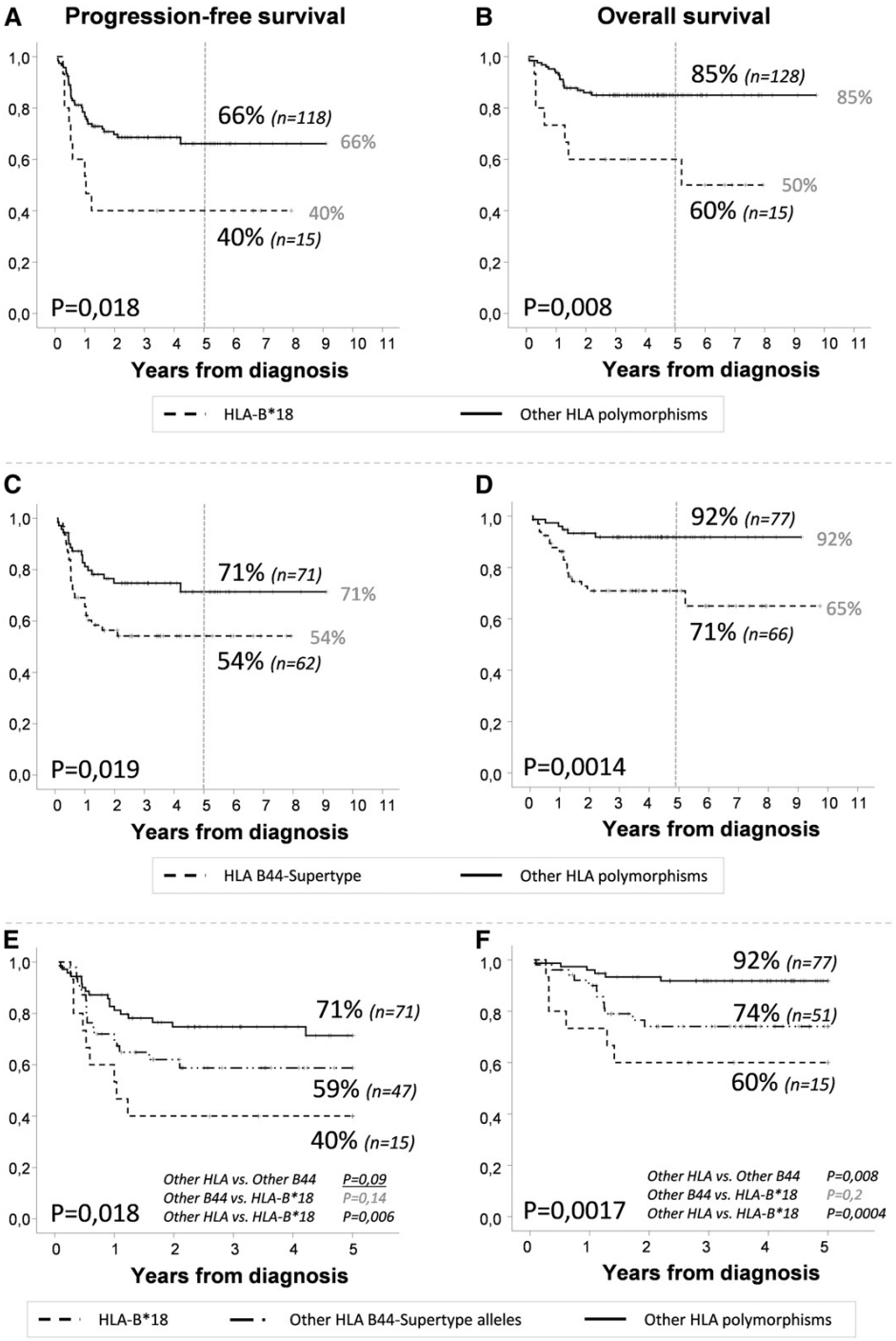


Figure 2. Progression-free survival and OS in patients receiving CHOP or CHOP-like plus rituximab by the HLA. (A,B) Presence of HLA-B*18 vs other polymorphisms. (C,D) Presence of HLA-B44 supertype vs other polymorphisms. (E,F) Presence of HLA-B*18 vs other HLA-B44 supertype polymorphisms vs other HLA polymorphisms.

more focused on toxicity improvement rather than on increasing efficacy.

In chronic myelogenous leukemia patients, the HLA-B*18 polymorphism has been associated with poor response and a trend toward shorter survival.[48] Moreover, HLA-B*44 is more easily lost than other HLA polymorphisms in some types of solid tumor,[49] which indicates that this allele may have a central role in the mechanisms of immune escape. An alternative explanation for these differences in prognosis is that the HLA-B44 supertype does not efficiently present common tumor peptides to the cytolytic

Table 2. Univariate and multivariate analysis of factors influencing PFS and OS in the whole series and in the rituximab-treated group

	PFS (n = 183/127)					OS (n = 193/136)				
Group and variable	**n**	**% 5 y**	**U**	**M**	**OR (95% CI)**	**N**	**% 5 y**	**U**	**M**	**OR (95% CI)**
Whole series										
HLA status										
HLA-B*18	26	35	.002	.002	2.7 (1.4-4.9)	103	53	.0001	.0004	4.4 (1.9-10.2)
Other HLA-B44 supertype alleles	70	64	NS	NS	—	75	76	.064	*.1*	1.9 (0.9-4.2)
Other HLA alleles (reference)	97	63	—	—	—	26	86	—	—	—
Age-adjusted IPI										
<2	78	70	*.064*	*.057*	1.7 (1.0-2.9)	81	86	*.074*	*.072*	2.0 (0.9-4.3)
≥2	125	60				132	77			
Rituximab group										
HLA status										
HLA-B*18	15	40	.006	.002	4.0 (1.7-9.6)	15	60	.0004	.001	7.6 (2.2-26.7)
Other HLA-B44 supertype alleles	47	59	*.09*	*.052*	2.0 (1.0-4.0)	51	74	.008	.017	3.7 (1.3-11.0)
Other HLA alleles (reference)	71	71	—	—	—	77	92	—	—	—
Age-adjusted IPI										
<2	57	78	.018	.006	2.7 (1.3-5.5)	60	91	*.087*	.049	2.8 (1.0-7.7)
≥2	80	60				86	81			

Numbers in italics indicate *P* values >.05.
M, multivariate analysis; NS, not significant; U, univariate analysis.

T lymphocytes, leading to a poor response. Our data indicate that although patients carrying any HLA-B44 supertype allele are not more susceptible to developing DLBCL, tumor cells for patients with these HLA variants may easily escape the immune surveillance, resulting in poor disease control. However, although our study suggests that the HLA typing of patients with DLBCL could be of interest to stratify their risk of progression and survival, it must be taken into account that the present results are preliminary and need further validation. We have thus designed a study to evaluate prospectively HLA-B*18 and the HLA-B44 supertype as part of the trials currently conducted by the GELTAMO Spanish group.

In conclusion, our results suggest a role for HLA in DLBCL susceptibility and prognosis and are consistent with the increasing evidence of the influence of 6p21 on B-NHL development and, in particular, on DLBCL. These findings could be helpful in the future classification, treatment, and monitoring of DLBCL patients.

Acknowledgments

The authors thank Alicia Antón, Felicitación García, Montserrat Hernández-Ruano, Rebeca Maldonado, Mercedes Jiménez, Alejandra Martín, and Isabel Sánchez (University Hospital of Salamanca, Spain) for their technical support. In addition, the authors are grateful to all the members of GELTAMO and the patients who participated in this study.

This work was supported in part by grants from the Health Council of Castilla y León (GRS265/A/08), Red Temática de Investigación Cooperativa en Cáncer (RD12/0036/0069-RTICC, ISCIII), and the Health Research Program (PS09/01382 and CM10/00078-Río Hortega, ISCIII).

Authorship

Contribution: M.A., E.S., L.M., R.G.-S., and M.G. conceived of and designed the research; M.A. and E.S. generated the molecular results, with help from A.B., C.J., and R.C.; M.A. analyzed the data, performed the statistical analyses, and wrote the first draft of the manuscript; M.E.S., M.C.C., and L.M. provided statistical support; N.P., E.P., C.G., J.L.B., C.A., F.d.C., C.P., A.M., E.G.B., and M.D.C. provided patient samples and/or clinical data; E.S., L.M., A.M., R.G.-S., J.F.S.M., and M.G. reviewed the final manuscript; M.G. and L.M. obtained the financial support for the study; J.F.S.M. was the head of the group; and M.G. produced the final revision of the manuscript and gave final approval for the version to be published.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Marcos González Díaz, Molecular Biology Histocompatibility Unit, Department of Hematology, IBSAL University Hospital of Salamanca, Paseo de San Vicente, 58-182, 37007, Salamanca, Spain; e-mail: margondi@usal.es.

References

1. Rothman N, Skibola CF, Wang SS, et al. Genetic variation in TNF and IL10 and risk of non-Hodgkin lymphoma: a report from the InterLymph Consortium. *Lancet Oncol.* 2006;7(1):27-38.
2. Skibola CF, Bracci PM, Nieters A, et al. Tumor necrosis factor (TNF) and lymphotoxin-alpha (LTA) polymorphisms and risk of non-Hodgkin lymphoma in the InterLymph Consortium. *Am J Epidemiol.* 2010;171(3):267-276.
3. Skibola CF, Bracci PM, Halperin E, et al. Genetic variants at 6p21.33 are associated with susceptibility to follicular lymphoma. *Nat Genet.* 2009;41(8):873-875.
4. Conde L, Halperin E, Akers NK, et al. Genome-wide association study of follicular lymphoma identifies a risk locus at 6p21.32. *Nat Genet.* 2010;42(8):661-664.
5. Smedby KE, Foo JN, Skibola CF, et al. GWAS of follicular lymphoma reveals allelic heterogeneity at 6p21.32 and suggests shared genetic susceptibility with diffuse large B-cell lymphoma. *PLoS Genet.* 2011;7(4):e1001378.
6. Cozen W, Li D, Best T, et al. A genome-wide meta-analysis of nodular sclerosing Hodgkin lymphoma identifies risk loci at 6p21.32. *Blood.* 2012;119(2):469-475.
7. Urayama KY, Jarrett RF, Hjalgrim H, et al. Genome-wide association study of classical Hodgkin lymphoma and Epstein-Barr virus status-defined subgroups. *J Natl Cancer Inst.* 2012;104(3):240-253.
8. Chaperot L, Delfau-Larue MH, Jacob MC, et al. Differentiation of antitumor-specific cytotoxic T lymphocytes from autologous tumor

infiltrating lymphocytes in non-Hodgkin's lymphomas. *Exp Hematol.* 1999;27(7):1185-1193.

9. Montes-Ares O, Moya-Quiles MR, Montes-Casado M, et al. Human leucocyte antigen-C in B chronic lymphocytic leukaemia. *Br J Haematol.* 2006;135(4):517-519.
10. Alcoceba M, Marín L, Balanzategui A, et al. The presence of DRB1*01 allele in multiple myeloma patients is associated with an indolent disease. *Tissue Antigens.* 2008;71(6):548-551.
11. Urayama KY, Chokkalingam AP, Metayer C, et al. HLA-DP genetic variation, proxies for early life immune modulation and childhood acute lymphoblastic leukemia risk. *Blood.* 2012;120(15): 3039-3047.
12. Nathalang O, Tatsumi N, Hino M, Prayoonwiwat W, Yamane T, Suwanasophon C, Sriphaisal T. HLA class II polymorphism in Thai patients with non-Hodgkin's lymphoma. *Eur J Immunogenet.* 1999;26(6):389-392.
13. Juszczynski P, Kalinka E, Bienvenu J, et al. Human leukocyte antigens class II and tumor necrosis factor genetic polymorphisms are independent predictors of non-Hodgkin lymphoma outcome. *Blood.* 2002;100(8):3037-3040.
14. Nowak J, Kalinka-Warzocha E, Juszczyński P, et al. Haplotype-specific pattern of association of human major histocompatibility complex with non-Hodgkin's lymphoma outcome. *Tissue Antigens.* 2008;71(1):16-26.
15. Wang SS, Abdou AM, Morton LM, et al. Human leukocyte antigen class I and II alleles in non-Hodgkin lymphoma etiology. *Blood.* 2010; 115(23):4820-4823.
16. Lu Y, Abdou AM, Cerhan JR, et al. Human leukocyte antigen class I and II alleles and overall survival in diffuse large B-cell lymphoma and follicular lymphoma. *ScientificWorldJournal.* 2011; 11:2062-2070.
17. Swerdlow SH, Campo E, Harris NL, et al. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. 4th ed. Lyon, France: IARC Press; 2008.
18. A clinical evaluation of the International Lymphoma Study Group classification of non-Hodgkin's lymphoma. The Non-Hodgkin's Lymphoma Classification Project. *Blood.* 1997; 89(11):3909-3918.
19. Hans CP, Weisenburger DD, Greiner TC, et al. Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. *Blood.* 2004;103(1):275-282.
20. Gonzalez-Barca E, Salar A, Bargay J, et al. Efficacy of dose-dense R-CHOP-14 Chemoimmunotherapy plus pegfilgrastim for first-line treatment of diffuse large B-cell lymphoma (DLBCL) in low risk patients: an interim analysis of an open-label clinical trial in Spain. On behalf of GEL/TAMO (Spanish Group of Lymphoma) [abstract]. *Blood.* 2007;110(11):3443.
21. Arranz R, Conde E, Grande C, et al; Grupo Español de Linfomas/Trasplante Autólogo de Médula Osea (GEL-TAMO). Dose-escalated CHOP and tailored intensification with IFE according to early response and followed by BEAM/autologous stem-cell transplantation in poor-risk aggressive B-cell lymphoma: a prospective study from the GEL-TAMO Study Group. *Eur J Haematol.* 2008;80(3):227-235.
22. Coronado M, Pardal E, Couto R, et al. Prognostic value of interim PET/CT in patients with poor risk diffuse large B-cell lymphoma (DLBCL): experience of the Spanish GELTAMO Group [abstract]. *Blood.* 2012;120(21):5076.
23. Alcoceba M, Marín L, Balanzategui A, et al. Frequency of HLA-A, -B and -DRB1 specificities and haplotypic associations in the population of Castilla y León (northwest-central Spain). *Tissue Antigens.* 2011;78(4):249-255.
24. González M, González D, López-Pérez R, et al. Heteroduplex analysis of VDJ amplified segments from rearranged IgH genes for clonality assessments in B-cell non-Hodgkin's lymphoma. A comparison between different strategies. *Haematologica.* 1999;84(9):779-784.
25. Sidney J, Peters B, Frahm N, Brander C, Sette A. HLA class I supertypes: a revised and updated classification. *BMC Immunol.* 2008;9:1.
26. Greenbaum J, Sidney J, Chung J, Brander C, Peters B, Sette A. Functional classification of class II human leukocyte antigen (HLA) molecules reveals seven different supertypes and a surprising degree of repertoire sharing across supertypes. *Immunogenetics.* 2011;63(6): 325-335.
27. Excoffier L, Lischer HE. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour.* 2010;10(3):564-567.
28. Cheson BD, Pfistner B, Juweid ME, et al; International Harmonization Project on Lymphoma. Revised response criteria for malignant lymphoma. *J Clin Oncol.* 2007;25(5): 579-586.
29. Vesely MD, Kershaw MH, Schreiber RD, Smyth MJ. Natural innate and adaptive immunity to cancer. *Annu Rev Immunol.* 2011;29:235-271.
30. Drénou B, Le Friec G, Bernard M, et al. Major histocompatibility complex abnormalities in non-Hodgkin lymphomas. *Br J Haematol.* 2002;119(2): 417-424.
31. Drénou B, Tilanus M, Semana G, et al. Loss of heterozygosity, a frequent but a non-exclusive mechanism responsible for HLA dysregulation in non-Hodgkin's lymphomas. *Br J Haematol.* 2004; 127(1):40-49.
32. Seliger B, Ritz U, Ferrone S. Molecular mechanisms of HLA class I antigen abnormalities following viral infection and transformation. *Int J Cancer.* 2006;118(1):129-138.
33. Borgerding A, Hasenkamp J, Engelke M, Burkhart N, Trümper L, Wienands J, Glass B. B-lymphoma cells escape rituximab-triggered elimination by NK cells through increased HLA class I expression. *Exp Hematol.* 2010;38(3):213-221.
34. Wilkinson ST, Vanpatten KA, Fernandez DR, et al. Partial plasma cell differentiation as a mechanism of lost major histocompatibility complex class II expression in diffuse large B-cell lymphoma. *Blood.* 2012;119(6):1459-1467.
35. Topalian SL, Drake CG, Pardoll DM. Targeting the PD-1/B7-H1(PD-L1) pathway to activate anti-tumor immunity. *Curr Opin Immunol.* 2012;24(2): 207-212.
36. Ghodke Y, Joshi K, Chopra A, Patwardhan B. HLA and disease. *Eur J Epidemiol.* 2005;20(6): 475-488.
37. Rioux JD, Goyette P, Vyse TJ, et al; International MHC and Autoimmunity Genetics Network. Mapping of multiple susceptibility variants within the MHC region for 7 immune-mediated diseases. *Proc Natl Acad Sci USA.* 2009;106(44): 18680-18685.
38. Holoshitz J. The rheumatoid arthritis HLA-DRB1 shared epitope. *Curr Opin Rheumatol.* 2010; 22(3):293-298.
39. Newman B, Silverberg MS, Gu X, et al. CARD15 and HLA DRB1 alleles influence susceptibility and disease localization in Crohn's disease. *Am J Gastroenterol.* 2004;99(2):306-315.
40. Ekström Smedby K, Vajdic CM, Falster M, et al. Autoimmune disorders and risk of non-Hodgkin lymphoma subtypes: a pooled analysis within the InterLymph Consortium. *Blood.* 2008;111(8): 4029-4038.
41. Colonna M, Brooks EG, Falco M, Ferrara GB, Strominger JL. Generation of allospecific natural killer cells by stimulation across a polymorphism of HLA-C. *Science.* 1993;260(5111):1121-1124.
42. Yen JH, Moore BE, Nakajima T, et al. Major histocompatibility complex class I-recognizing receptors are disease risk genes in rheumatoid arthritis. *J Exp Med.* 2001;193(10):1159-1167.
43. Martin MP, Nelson G, Lee JH, et al. Cutting edge: susceptibility to psoriatic arthritis: influence of activating killer Ig-like receptor genes in the absence of specific HLA-C alleles. *J Immunol.* 2002;169(6):2818-2822.
44. Martin MP, Gao X, Lee JH, et al. Epistatic interaction between KIR3DS1 and HLA-B delays the progression to AIDS. *Nat Genet.* 2002;31(4): 429-434.
45. Khakoo SI, Thio CL, Martin MP, et al. HLA and NK cell inhibitory receptor genes in resolving hepatitis C virus infection. *Science.* 2004;305(5685): 872-874.
46. Feugier P, Van Hoof A, Sebban C, et al. Long-term results of the R-CHOP study in the treatment of elderly patients with diffuse large B-cell lymphoma: a study by the Groupe d'Etude des Lymphomes de l'Adulte. *J Clin Oncol.* 2005; 23(18):4117-4126.
47. Merli F, Luminari S, Rossi G, et al; for the Fondazione Italiana Linfomi. Outcome of frail elderly patients with diffuse large B-cell lymphoma prospectively identified by Comprehensive Geriatric Assessment: results from a study of the Fondazione Italiana Linfomi [published online ahead of print April 30, 2013]. *Leuk Lymphoma.*
48. Cortes J, Fayad L, Kantarjian H, O'Brien S, Lee MS, Talpaz M. Association of HLA phenotype and response to interferon-alpha in patients with chronic myelogenous leukemia. *Leukemia.* 1998; 12(4):455-462.
49. Cabrera T, Maleno I, Lopez-Nevot MA, et al. High frequency of HLA-B44 allelic losses in human solid tumors. *Hum Immunol.* 2003;64(10): 941-950.

Table S1.

Table S2. Clinical-biological characteristics of DLBCL patients according to HLA status in Rituximab-treated patients

Variable	HLA-B44 supertype (n=51) n (% of patients)	HLA-B*18 (n=15) n (% of patients)	Other HLA (n=77) n (% of patients)
Age [years, median (range)]	58 (19-81)	57 (25-76)	55 (15-83)
≤ 60	20 (61)	10 (67)	25 (67)
> 60	31 (39)	5 (33)	52 (33)
Sex			
Male	29 (57)	7 (47)	38 (49)
Female	22 (43)	8 (53)	39 (51)
Performance status ECOG ≥2	13 (27)	5 (36)	23 (32)
β2-microglobulin >3mg/l	12 (33)	4 (36)	23 (39)
Albumin <3.5g/dl	10 (26)	8 (73)*	21 (37)
Elevated LDH	25 (52)	6 (43)	42 (59)
International Prognostic Index (IPI)			
IPI 0-1	16 (33)	3 (21.5)	15 (20)
IPI 2	10 (20)	6 (43)	19 (26)
IPI 3	16 (33)	3 (21.5)	29 (39)
IPI 4-5	7 (14)	2 (14)	11 (15)
Age-adjusted IPI			
aaIPI 0-1	25 (51)	6 (43)	25 (35)
aaIPI 2-3	24 (49)	8 (57)	48 (66)
Immunostaining			
GCB	15 (54)	5 (42)	12 (27)
Non-GCB	13 (46)	7 (58)	33 (73)
Treatment (n=153)			
R-CHOP / R-COMP	8 (16)	2 (13)	13 (17)
R-MegaCHOP	14 (27)	3 (20)	27 (35)
Others with R	29 (47)	10 (67)	37 (48)

* *P*-value <0.05 as compared to any of the other groups

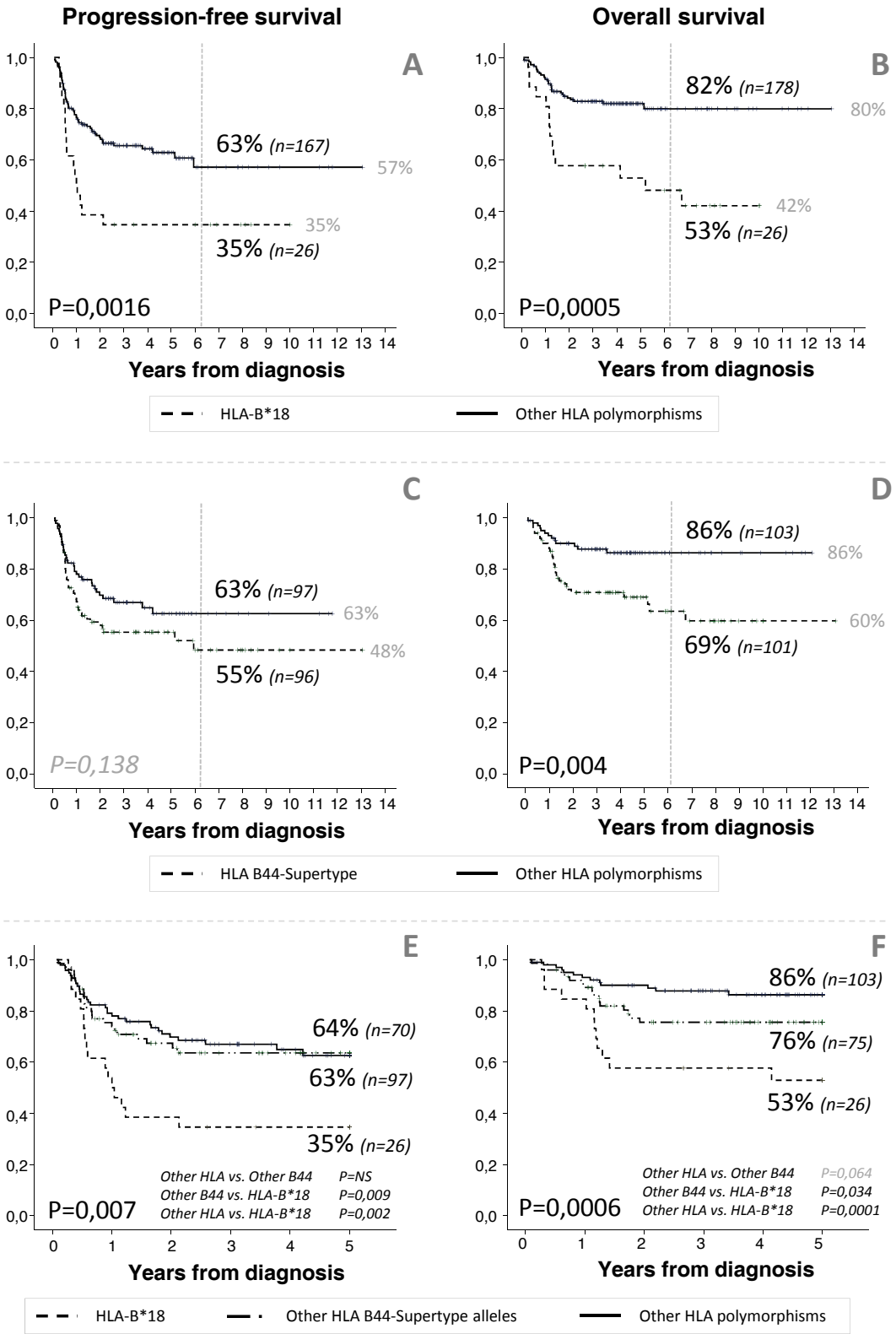
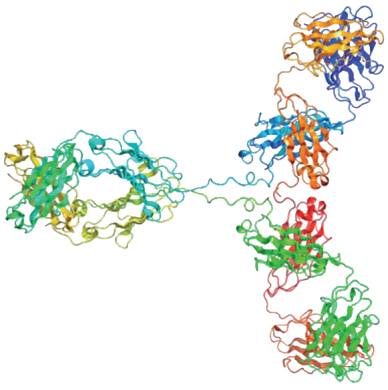


Figure S1. Progression-free survival and overall survival in the whole series according to the HLA: A and B, presence of HLA-B*18 vs. other polymorphisms; C and D, presence of HLA B44-supertype vs. other HLA polymorphisms; E and F, presence of HLA-B*18 vs. other HLA B44-supertype polymorphisms vs. other HLA polymorphisms.

Conclusiones



PRIMER TRABAJO

1. Los linfomas B difusos de célula grande usan genes de la región variable de la cadena pesada de las inmunoglobulinas con características propias que les diferencian de otros síndromes linfoproliferativos B, tanto en el tipo (familias y genes VH) como en sus características: proceso canónico de hipermutación somática, grado de mutación, cambios recurrentes en ciertos aminoácidos y longitud de la región clonotípica CDR3). Estos hallazgos indican que, al menos en un subgrupo de linfomas B difusos de célula grande, los genes VH tienen relevancia en su génesis y desarrollo, posiblemente por un mecanismo de selección antigénica.

2. En los linfomas B difusos de célula grande hay secuencias estereotipadas, al igual que en otros síndromes linfoproliferativos B, aunque con menor incidencia. Esta es la primera vez que se describe este hallazgo, que sugiere que al menos en algunos casos, haya antígenos específicos cuyo estímulo puede ser un posible mecanismo de desarrollo para estos linfomas.

3. Identificamos algunos linfomas B difusos de célula grande en los que el uso de un gen VH concreto se correlaciona con las características clínico-biológicas e histológicas. Así, las formas no mutadas y los linfomas que usan el gen *IGHV34-34* se asocian con el subtipo histológico no centro germinal. Estos hallazgos sugieren que dentro de los subtipos de estos linfomas existen subgrupos particulares que deben ser validados en series independientes donde se confirmen sus características diferenciales.

SEGUNDO TRABAJO

1. El estudio del número de copias génicas identifica ganancias y pérdidas recurrentes y puede ayudar a comprender la patogenia y comportamiento del linfoma B difuso de célula grande. No obstante, es imprescindible analizar simultáneamente muestra tumoral y normal del mismo paciente, para evitar la confusión entre alteraciones y polimorfismos, e identificar pérdidas de heterocigosidad reales, en el que el número de copias es normal en la muestra no tumoral. Nuestro trabajo por tanto, cuestiona los estudios llevados a cabo sin análisis de muestras no tumorales en paralelo.

2. El estudio de anomalías en el número de copias ayuda a clasificar los linfomas B difusos de célula grande en subtipos histológicos. La pérdida de *β2M* y la ganancia del cromosoma 7 son las alteraciones más específicas del subtipo centro germinal, mientras que la pérdida del gen *ERICH1*, y la ganancia de *ETS1* y *FLI1* son características del subtipo no centro germinal.

3. La región 6p21, que contiene los genes del sistema HLA, es una región diana para pérdida de heterocigosidad en muestras tumorales de linfoma B difuso de célula grande. Ello sugiere que el sistema HLA juega un importante papel en la patogenia y desarrollo de estos linfomas por un mecanismo de escape inmune.

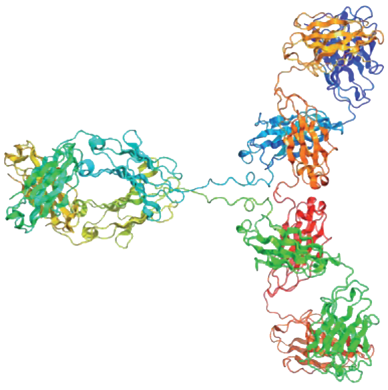
4. Nuestro estudio muestra que el análisis de alteraciones numéricas permite refinar el pronóstico de los linfomas B difusos de célula grande. Así, tanto la doble inactivación del gen *TP53* (pérdida de heterocigosidad y mutación), como la pérdida de *METTL4* son factores pronósticos adversos en estos linfomas; por el contrario, la ganancia de *ETS1 y FLI1,* se relaciona con un mejor pronóstico.

TERCER TRABAJO

1. El sistema HLA, que tiene un papel esencial en los mecanismos antitumorales del sistema inmune, puede jugar un papel relevante en la patogenia del linfoma B difuso de célula grande, ya que la frecuencia de algunas especificidades HLA son significativamente distintas entre los pacientes con este tipo de linfoma y la población sana. Así, el antígeno HLA-DRB1*01 es más frecuente en los pacientes con linfoma B difuso de célula grande, mientras que el antígeno HLA-C*03 es más frecuente en la población normal sana.

2. En los pacientes con linfoma B difuso de célula grande tratados con poliquimioterapia de tipo CHOP con Rituximab, la presencia del supertipo HLA-B44, y en especial con especificidad HLA-B*18, tienen una menor supervivencia libre de progresión y peor supervivencia global. Este hallazgo permite plantear la hipótesis de que las células tumorales con estas especificidades escapan al reconocimiento inmune con mayor facilidad.

Referencias



1. Swerdlow SH, Cancer IA for R on. WHO classification of tumours of haematopoietic and lymphoid tissues: Vol. 2: International Agency for Research on Cancer. WHO; 2008.

2. Salaverria I, Philipp C, Oschlies I, et al. Translocations activating IRF4 identify a subtype of germinal center-derived B-cell lymphoma affecting predominantly children and young adults. *Blood*. 2011;118(1):139-147.

3. Klapper W, Kreuz M, Kohler CW, et al. Patient age at diagnosis is associated with the molecular characteristics of diffuse large B-cell lymphoma. *Blood*. 2012;119(8):1882-1887.

4. Montes-Moreno S, Odqvist L, Diaz-Perez JA, et al. EBV-positive diffuse large B-cell lymphoma of the elderly is an aggressive post-germinal center B-cell neoplasm characterized by prominent nuclear factor-kB activation. *Mod. Pathol. Off. J. U. S. Can. Acad. Pathol. Inc*. 2012;25(7):968-982.

5. Kato H, Karube K, Yamamoto K, et al. Gene expression profiling of Epstein-Barr virus-positive diffuse large B-cell lymphoma of the elderly reveals alterations of characteristic oncogenetic pathways. *Cancer Sci*. 2014;105(5):537-544.

6. Rossi D, Gaidano G. Richter syndrome: molecular insights and clinical perspectives. *Hematol. Oncol*. 2009;27(1):1-10.

7. Chigrinova E, Rinaldi A, Kwee I, et al. Two main genetic pathways lead to the transformation of chronic lymphocytic leukemia to Richter syndrome. *Blood*. 2013;122(15):2673-2682.

8. Martinez-Climent JA, Alizadeh AA, Segraves R, et al. Transformation of follicular lymphoma to diffuse large cell lymphoma is associated with a heterogeneous set of DNA copy number and gene expression alterations. *Blood*. 2003;101(8):3109-3117.

9. Pasqualucci L, Trifonov V, Fabbri G, et al. Analysis of the coding genome of diffuse large B-cell lymphoma. *Nat. Genet*. 2011;43(9):830-837.

10. Tiirikainen MI, Mullaney BP, Holly EA, Pallavicini MG, Jensen RH. DNA copy number alterations in HIV-positive and HIV-negative patients with diffuse large-cell lymphomas. *J. Acquir. Immune Defic. Syndr. 1999*. 2001;27(3):272-276.

11. Carbone A, Gloghini A, Capello D, Gaidano G. Genetic pathways and histogenetic models of AIDS-related lymphomas. *Eur. J. Cancer Oxf. Engl. 1990*. 2001;37(10):1270-1275.

12. Gucalp A, Noy A. Spectrum of HIV lymphoma 2009. *Curr. Opin. Hematol*. 2010;17 (4):362-367.

13. Park S, Lee J, Ko YH, et al. The impact of Epstein-Barr virus status on clinical outcome in diffuse large B-cell lymphoma. *Blood*. 2007;110(3):972-978.

14. Engels EA, Cho ER, Jee SH. Hepatitis B virus infection and risk of non-Hodgkin lymphoma in South Korea: a cohort study. *Lancet Oncol*. 2010;11(9):827-834.

15. Dalia S, Chavez J, Castillo JJ, Sokol L. Hepatitis B infection increases the risk of non-Hodgkin lymphoma: a meta-analysis of observational studies. *Leuk. Res*. 2013;37(9):1107-1115.

16. Starkebaum G, Sasso EH. Hepatitis C and B cells: induction of autoimmunity and lymphoproliferation may reflect chronic stimulation through cell-surface receptors. *J. Rheumatol.* 2004;31(3):416-418.

17. Campbell J, Seymour JF, Matthews J, et al. The prognostic impact of bone marrow involvement in patients with diffuse large cell lymphoma varies according to the degree of infiltration and presence of discordant marrow involvement. *Eur. J. Haematol.* 2006;76(6): 473-480.

18. Chung R, Lai R, Wei P, et al. Concordant but not discordant bone marrow involvement in diffuse large B-cell lymphoma predicts a poor clinical outcome independent of the International Prognostic Index. *Blood.* 2007;110(4):1278-1282.

19. Talaulikar D, Dahlstrom JE, Shadbolt B, et al. Occult bone marrow involvement in patients with diffuse large B-cell lymphoma: results of a pilot study. *Pathology (Phila.).* 2007;39(6): 580-585.

20. Cabezas-Quintario MA, Gomez P, Yuste-Del Pozo V, et al. Bone marrow trephine biopsy involvement by lymphoma: pattern of involvement and concordance with flow cytometry, in 10 years from a single institution. *Clin. Transl. Oncol. Off. Publ. Fed. Span. Oncol. Soc. Natl. Cancer Inst. Mex.* 2015;

21. Kremer M, Spitzer M, Mandl-Weber S, et al. Discordant bone marrow involvement in diffuse large B-cell lymphoma: comparative molecular analysis reveals a heterogeneous group of disorders. *Lab. Investig. J. Tech. Methods Pathol.* 2003;83(1):107-114.

22. Alizadeh AA, Eisen MB, Davis RE, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature.* 2000;403(6769):503-511.

23. Wright G, Tan B, Rosenwald A, et al. A gene expression-based method to diagnose clinically distinct subgroups of diffuse large B cell lymphoma. *Proc. Natl. Acad. Sci. U. S. A.* 2003;100(17):9991-9996.

24. Rosenwald A, Wright G, Leroy K, et al. Molecular diagnosis of primary mediastinal B cell lymphoma identifies a clinically favorable subgroup of diffuse large B cell lymphoma related to Hodgkin lymphoma. *J. Exp. Med.* 2003;198(6):851-862.

25. Lenz G, Wright G, Dave SS, et al. Stromal gene signatures in large-B-cell lymphomas. *N. Engl. J. Med.* 2008;359(22):2313-2323.

26. Lenz G, Staudt LM. Aggressive lymphomas. *N. Engl. J. Med.* 2010;362(15):1417-1429.

27. Meyer PN, Fu K, Greiner TC, et al. Immunohistochemical methods for predicting cell of origin and survival in patients with diffuse large B-cell lymphoma treated with rituximab. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 2011;29(2):200-207.

28. Visco C, Li Y, Xu-Monette ZY, et al. Comprehensive gene expression profiling and immunohistochemical studies support application of immunophenotypic algorithm for molecular subtype classification in diffuse large B-cell lymphoma: a report from the International DLBCL Rituximab-CHOP Consortium Program Study. *Leukemia.* 2012;26(9): 2103-2113.

29. Rosenwald A, Wright G, Chan WC, et al. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N. Engl. J. Med.* 2002;346(25):1937-1947.

30. Dunleavy K, Pittaluga S, Czuczman MS, et al. Differential efficacy of bortezomib plus chemotherapy within molecular subtypes of diffuse large B-cell lymphoma. *Blood*. 2009;113(24):6069-6076.

31. Lenz G, Davis RE, Ngo VN, et al. Oncogenic CARD11 mutations in human diffuse large B cell lymphoma. *Science*. 2008;319(5870):1676-1679.

32. Milhollen MA, Traore T, Adams-Duffy J, et al. MLN4924, a NEDD8-activating enzyme inhibitor, is active in diffuse large B-cell lymphoma models: rationale for treatment of NF-{kappa}B-dependent lymphoma. *Blood*. 2010;116(9):1515-1523.

33. Offner F, Samoilova O, Osmanov E, et al. Frontline rituximab, cyclophosphamide, doxorubicin, and prednisone with bortezomib (VR-CAP) or vincristine (R-CHOP) for non-GCB DLBCL. *Blood*. 2015;

34. Hans CP, Weisenburger DD, Greiner TC, et al. Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. *Blood*. 2004;103(1):275-282.

35. Choi WWL, Weisenburger DD, Greiner TC, et al. A new immunostain algorithm classifies diffuse large B-cell lymphoma into molecular subtypes with high accuracy. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 2009;15(17):5494-5502.

36. Muris JJF, Meijer CJLM, Vos W, et al. Immunohistochemical profiling based on Bcl-2, CD10 and MUM1 expression improves risk stratification in patients with primary nodal diffuse large B cell lymphoma. *J. Pathol.* 2006;208(5):714-723.

37. Nyman H, Adde M, Karjalainen-Lindsberg M-L, et al. Prognostic impact of immunohistochemically defined germinal center phenotype in diffuse large B-cell lymphoma patients treated with immunochemotherapy. *Blood*. 2007;109(11):4930-4935.

38. Gutiérrez-García G, Cardesa-Salzmann T, Climent F, et al. Gene-expression profiling and not immunophenotypic algorithms predicts prognosis in patients with diffuse large B-cell lymphoma treated with immunochemotherapy. *Blood*. 2011;117(18):4836-4843.

39. Zu Y, Steinberg SM, Campo E, et al. Validation of tissue microarray immunohistochemistry staining and interpretation in diffuse large B-cell lymphoma. *Leuk. Lymphoma*. 2005;46(5):693-701.

40. de Jong D, Xie W, Rosenwald A, et al. Immunohistochemical prognostic markers in diffuse large B-cell lymphoma: validation of tissue microarray as a prerequisite for broad clinical applications (a study from the Lunenburg Lymphoma Biomarker Consortium). *J. Clin. Pathol.* 2009;62(2):128-138.

41. Culpin RE, Sieniawski M, Angus B, et al. Prognostic significance of immunohistochemistry-based markers and algorithms in immunochemotherapy-treated diffuse large B cell lymphoma patients. *Histopathology*. 2013;63(6):788-801.

42. Green TM, Young KH, Visco C, et al. Immunohistochemical double-hit score is a strong predictor of outcome in patients with diffuse large B-cell lymphoma treated with rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 2012;30(28):3460-3467.

43. Johnson NA, Slack GW, Savage KJ, et al. Concurrent Expression of MYC and BCL2 in Diffuse Large B-Cell Lymphoma Treated With Rituximab Plus Cyclophosphamide, Doxorubicin, Vincristine, and Prednisone. *J. Clin. Oncol.* 2012;30(28):3452-3459.

44. Hu S, Xu-Monette ZY, Tzankov A, et al. MYC/BCL2 protein coexpression contributes to the inferior survival of activated B-cell subtype of diffuse large B-cell lymphoma and demonstrates high-risk gene expression signatures: a report from The International DLBCL Rituximab-CHOP Consortium Program. *Blood*. 2013;121(20):4021-4031; quiz 4250.

45. Hu S, Xu-Monette ZY, Balasubramanyam A, et al. CD30 expression defines a novel subgroup of diffuse large B-cell lymphoma with favorable prognosis and distinct gene expression signature: a report from the International DLBCL Rituximab-CHOP Consortium Program Study. *Blood*. 2013;121(14):2715-2724.

46. Montes-Moreno S, Mollejo M, Fraga Máximo, et al. Recomendaciones para el estudio histopatológico, inmunohistoquímico, citogenético y molecular e informe diagnóstico de los procesos neoplásicos linfoides. Consenso SEAP-SEHH-GOTEL. Libro Blanco de Anatomía Patológica en España 2015.

47. van Dongen JJM, Langerak AW, Brüggemann M, et al. Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936. *Leukemia*. 2003;17(12):2257-2317.

48. Langerak AW, Groenen PJTA, Brüggemann M, et al. EuroClonality/BIOMED-2 guidelines for interpretation and reporting of Ig/TCR clonality testing in suspected lymphoproliferations. *Leukemia*. 2012;26(10):2159-2171.

49. Aukema SM, Siebert R, Schuuring E, et al. Double-hit B-cell lymphomas. *Blood*. 2011; 117(8):2319-2331.

50. Ruzinova MB, Caron T, Rodig SJ. Altered subcellular localization of c-Myc protein identifies aggressive B-cell lymphomas harboring a c-MYC translocation. *Am. J. Surg. Pathol.* 2010;34(6):882-891.

51. Pillai RK, Sathanoori M, Van Oss SB, Swerdlow SH. Double-hit B-cell lymphomas with BCL6 and MYC translocations are aggressive, frequently extranodal lymphomas distinct from BCL2 double-hit B-cell lymphomas. *Am. J. Surg. Pathol.* 2013;37(3):323-332.

52. Barrans S, Crouch S, Smith A, et al. Rearrangement of MYC is associated with poor prognosis in patients with diffuse large B-cell lymphoma treated in the era of rituximab. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 2010;28(20):3360-3365.

53. Ok CY, Xu-Monette ZY, Tzankov A, et al. Prevalence and clinical implications of cyclin D1 expression in diffuse large B-cell lymphoma (DLBCL) treated with immunochemotherapy:

a report from the International DLBCL Rituximab-CHOP Consortium Program. *Cancer*. 2014;120(12):1818-1829.

54. Valera A, Balagué O, Colomo L, et al. IG/MYC rearrangements are the main cytogenetic alteration in plasmablastic lymphomas. *Am. J. Surg. Pathol.* 2010;34(11):1686-1694.

55. Valera A, Colomo L, Martínez A, et al. ALK-positive large B-cell lymphomas express a terminal B-cell differentiation program and activated STAT3 but lack MYC rearrangements. *Mod. Pathol. Off. J. U. S. Can. Acad. Pathol. Inc.* 2013;26(10):1329-1337.

56. Testoni M, Zucca E, Young KH, Bertoni F. Genetic lesions in diffuse large B-cell lymphomas. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol. ESMO.* 2015;26(6):1069-1080.

57. Pasqualucci L, Dalla-Favera R. The genetic landscape of diffuse large B-cell lymphoma. *Semin. Hematol.* 2015;52(2):67-76.

58. Iqbal J, Meyer PN, Smith LM, et al. BCL2 predicts survival in germinal center B-cell-like diffuse large B-cell lymphoma treated with CHOP-like therapy and rituximab. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 2011;17(24):7785-7795.

59. Lenz G, Wright GW, Emre NCT, et al. Molecular subtypes of diffuse large B-cell lymphoma arise by distinct genetic pathways. *Proc. Natl. Acad. Sci. U. S. A.* 2008;105(36):13520-13525.

60. Lohr JG, Stojanov P, Lawrence MS, et al. Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. *Proc. Natl. Acad. Sci. U. S. A.* 2012;109(10):3879-3884.

61. Morin RD, Mungall K, Pleasance E, et al. Mutational and structural analysis of diffuse large B-cell lymphoma using whole-genome sequencing. *Blood.* 2013;122(7):1256-1265.

62. Saito M, Novak U, Piovan E, et al. BCL6 suppression of BCL2 via Miz1 and its disruption in diffuse large B cell lymphoma. *Proc. Natl. Acad. Sci. U. S. A.* 2009;106(27):11294-11299.

63. Dominguez-Sola D, Victora GD, Ying CY, et al. The proto-oncogene MYC is required for selection in the germinal center and cyclic reentry. *Nat. Immunol.* 2012;13(11):1083-1091.

64. Pfeifer M, Grau M, Lenze D, et al. PTEN loss defines a PI3K/AKT pathway-dependent germinal center subtype of diffuse large B-cell lymphoma. *Proc. Natl. Acad. Sci. U. S. A.* 2013;110(30):12420-12425.

65. Rao PH, Houldsworth J, Dyomina K, et al. Chromosomal and gene amplification in diffuse large B-cell lymphoma. *Blood.* 1998;92(1):234-240.

66. Houldsworth J, Olshen AB, Cattoretti G, et al. Relationship between REL amplification, REL function, and clinical and biologic features in diffuse large B-cell lymphomas. *Blood.* 2004;103(5):1862-1868.

67. Fukuhara N, Tagawa H, Kameoka Y, et al. Characterization of target genes at the 2p15-16 amplicon in diffuse large B-cell lymphoma. *Cancer Sci.* 2006;97(6):499-504.

68. O'Donnell KA, Wentzel EA, Zeller KI, Dang CV, Mendell JT. c-Myc-regulated microRNAs modulate E2F1 expression. *Nature.* 2005;435(7043):839-843.

69. He L, Thomson JM, Hemann MT, et al. A microRNA polycistron as a potential human oncogene. *Nature*. 2005;435(7043):828-833.

70. Mendell JT. miRiad roles for the miR-17-92 cluster in development and disease. *Cell*. 2008;133(2):217-222.

71. Inomata M, Tagawa H, Guo Y-M, et al. MicroRNA-17-92 down-regulates expression of distinct targets in different B-cell lymphoma subtypes. *Blood*. 2009;113(2):396-402.

72. Scandurra M, Rossi D, Deambrogi C, et al. Genomic profiling of Richter's syndrome: recurrent lesions and differences with de novo diffuse large B-cell lymphomas. *Hematol. Oncol.* 2010;28(2):62-67.

73. Schiffman JD, Lorimer PD, Rodic V, et al. Genome wide copy number analysis of paediatric Burkitt lymphoma using formalin-fixed tissues reveals a subset with gain of chromosome 13q and corresponding miRNA over expression. *Br. J. Haematol.* 2011;155(4):477-486.

74. Rao E, Jiang C, Ji M, et al. The miRNA-17~92 cluster mediates chemoresistance and enhances tumor growth in mantle cell lymphoma via PI3K/AKT pathway activation. *Leukemia*. 2012;26(5):1064-1072.

75. Chigrinova E, Mian M, Shen Y, et al. Integrated profiling of diffuse large B-cell lymphoma with 7q gain. *Br. J. Haematol.* 2011;153(4):499-503.

76. Chigrinova E, Mian M, Scandurra M, et al. Diffuse large B-cell lymphoma with concordant bone marrow involvement has peculiar genomic profile and poor clinical outcome. *Hematol. Oncol.* 2011;29(1):38-41.

77. Morin RD, Johnson NA, Severson TM, et al. Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. *Nat. Genet.* 2010;42(2):181-185.

78. Pasqualucci L, Dominguez-Sola D, Chiarenza A, et al. Inactivating mutations of acetyltransferase genes in B-cell lymphoma. *Nature*. 2011;471(7337):189-195.

79. Ryan RJH, Nitta M, Borger D, et al. EZH2 codon 641 mutations are common in BCL2-rearranged germinal center B cell lymphomas. *PloS One*. 2011;6(12):e28585.

80. Béguelin W, Popovic R, Teater M, et al. EZH2 is required for germinal center formation and somatic EZH2 mutations promote lymphoid transformation. *Cancer Cell*. 2013;23(5):677-692.

81. McCabe MT, Ott HM, Ganji G, et al. EZH2 inhibition as a therapeutic strategy for lymphoma with EZH2-activating mutations. *Nature*. 2012;492(7427):108-112.

82. Qi W, Chan H, Teng L, et al. Selective inhibition of Ezh2 by a small molecule inhibitor blocks tumor cells proliferation. *Proc. Natl. Acad. Sci. U. S. A.* 2012;109(52):21360-21365.

83. Morin RD, Mendez-Lago M, Mungall AJ, et al. Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature*. 2011;476(7360):298-303.

84. Zhang J, Grubor V, Love CL, et al. Genetic heterogeneity of diffuse large B-cell lymphoma. *Proc. Natl. Acad. Sci. U. S. A.* 2013;110(4):1398-1403.

85. Muppidi JR, Schmitz R, Green JA, et al. Loss of signalling via G 13 in germinal centre B-cell-derived lymphoma. *Nature*. 2014;516(7530):254-258.

86. Mandelbaum J, Bhagat G, Tang H, et al. BLIMP1 is a tumor suppressor gene frequently disrupted in activated B cell-like diffuse large B cell lymphoma. *Cancer Cell*. 2010;18(6): 568-579.

87. Shapiro-Shelef M, Lin K-I, McHeyzer-Williams LJ, et al. Blimp-1 is required for the formation of immunoglobulin secreting plasma cells and pre-plasma memory B cells. *Immunity*. 2003;19(4):607-620.

88. Pasqualucci L, Compagno M, Houldsworth J, et al. Inactivation of the PRDM1/BLIMP1 gene in diffuse large B cell lymphoma. *J. Exp. Med.* 2006;203(2):311-317.

89. Iqbal J, Greiner TC, Patel K, et al. Distinctive patterns of BCL6 molecular alterations and their functional consequences in different subgroups of diffuse large B-cell lymphoma. *Leukemia*. 2007;21(11):2332-2343.

90. Calado DP, Zhang B, Srinivasan L, et al. Constitutive canonical NF- B activation cooperates with disruption of BLIMP1 in the pathogenesis of activated B cell-like diffuse large cell lymphoma. *Cancer Cell*. 2010;18(6):580-589.

91. Lenz G, Nagel I, Siebert R, et al. Aberrant immunoglobulin class switch recombination and switch translocations in activated B cell-like diffuse large B cell lymphoma. *J. Exp. Med.* 2007;204(3):633-643.

92. Davis RE, Brown KD, Siebenlist U, Staudt LM. Constitutive nuclear factor kappaB activity is required for survival of activated B cell-like diffuse large B cell lymphoma cells. *J. Exp. Med.* 2001;194(12):1861-1874.

93. Compagno M, Lim WK, Grunn A, et al. Mutations of multiple genes cause deregulation of NF-kappaB in diffuse large B-cell lymphoma. *Nature*. 2009;459(7247):717-721.

94. Davis RE, Ngo VN, Lenz G, et al. Chronic active B-cell-receptor signalling in diffuse large B-cell lymphoma. *Nature*. 2010;463(7277):88-92.

95. Ngo VN, Young RM, Schmitz R, et al. Oncogenically active MYD88 mutations in human lymphoma. *Nature*. 2011;470(7332):115-119.

96. Thome M. CARMA1, BCL-10 and MALT1 in lymphocyte development and activation. *Nat. Rev. Immunol.* 2004;4(5):348-359.

97. Jiménez C, Sebastián E, Chillón MC, et al. MYD88 L265P is a marker highly characteristic of, but not restricted to, Waldenström's macroglobulinemia. *Leukemia*. 2013;27(8):1722-1728.

98. Kato M, Sanada M, Kato I, et al. Frequent inactivation of A20 in B-cell lymphomas. *Nature*. 2009;459(7247):712-716.

99. Roschewski M, Staudt LM, Wilson WH. Diffuse large B-cell lymphoma-treatment approaches in the molecular era. *Nat. Rev. Clin. Oncol.* 2014;11(1):12-23.

100. Jardin F, Jais J-P, Molina T-J, et al. Diffuse large B-cell lymphomas with CDKN2A deletion have a distinct gene expression signature and a poor prognosis under R-CHOP treatment: a GELA study. *Blood*. 2010;116(7):1092-1104.

101. Monti S, Chapuy B, Takeyama K, et al. Integrative analysis reveals an outcome-associated and targetable pattern of p53 and cell cycle deregulation in diffuse large B cell lymphoma. *Cancer Cell*. 2012;22(3):359-372.

102. Fabbri G, Khiabanian H, Holmes AB, et al. Genetic lesions associated with chronic lymphocytic leukemia transformation to Richter syndrome. *J. Exp. Med*. 2013;210(11):2273-2288.

103. Pasqualucci L, Khiabanian H, Fangazio M, et al. Genetics of follicular lymphoma transformation. *Cell Rep*. 2014;6(1):130-140.

104. Bouska A, McKeithan TW, Deffenbacher KE, et al. Genome-wide copy-number analyses reveal genomic abnormalities involved in transformation of follicular lymphoma. *Blood*. 2014;123(11):1681-1690.

105. Phan RT, Dalla-Favera R. The BCL6 proto-oncogene suppresses p53 expression in germinal-centre B cells. *Nature*. 2004;432(7017):635-639.

106. Goodman RH, Smolik S. CBP/p300 in cell growth, transformation, and development. *Genes Dev*. 2000;14(13):1553-1577.

107. Ying CY, Dominguez-Sola D, Fabi M, et al. MEF2B mutations lead to deregulated expression of the oncogene BCL6 in diffuse large B cell lymphoma. *Nat. Immunol*. 2013;14(10):1084-1092.

108. Green MR, Gentles AJ, Nair RV, et al. Hierarchy in somatic mutations arising during genomic evolution and progression of follicular lymphoma. *Blood*. 2013;121(9):1604-1611.

109. Pasqualucci L, Migliazza A, Fracchiolla N, et al. BCL-6 mutations in normal germinal center B cells: evidence of somatic hypermutation acting outside Ig loci. *Proc. Natl. Acad. Sci. U. S. A*. 1998;95(20):11816-11821.

110. Shen HM, Peters A, Baron B, Zhu X, Storb U. Mutation of BCL-6 gene in normal B cells by the process of somatic hypermutation of Ig genes. *Science*. 1998;280(5370):1750-1752.

111. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011;144(5):646-674.

112. Rimsza LM, Roberts RA, Miller TP, et al. Loss of MHC class II gene and protein expression in diffuse large B-cell lymphoma is related to decreased tumor immunosurveillance and poor patient survival regardless of other prognostic factors: a follow-up study from the Leukemia and Lymphoma Molecular Profiling Project. *Blood*. 2004;103(11):4251-4258.

113. Rimsza LM, Roberts RA, Campo E, et al. Loss of major histocompatibility class II expression in non-immune-privileged site diffuse large B-cell lymphoma is highly coordinated and not due to chromosomal deletions. *Blood*. 2006;107(3):1101-1107.

114. Cycon KA, Rimsza LM, Murphy SP. Alterations in CIITA constitute a common mechanism accounting for downregulation of MHC class II expression in diffuse large B-cell lymphoma (DLBCL). *Exp. Hematol.* 2009;37(2):184-194.

115. Challa-Malladi M, Lieu YK, Califano O, et al. Combined genetic inactivation of β2-Microglobulin and CD58 reveals frequent escape from immune recognition in diffuse large B cell lymphoma. *Cancer Cell.* 2011;20(6):728-740.

116. Riemersma SA, Jordanova ES, Schop RF, et al. Extensive genetic alterations of the HLA region, including homozygous deletions of HLA class II genes in B-cell lymphomas arising in immune-privileged sites. *Blood.* 2000;96(10):3569-3577.

117. Steidl C, Shah SP, Woolcock BW, et al. MHC class II transactivator CIITA is a recurrent gene fusion partner in lymphoid cancers. *Nature.* 2011;471(7338):377-381.

118. Rimsza LM, Chan WC, Gascoyne RD, et al. CIITA or RFX coding region loss of function mutations occur rarely in diffuse large B-cell lymphoma cases and cell lines with low levels of major histocompatibility complex class II expression. *Haematologica.* 2009;94(4):596-598.

119. Drénou B, Tilanus M, Semana G, et al. Loss of heterozygosity, a frequent but a non-exclusive mechanism responsible for HLA dysregulation in non-Hodgkin's lymphomas. *Br. J. Haematol.* 2004;127(1):40-49.

120. Capello D, Scandurra M, Poretti G, et al. Genome wide DNA-profiling of HIV-related B-cell lymphomas. *Br. J. Haematol.* 2010;148(2):245-255.

121. Rinaldi A, Capello D, Scandurra M, et al. Single nucleotide polymorphism-arrays provide new insights in the pathogenesis of post-transplant diffuse large B-cell lymphoma. *Br. J. Haematol.* 2010;149(4):569-577.

122. Cheung K-JJ, Johnson NA, Affleck JG, et al. Acquired TNFRSF14 mutations in follicular lymphoma are associated with worse prognosis. *Cancer Res.* 2010;70(22):9166-9174.

123. Grønbaek K, Straten PT, Ralfkiaer E, et al. Somatic Fas mutations in non-Hodgkin's lymphoma: association with extranodal disease and autoimmunity. *Blood.* 1998;92(9):3018-3024.

124. Wohlfart S, Sebinger D, Gruber P, et al. FAS (CD95) mutations are rare in gastric MALT lymphoma but occur more frequently in primary gastric diffuse large B-cell lymphoma. *Am. J. Pathol.* 2004;164(3):1081-1089.

125. Takahashi H, Feuerhake F, Kutok JL, et al. FAS death domain deletions and cellular FADD-like interleukin 1beta converting enzyme inhibitory protein (long) overexpression: alternative mechanisms for deregulating the extrinsic apoptotic pathway in diffuse large B-cell lymphoma subtypes. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 2006;12(11 Pt 1):3265-3271.

126. Rubio-Moscardo F, Blesa D, Mestre C, et al. Characterization of 8p21.3 chromosomal deletions in B-cell lymphoma: TRAIL-R1 and TRAIL-R2 as candidate dosage-dependent tumor suppressor genes. *Blood.* 2005;106(9):3214-3222.

127. Mian M, Scandurra M, Chigrinova E, et al. Clinical and molecular characterization of diffuse large B-cell lymphomas with 13q14.3 deletion. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol. ESMO.* 2012;23(3):729-735.

128. Scandurra M, Mian M, Greiner TC, et al. Genomic lesions associated with a different clinical outcome in diffuse large B-Cell lymphoma treated with R-CHOP-21. *Br. J. Haematol.* 2010;151(3):221-231.

129. Kojima Y, Tsurumi H, Goto N, et al. Fas and Fas ligand expression on germinal center type-diffuse large B-cell lymphoma is associated with the clinical outcome. *Eur. J. Haematol.* 2006;76(6):465-472.

130. Kramer MH, Hermans J, Parker J, et al. Clinical significance of bcl2 and p53 protein expression in diffuse large B-cell lymphoma: a population-based study. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 1996;14(7):2131-2138.

131. Klapproth K, Wirth T. Advances in the understanding of MYC-induced lymphomagenesis. *Br. J. Haematol.* 2010;149(4):484-497.

132. Kluk MJ, Chapuy B, Sinha P, et al. Immunohistochemical detection of MYC-driven diffuse large B-cell lymphomas. *PloS One.* 2012;7(4):e33813.

133. Horn H, Ziepert M, Becher C, et al. MYC status in concert with BCL2 and BCL6 expression predicts outcome in diffuse large B-cell lymphoma. *Blood.* 2013;121(12):2253-2263.

134. Savage KJ, Johnson NA, Ben-Neriah S, et al. MYC gene rearrangements are associated with a poor prognosis in diffuse large B-cell lymphoma patients treated with R-CHOP chemotherapy. *Blood.* 2009;114(17):3533-3537.

135. Johnson NA, Savage KJ, Ludkovski O, et al. Lymphomas with concurrent BCL2 and MYC translocations: the critical factors associated with survival. *Blood.* 2009;114(11):2273-2279.

136. Stasik CJ, Nitta H, Zhang W, et al. Increased MYC gene copy number correlates with increased mRNA levels in diffuse large B-cell lymphoma. *Haematologica.* 2010;95(4):597-603.

137. Snuderl M, Kolman OK, Chen Y-B, et al. B-cell lymphomas with concurrent IGH-BCL2 and MYC rearrangements are aggressive neoplasms with clinical and pathologic features distinct from Burkitt lymphoma and diffuse large B-cell lymphoma. *Am. J. Surg. Pathol.* 2010;34(3):327-340.

138. Barrans S, Crouch S, Smith A, et al. Rearrangement of MYC is associated with poor prognosis in patients with diffuse large B-cell lymphoma treated in the era of rituximab. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 2010;28(20):3360-3365.

139. Salaverria I, Siebert R. The gray zone between Burkitt's lymphoma and diffuse large B-cell lymphoma from a genetics perspective. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 2011; 29(14):1835-1843.

140. Testoni M, Kwee I, Greiner TC, et al. Gains of MYC locus and outcome in patients with diffuse large B-cell lymphoma treated with R-CHOP. *Br. J. Haematol.* 2011;155(2):274-277.

141. Li S, Seegmiller AC, Lin P, et al. B-cell lymphomas with concurrent MYC and BCL2 abnormalities other than translocations behave similarly to MYC/BCL2 double-hit lymphomas. *Mod. Pathol. Off. J. U. S. Can. Acad. Pathol. Inc.* 2015;28(2):208-217.

142. Cuccuini W, Briere J, Mounier N, et al. MYC+ diffuse large B-cell lymphoma is not salvaged by classical R-ICE or R-DHAP followed by BEAM plus autologous stem cell transplantation. *Blood.* 2012;119(20):4619-4624.

143. Friedberg JW. Double-hit diffuse large B-cell lymphoma. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 2012;30(28):3439-3443.

144. Oki Y, Noorani M, Lin P, et al. Double hit lymphoma: the MD Anderson Cancer Center clinical experience. *Br. J. Haematol.* 2014;166(6):891-901.

145. Petrich AM, Gandhi M, Jovanovic B, et al. Impact of induction regimen and stem cell transplantation on outcomes in double-hit lymphoma: a multicenter retrospective analysis. *Blood.* 2014;124(15):2354-2361.

146. A predictive model for aggressive non-Hodgkin's lymphoma. The International Non-Hodgkin's Lymphoma Prognostic Factors Project. *N. Engl. J. Med.* 1993;329(14):987-994.

147. Sehn LH, Donaldson J, Chhanabhai M, et al. Introduction of combined CHOP plus rituximab therapy dramatically improved outcome of diffuse large B-cell lymphoma in British Columbia. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 2005;23(22):5027-5033.

148. Ziepert M, Hasenclever D, Kuhnt E, et al. Standard International prognostic index remains a valid predictor of outcome for patients with aggressive CD20+ B-cell lymphoma in the rituximab era. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 2010;28(14):2373-2380.

149. Huang H-H, Xiao F, Chen F-Y, et al. Reassessment of the prognostic value of the International Prognostic Index and the revised International Prognostic Index in patients with diffuse large B-cell lymphoma: A multicentre study. *Exp. Ther. Med.* 2012;4(3):475-480.

150. Sehn LH, Berry B, Chhanabhai M, et al. The revised International Prognostic Index (R-IPI) is a better predictor of outcome than the standard IPI for patients with diffuse large B-cell lymphoma treated with R-CHOP. *Blood.* 2007;109(5):1857-1861.

151. Avilés A, Zepeda G, Díaz-Maqueo JC, et al. Beta 2 microglobulin level as an indicator of prognosis in diffuse large cell lymphoma. *Leuk. Lymphoma.* 1992;7(1-2):135-138.

152. Suki S, Swan F, Tucker S, et al. Risk classification for large cell lymphoma using lactate dehydrogenase, beta-2 microglobulin, and thymidine kinase. *Leuk. Lymphoma.* 1995;18(1-2):87-92.

153. Wu L, Wang T, Gui W, et al. Prognostic significance of serum beta-2 microglobulin in patients with non-Hodgkin lymphoma. *Oncology.* 2014;87(1):40-47.

154. Advani RH, Chen H, Habermann TM, et al. Comparison of conventional prognostic indices in patients older than 60 years with diffuse large B-cell lymphoma treated with R-CHOP in the US Intergroup Study (ECOG 4494, CALGB 9793): consideration of age greater than 70 years in an elderly prognostic index (E-IPI). *Br. J. Haematol.* 2010;151(2):143-151.

155. Coiffier B. Rituximab therapy in malignant lymphoma. *Oncogene*. 2007;26(25):3603-3613.

156. Cox MC, Nofroni I, Ruco L, et al. Low absolute lymphocyte count is a poor prognostic factor in diffuse-large-B-cell-lymphoma. *Leuk. Lymphoma*. 2008;49(9):1745-1751.

157. Chae YS, Shin H, Sohn SK, et al. Absolute lymphocyte count at day + 21 predicts survival in patients with early-stage diffuse large B-cell lymphoma treated with rituximab, cyclophosphamide, adriamycin, vincristine and prednisone. *Leuk. Lymphoma*. 2012;53(9): 1757-1763.

158. Sehn LH, Scott DW, Chhanabhai M, et al. Impact of concordant and discordant bone marrow involvement on outcome in diffuse large B-cell lymphoma treated with R-CHOP. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 2011;29(11):1452-1457.

159. Troppan KT, Schlick K, Deutsch A, et al. C-reactive protein level is a prognostic indicator for survival and improves the predictive ability of the R-IPI score in diffuse large B-cell lymphoma patients. *Br. J. Cancer*. 2014;111(1):55-60.

160. Huang Y-C, Liu C-Y, Lu H-J, et al. Comparison of prognostic models for patients with diffuse large B-cell lymphoma in the rituximab era. *Ann. Hematol.* 2013;92(11):1513-1520.

161. Sehn LH. Paramount prognostic factors that guide therapeutic strategies in diffuse large B-cell lymphoma. *Hematol. Educ. Program Am. Soc. Hematol. Am. Soc. Hematol. Educ. Program*. 2012;2012:402-409.

162. Sehn LH. Early detection of patients with poor risk diffuse large B-cell lymphoma. *Leuk. Lymphoma*. 2009;50(11):1744-1747.

163. Pfreundschuh M, Ho AD, Cavallin-Stahl E, et al. Prognostic significance of maximum tumour (bulk) diameter in young patients with good-prognosis diffuse large-B-cell lymphoma treated with CHOP-like chemotherapy with or without rituximab: an exploratory analysis of the MabThera International Trial Group (MInT) study. *Lancet Oncol.* 2008;9(5):435-444.

164. Müller C, Murawski N, Wiesen MHJ, et al. The role of sex and weight on rituximab clearance and serum elimination half-life in elderly patients with DLBCL. *Blood*. 2012;119(14): 3276-3284.

165. Maurer MJ, Cerhan JR, Katzmann JA, et al. Monoclonal and polyclonal serum free light chains and clinical outcome in chronic lymphocytic leukemia. *Blood*. 2011;118(10):2821-2826.

166. Haioun C, Itti E, Rahmouni A, et al. [18F]fluoro-2-deoxy-D-glucose positron emission tomography (FDG-PET) in aggressive lymphoma: an early prognostic tool for predicting patient outcome. *Blood*. 2005;106(4):1376-1381.

167. Pardal E, Coronado M, Martín A, et al. Intensification treatment based on early FDG-PET in patients with high-risk diffuse large B-cell lymphoma: a phase II GELTAMO trial. *Br. J. Haematol.* 2014;167(3):327-336.

168. Moskowitz CH, Schöder H. Current status of the role of PET imaging in diffuse large B-cell lymphoma. *Semin. Hematol.* 2015;52(2):138-142.

169. Ninan MJ, Wadhwa PD, Gupta P. Prognostication of diffuse large B-cell lymphoma in the rituximab era. *Leuk. Lymphoma.* 2011;52(3):360-373.

170. Salles G, de Jong D, Xie W, et al. Prognostic significance of immunohistochemical biomarkers in diffuse large B-cell lymphoma: a study from the Lunenburg Lymphoma Biomarker Consortium. *Blood.* 2011;117(26):7070-7078.

171. Sohn SK, Jung JT, Kim DH, et al. Prognostic significance of bcl-2, bax, and p53 expression in diffuse large B-cell lymphoma. *Am. J. Hematol.* 2003;73(2):101-107.

172. Mounier N, Briere J, Gisselbrecht C, et al. Rituximab plus CHOP (R-CHOP) overcomes bcl-2–associated resistance to chemotherapy in elderly patients with diffuse large B-cell lymphoma (DLBCL). *Blood.* 2003;101(11):4279-4284.

173. Iqbal J, Neppalli VT, Wright G, et al. BCL2 expression is a prognostic marker for the activated B-cell-like type of diffuse large B-cell lymphoma. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 2006;24(6):961-968.

174. Iqbal J, Sanger WG, Horsman DE, et al. BCL2 translocation defines a unique tumor subset within the germinal center B-cell-like diffuse large B-cell lymphoma. *Am. J. Pathol.* 2004;165(1):159-166.

175. Puvvada S, Kendrick S, Rimsza L. Molecular classification, pathway addiction, and therapeutic targeting in diffuse large B cell lymphoma. *Cancer Genet.* 2013;206(7-8):257-265.

176. Levine AJ. p53, the cellular gatekeeper for growth and division. *Cell.* 1997;88(3):323-331.

177. Hainaut P, Hollstein M. p53 and human cancer: the first ten thousand mutations. *Adv. Cancer Res.* 2000;77:81-137.

178. Lossos IS, Morgensztern D. Prognostic biomarkers in diffuse large B-cell lymphoma. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 2006;24(6):995-1007.

179. Young KH, Weisenburger DD, Dave BJ, et al. Mutations in the DNA-binding codons of TP53, which are associated with decreased expression of TRAILreceptor-2, predict for poor survival in diffuse large B-cell lymphoma. *Blood.* 2007;110(13):4396-4405.

180. Young KH, Leroy K, Møller MB, et al. Structural profiles of TP53 gene mutations predict clinical outcome in diffuse large B-cell lymphoma: an international collaborative study. *Blood.* 2008;112(8):3088-3098.

181. Ichikawa A, Kinoshita T, Watanabe T, et al. Mutations of the p53 gene as a prognostic factor in aggressive B-cell lymphoma. *N. Engl. J. Med.* 1997;337(8):529-534.

182. Zainuddin N, Berglund M, Wanders A, et al. TP53 mutations predict for poor survival in de novo diffuse large B-cell lymphoma of germinal center subtype. *Leuk. Res.* 2009;33(1):60-66.

183. Osada M, Ishioka C, Ichinohasama R, et al. Influence of p53 mutation on pathological grade, but not prognosis of non-Hodgkin's lymphoma. *Anticancer. Drug Des.* 1999;14(2):107-114.

184. el-Deiry WS, Tokino T, Velculescu VE, et al. WAF1, a potential mediator of p53 tumor suppression. *Cell.* 1993;75(4):817-825.

185. Winter JN, Li S, Aurora V, et al. Expression of p21 protein predicts clinical outcome in DLBCL patients older than 60 years treated with R-CHOP but not CHOP: a prospective ECOG and Southwest Oncology Group correlative study on E4494. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 2010;16(8):2435-2442.

186. Visco C, Canal F, Parolini C, et al. The impact of P53 and P21(waf1) expression on the survival of patients with the germinal center phenotype of diffuse large B-cell lymphoma. *Haematologica.* 2006;91(5):687-690.

187. Miller TP, Lippman SM, Spier CM, Slymen DJ, Grogan TM. HLA-DR (Ia) immune phenotype predicts outcome for patients with diffuse large cell lymphoma. *J. Clin. Invest.* 1988;82(1): 370-372.

188. Keir ME, Butte MJ, Freeman GJ, Sharpe AH. PD-1 and its ligands in tolerance and immunity. *Annu. Rev. Immunol.* 2008;26:677-704.

189. Topalian SL, Hodi FS, Brahmer JR, et al. Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *N. Engl. J. Med.* 2012;366(26):2443-2454.

190. Kiyasu J, Miyoshi H, Hirata A, et al. Expression of programmed cell death ligand 1 is associated with poor overall survival in patients with diffuse large B-cell lymphoma. *Blood.* 2015;126(19):2193-2201.

191. Chen J, Xu-Monette ZY, Deng L, et al. Dysregulated CXCR4 expression promotes lymphoma cell survival and independently predicts disease progression in germinal center B-cell-like diffuse large B-cell lymphoma. *Oncotarget.* 2015;6(8):5597-5614.

192. Moreno MJ, Bosch R, Dieguez-Gonzalez R, et al. CXCR4 expression enhances diffuse large B cell lymphoma dissemination and decreases patient survival. *J. Pathol.* 2015;235(3):445-455.

193. Xue X, Zeng N, Gao Z, Du M-Q. Diffuse large B-cell lymphoma: sub-classification by massive parallel quantitative RT-PCR. *Lab. Investig. J. Tech. Methods Pathol.* 2015;95(1):113-120.

194. Schmitz N, Nickelsen M, Ziepert M, et al. Conventional chemotherapy (CHOEP-14) with rituximab or high-dose chemotherapy (MegaCHOEP) with rituximab for young, high-risk patients with aggressive B-cell lymphoma: an open-label, randomised, phase 3 trial (DSHNHL 2002-1). *Lancet Oncol.* 2012;13(12):1250-1259.

195. Stiff PJ, Unger JM, Cook JR, et al. Autologous transplantation as consolidation for aggressive non-Hodgkin's lymphoma. *N. Engl. J. Med.* 2013;369(18):1681-1690.

196. Stiff PJ, Unger JM, Fisher RI. The role of myeloablation for lymphoma. *N. Engl. J. Med.* 2014;370(6):575-576.

197. García-Suárez J, Bañas H, Arribas I, et al. Dose-adjusted EPOCH plus rituximab is an effective regimen in patients with poor-prognostic untreated diffuse large B-cell lymphoma: results from a prospective observational study. *Br. J. Haematol.* 2007;136(2):276-285.

198. Howlett C, Snedecor SJ, Landsburg DJ, et al. Front-line, dose-escalated immunochemotherapy is associated with a significant progression-free survival advantage in patients with double-hit lymphomas: a systematic review and meta-analysis. *Br. J. Haematol.* 2015;170(4): 504-514.

199. Cerchietti LC, Ghetu AF, Zhu X, et al. A small-molecule inhibitor of BCL6 kills DLBCL cells in vitro and in vivo. *Cancer Cell.* 2010;17(4):400-411.

200. Younes A, Thieblemont C, Morschhauser F, et al. Combination of ibrutinib with rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone (R-CHOP) for treatment-naive patients with CD20-positive B-cell non-Hodgkin lymphoma: a non-randomised, phase 1b study. *Lancet Oncol.* 2014;15(9):1019-1026.

201. Nagel D, Spranger S, Vincendeau M, et al. Pharmacologic inhibition of MALT1 protease by phenothiazines as a therapeutic approach for the treatment of aggressive ABC-DLBCL. *Cancer Cell.* 2012;22(6):825-837.

202. Flinn IW, Kahl BS, Leonard JP, et al. Idelalisib, a selective inhibitor of phosphatidylinositol 3-kinase- , as therapy for previously treated indolent non-Hodgkin lymphoma. *Blood.* 2014;123(22):3406-3413.

203. Yang Y, Shaffer AL, Emre NCT, et al. Exploiting synthetic lethality for the therapy of ABC diffuse large B cell lymphoma. *Cancer Cell.* 2012;21(6):723-737.

204. Nowakowski GS, LaPlant B, Macon WR, et al. Lenalidomide combined with R-CHOP overcomes negative prognostic impact of non-germinal center B-cell phenotype in newly diagnosed diffuse large B-Cell lymphoma: a phase II study. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 2015;33(3):251-257.

205. Liesveld JL, Abboud CN, Duerst RE, et al. Characterization of human marrow stromal cells: role in progenitor cell binding and granulopoiesis. *Blood.* 1989;73(7):1794-1800.

206. Hemler ME. Adhesive protein receptors on hematopoietic cells. *Immunol. Today.* 1988;9(4):109-113.

207. Tarlinton D. Germinal centers: getting there is half the fun. *Curr. Biol. CB.* 1998;8(21):R753-756.

208. Young AJ. The physiology of lymphocyte migration through the single lymph node in vivo. *Semin. Immunol.* 1999;11(2):73-83.

209. Delves PJ, Roitt IM. The immune system. Second of two parts. *N. Engl. J. Med.* 2000;343(2): 108-117.

210. Delves PJ, Roitt IM. The immune system. First of two parts. *N. Engl. J. Med.* 2000;343(1): 37-49.

211. Kelsoe G. V(D)J hypermutation and receptor revision: coloring outside the lines. *Curr. Opin. Immunol.* 1999;11(1):70-75.

212. Slifka MK, Antia R, Whitmire JK, Ahmed R. Humoral immunity due to long-lived plasma cells. *Immunity.* 1998;8(3):363-372.

213. van Dongen JJ, Wolvers-Tettero IL. Analysis of immunoglobulin and T cell receptor genes. Part I: Basic and technical aspects. *Clin. Chim. Acta Int. J. Clin. Chem.* 1991;198(1-2):1-91.

214. Jensen GS, Mant MJ, Pilarski LM. Sequential maturation stages of monoclonal B lineage cells from blood, spleen, lymph node, and bone marrow from a terminal myeloma patient. *Am. J. Hematol.* 1992;41(3):199-208.

215. Reth M. Antigen receptors on B lymphocytes. *Annu. Rev. Immunol.* 1992;10:97-121.

216. Edelman GM. Antibody structure and molecular immunology. *Science.* 1973;180(4088):830-840.

217. Porter RR. Structural studies of immunoglobulins. *Science.* 1973;180(4087):713-716.

218. Alt FW, Blackwell TK, Yancopoulos GD. Development of the primary antibody repertoire. *Science.* 1987;238(4830):1079-1087.

219. Tonegawa S. Somatic generation of antibody diversity. *Nature.* 1983;302(5909):575-581.

220. Frippiat JP, Williams SC, Tomlinson IM, et al. Organization of the human immunoglobulin lambda light-chain locus on chromosome 22q11.2. *Hum. Mol. Genet.* 1995;4(6):983-991.

221. Erikson J, Martinis J, Croce CM. Assignment of the genes for human lambda immunoglobulin chains to chromosome 22. *Nature.* 1981;294(5837):173-175.

222. Ruiz M, Pallarès N, Contet V, Barbi V, Lefranc MP. The human immunoglobulin heavy diversity (IGHD) and joining (IGHJ) segments. *Exp. Clin. Immunogenet.* 1999;16(3):173-184.

223. Pallarès N, Lefebvre S, Contet V, Matsuda F, Lefranc MP. The human immunoglobulin heavy variable genes. *Exp. Clin. Immunogenet.* 1999;16(1):36-60.

224. Lefranc M-P. IMGT, the international ImMunoGeneTics database: a high-quality information system for comparative immunogenetics and immunology. *Dev. Comp. Immunol.* 2002;26(8):697-705.

225. Hofker MH, Walter MA, Cox DW. Complete physical map of the human immunoglobulin heavy chain constant region gene complex. *Proc. Natl. Acad. Sci. U. S. A.* 1989;86(14):5567-5571.

226. Dunnick W, Hertz GZ, Scappino L, Gritzmacher C. DNA sequences at immunoglobulin switch region recombination sites. *Nucleic Acids Res.* 1993;21(3):365-372.

227. Alt FW, Yancopoulos GD, Blackwell TK, et al. Ordered rearrangement of immunoglobulin heavy chain variable region segments. *EMBO J.* 1984;3(6):1209-1219.

228. Sleckman BP, Gorman JR, Alt FW. Accessibility control of antigen-receptor variable-region gene assembly: role of cis-acting elements. *Annu. Rev. Immunol.* 1996;14:459-481.

229. Melchers F. The pre-B-cell receptor: selector of fitting immunoglobulin heavy chains for the B-cell repertoire. *Nat. Rev. Immunol.* 2005;5(7):578-584.

230. Lanig H, Bradl H, Jäck H-M. Three-dimensional modeling of a pre-B-cell receptor. *Mol. Immunol.* 2004;40(17):1263-1272.

231. Schiff C, Bensmana M, Guglielmi P, et al. The immunoglobulin lambda-like gene cluster (14.1, 16.1 and F lambda 1) contains gene(s) selectively expressed in pre-B cells and is the human counterpart of the mouse lambda 5 gene. *Int. Immunol.* 1990;2(3):201-207.

232. Minaee S, Farmer D, Georgiou A, et al. Mapping and functional analysis of regulatory sequences in the mouse lambda5-VpreB1 domain. *Mol. Immunol.* 2005;42(11):1283-1292.

233. Sabbattini P, Dillon N. The lambda5-VpreB1 locus–a model system for studying gene regulation during early B cell development. *Semin. Immunol.* 2005;17(2):121-127.

234. Melchers F, ten Boekel E, Yamagami T, Andersson J, Rolink A. The roles of preB and B cell receptors in the stepwise allelic exclusion of mouse IgH and L chain gene loci. *Semin. Immunol.* 1999;11(5):307-317.

235. Wang LD, Clark MR. B-cell antigen-receptor signalling in lymphocyte development. *Immunology.* 2003;110(4):411-420.

236. Graninger WB, Goldman PL, Morton CC, O'Brien SJ, Korsmeyer SJ. The kappa-deleting element. Germline and rearranged, duplicated and dispersed forms. *J. Exp. Med.* 1988;167(2): 488-501.

237. Siminovitch KA, Bakhshi A, Goldman P, Korsmeyer SJ. A uniform deleting element mediates the loss of kappa genes in human B cells. *Nature.* 1985;316(6025):260-262.

238. Gay D, Saunders T, Camper S, Weigert M. Receptor editing: an approach by autoreactive B cells to escape tolerance. *J. Exp. Med.* 1993;177(4):999-1008.

239. Sandel PC, Monroe JG. Negative selection of immature B cells by receptor editing or deletion is determined by site of antigen encounter. *Immunity.* 1999;10(3):289-299.

240. Nemazee D. Receptor editing in B cells. *Adv. Immunol.* 2000;74:89-126.

241. Notarangelo LD, Santagata S, Villa A. Recombinase activating gene enzymes of lymphocytes. *Curr. Opin. Hematol.* 2001;8(1):41-46.

242. Grawunder U, West RB, Lieber MR. Antigen receptor gene rearrangement. *Curr. Opin. Immunol.* 1998;10(2):172-180.

243. van Gent DC, Mizuuchi K, Gellert M. Similarities between initiation of V(D)J recombination and retroviral integration. *Science.* 1996;271(5255):1592-1594.

244. Gottlieb TM, Jackson SP. The DNA-dependent protein kinase: requirement for DNA ends and association with Ku antigen. *Cell.* 1993;72(1):131-142.

245. Gu Y, Jin S, Gao Y, Weaver DT, Alt FW. Ku70-deficient embryonic stem cells have increased ionizing radiosensitivity, defective DNA end-binding activity, and inability to support V(D)J recombination. *Proc. Natl. Acad. Sci. U. S. A.* 1997;94(15):8076-8081.

246. Nussenzweig A, Chen C, da Costa Soares V, et al. Requirement for Ku80 in growth and immunoglobulin V(D)J recombination. *Nature.* 1996;382(6591):551-555.

247. Lewis SM. The mechanism of V(D)J joining: lessons from molecular, immunological, and comparative analyses. *Adv. Immunol.* 1994;56:27-150.

248. Goodnow CC, Crosbie J, Adelstein S, et al. Altered immunoglobulin expression and functional silencing of self-reactive B lymphocytes in transgenic mice. *Nature*. 1988;334(6184): 676-682.

249. Gay D, Saunders T, Camper S, Weigert M. Receptor editing: an approach by autoreactive B cells to escape tolerance. *J. Exp. Med.* 1993;177(4):999-1008.

250. Prak EL, Trounstine M, Huszar D, Weigert M. Light chain editing in kappa-deficient animals: a potential mechanism of B cell tolerance. *J. Exp. Med.* 1994;180(5):1805-1815.

251. Chen C, Nagy Z, Prak EL, Weigert M. Immunoglobulin heavy chain gene replacement: a mechanism of receptor editing. *Immunity*. 1995;3(6):747-755.

252. Steenbergen EJ, Verhagen OJ, van den Berg H, et al. Rearrangement status of the malignant cell determines type of secondary IgH rearrangement (V-replacement or V to DJ joining) in childhood B precursor acute lymphoblastic leukemia. *Leukemia*. 1997;11(8):1258-1265.

253. Gilfillan S, Dierich A, Lemeur M, Benoist C, Mathis D. Mice lacking TdT: mature animals with an immature lymphocyte repertoire. *Science*. 1993;261(5125):1175-1178.

254. Tumas-Brundage K, Manser T. The transcriptional promoter regulates hypermutation of the antibody heavy chain locus. *J. Exp. Med.* 1997;185(2):239-250.

255. Fukita Y, Jacobs H, Rajewsky K. Somatic hypermutation in the heavy chain locus correlates with transcription. *Immunity*. 1998;9(1):105-114.

256. Betz AG, Milstein C, González-Fernández A, et al. Elements regulating somatic hypermutation of an immunoglobulin kappa gene: critical role for the intron enhancer/matrix attachment region. *Cell*. 1994;77(2):239-248.

257. Lebecque SG, Gearhart PJ. Boundaries of somatic mutation in rearranged immunoglobulin genes: 5' boundary is near the promoter, and 3' boundary is approximately 1 kb from V(D)J gene. *J. Exp. Med.* 1990;172(6):1717-1727.

258. Müller-Hermelink HK, Greiner A. Molecular analysis of human immunoglobulin heavy chain variable genes (IgVH) in normal and malignant B cells. *Am. J. Pathol.* 1998;153(5): 1341-1346.

259. Goossens T, Klein U, Küppers R. Frequent occurrence of deletions and duplications during somatic hypermutation: implications for oncogene translocations and heavy chain disease. *Proc. Natl. Acad. Sci. U. S. A.* 1998;95(5):2463-2468.

260. Milstein C, Neuberger MS, Staden R. Both DNA strands of antibody genes are hypermutation targets. *Proc. Natl. Acad. Sci. U. S. A.* 1998;95(15):8791-8794.

261. Neuberger MS, Milstein C. Somatic hypermutation. *Curr. Opin. Immunol.* 1995;7(2): 248-254.

262. Stavnezer J. Antibody class switching. *Adv. Immunol.* 1996;61:79-146.

263. Stavnezer J. Immunoglobulin class switching. *Curr. Opin. Immunol.* 1996;8(2):199-205.

264. Iwasato T, Shimizu A, Honjo T, Yamagishi H. Circular DNA is excised by immunoglobulin class switch recombination. *Cell*. 1990;62(1):143-149.

265. Esser C, Radbruch A. Immunoglobulin class switching: molecular and cellular analysis. *Annu. Rev. Immunol.* 1990;8:717-735.

266. García-Sanz R, López-Pérez R, Langerak AW, et al. Heteroduplex PCR analysis of rearranged immunoglobulin genes for clonality assessment in multiple myeloma. *Haematologica*. 1999;84(4):328-335.

267. González M, González D, López-Pérez R, et al. Heteroduplex analysis of VDJ amplified segments from rearranged IgH genes for clonality assessments in B-cell non-Hodgkin's lymphoma. A comparison between different strategies. *Haematologica*. 1999;84(9):779-784.

268. Langerak AW, Szczepa ski T, van der Burg M, Wolvers-Tettero IL, van Dongen JJ. Heteroduplex PCR analysis of rearranged T cell receptor genes for clonality assessment in suspect T cell proliferations. *Leukemia*. 1997;11(12):2192-2199.

269. Segal GH, Jorgensen T, Masih AS, Braylan RC. Optimal primer selection for clonality assessment by polymerase chain reaction analysis: I. Low grade B-cell lymphoproliferative disorders of nonfollicular center cell type. *Hum. Pathol.* 1994;25(12):1269-1275.

270. Segal GH, Jorgensen T, Scott M, Braylan RC. Optimal primer selection for clonality assessment by polymerase chain reaction analysis: II. Follicular lymphomas. *Hum. Pathol.* 1994;25(12):1276-1282.

271. Aubin J, Davi F, Nguyen-Salomon F, et al. Description of a novel FR1 IgH PCR strategy and its comparison with three other strategies for the detection of clonality in B cell malignancies. *Leukemia*. 1995;9(3):471-479.

272. Deane M, McCarthy KP, Wiedemann LM, Norton JD. An improved method for detection of B-lymphoid clonality by polymerase chain reaction. *Leukemia*. 1991;5(8):726-730.

273. Segal GH, Wittwer CT, Fishleder AJ, et al. Identification of monoclonal B-cell populations by rapid cycle polymerase chain reaction. A practical screening method for the detection of immunoglobulin gene rearrangements. *Am. J. Pathol.* 1992;141(6):1291-1297.

274. Linke B, Pyttlich J, Tiemann M, et al. Identification and structural analysis of rearranged immunoglobulin heavy chain genes in lymphomas and leukemias. *Leukemia*.1995;9(5):840-847.

275. López-Pérez R, García-Sanz R, González D, et al. The detection of contaminating clonal cells in apheresis products is related to response and outcome in multiple myeloma undergoing autologous peripheral blood stem cell transplantation. *Leukemia*. 2000;14(8):1493-1499.

276. Schuh A, Becq J, Humphray S, et al. Monitoring chronic lymphocytic leukemia progression by whole genome sequencing reveals heterogeneous clonal evolution patterns. *Blood*. 2012;120(20):4191-4196.

277. Wu CJ. CLL clonal heterogeneity: an ecology of competing subpopulations. *Blood*. 2012;120(20):4117-4118.

278. Lossos IS, Levy R, Alizadeh AA. AID is expressed in germinal center B-cell-like and activated B-cell-like diffuse large-cell lymphomas and is not correlated with intraclonal heterogeneity. *Leukemia*. 2004;18(11):1775-1779.

279. Morin RD, Mungall K, Pleasance E, et al. Mutational and structural analysis of diffuse large B-cell lymphoma using whole-genome sequencing. *Blood*. 2013;122(7):1256-1265.

280. Melchardt T, Hufnagl C, Weigert O, et al. Clonal Evolution in Relapsed or Refractory Diffuse Large B Cell Lymphoma. *Blood*. 2014;124(21):77-77.

281. Sufficool KE, Lockwood CM, Abel HJ, et al. T-cell clonality assessment by next-generation sequencing improves detection sensitivity in mycosis fungoides. *J. Am. Acad. Dermatol.* 2015;73(2):228-236.e2.

282. Ramesh M, Hamm D, Simchoni N, Cunningham-Rundles C. Clonal and constricted T cell repertoire in Common Variable Immune Deficiency. *Clin. Immunol. Orlando Fla*. 2015;

283. Rechavi E, Lev A, Lee YN, et al. Timely and spatially regulated maturation of B and T cell repertoire during human fetal development. *Sci. Transl. Med.* 2015;7(276):276ra25.

284. Boyd SD, Marshall EL, Merker JD, et al. Measurement and clinical monitoring of human lymphocyte clonality by massively parallel VDJ pyrosequencing. *Sci. Transl. Med.* 2009;1(12):12ra23.

285. Moorhouse MJ, van Zessen D, IJspeert H, et al. ImmunoGlobulin galaxy (IGGalaxy) for simple determination and quantitation of immunoglobulin heavy chain rearrangements from NGS. *BMC Immunol.* 2014;15:59.

286. Gawad C, Pepin F, Carlton VEH, et al. Massive evolution of the immunoglobulin heavy chain locus in children with B precursor acute lymphoblastic leukemia. *Blood*. 2012;120(22):4407-4417.

287. Carlotti E, Wrench D, Rosignoli G, et al. High Throughput Sequencing Analysis of the Immunoglobulin Heavy Chain Gene from Flow-Sorted B Cell Sub-Populations Define the Dynamics of Follicular Lymphoma Clonal Evolution. *PloS One*. 2015;10(9):e0134833.

288. MacLennan IC. Germinal centers. *Annu. Rev. Immunol.* 1994;12:117-139.

289. Klein U, Dalla-Favera R. Germinal centres: role in B-cell physiology and malignancy. *Nat. Rev. Immunol.* 2008;8(1):22-33.

290. Basso K, Dalla-Favera R. Germinal centres and B cell lymphomagenesis. *Nat. Rev. Immunol.* 2015;15(3):172-184.

291. Küppers R, Dalla-Favera R. Mechanisms of chromosomal translocations in B cell lymphomas. *Oncogene*. 2001;20(40):5580-5594.

292. Küppers R. Mechanisms of B-cell lymphoma pathogenesis. *Nat. Rev. Cancer*. 2005;5(4):251-262.

293. Victora GD, Schwickert TA, Fooksman DR, et al. Germinal center dynamics revealed by multiphoton microscopy with a photoactivatable fluorescent reporter. *Cell*. 2010;143(4):592-605.

294. Victora GD, Dominguez-Sola D, Holmes AB, et al. Identification of human germinal center light and dark zone cells and their relationship to human B-cell lymphomas. *Blood*. 2012;120(11):2240-2248.

295. Allen CDC, Okada T, Tang HL, Cyster JG. Imaging of germinal center selection events during affinity maturation. *Science*. 2007;315(5811):528-531.

296. Caron G, Le Gallou S, Lamy T, Tarte K, Fest T. CXCR4 expression functionally discriminates centroblasts versus centrocytes within human germinal center B cells. *J. Immunol. Baltim. Md 1950*. 2009;182(12):7595-7602.

297. Schwickert TA, Lindquist RL, Shakhar G, et al. In vivo imaging of germinal centres reveals a dynamic open structure. *Nature*. 2007;446(7131):83-87.

298. Richter J, Schlesner M, Hoffmann S, et al. Recurrent mutation of the ID3 gene in Burkitt lymphoma identified by integrated genome, exome and transcriptome sequencing. *Nat. Genet*. 2012;44(12):1316-1320.

299. Scott DW, Gascoyne RD. The tumour microenvironment in B cell lymphomas. *Nat. Rev. Cancer*. 2014;14(8):517-534.

300. Muramatsu M, Kinoshita K, Fagarasan S, et al. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell*. 2000;102(5):553-563.

301. Dalla-Favera R, Bregni M, Erikson J, et al. Human c-myc onc gene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells. *Proc. Natl. Acad. Sci. U. S. A.* 1982;79(24):7824-7827.

302. Taub R, Kirsch I, Morton C, et al. Translocation of the c-myc gene into the immunoglobulin heavy chain locus in human Burkitt lymphoma and murine plasmacytoma cells. *Proc. Natl. Acad. Sci. U. S. A.* 1982;79(24):7837-7841.

303. Calado DP, Sasaki Y, Godinho SA, et al. The cell-cycle regulator c-Myc is essential for the formation and maintenance of germinal centers. *Nat. Immunol.* 2012;13(11):1092-1100.

304. Lin CY, Lovén J, Rahl PB, et al. Transcriptional amplification in tumor cells with elevated c-Myc. *Cell*. 2012;151(1):56-67.

305. Ye BH, Lista F, Lo Coco F, et al. Alterations of a zinc finger-encoding gene, BCL-6, in diffuse large-cell lymphoma. *Science*. 1993;262(5134):747-750.

306. Basso K, Dalla-Favera R. Roles of BCL6 in normal and transformed germinal center B cells. *Immunol. Rev.* 2012;247(1):172-183.

307. Kerfoot SM, Yaari G, Patel JR, et al. Germinal center B cell and T follicular helper cell development initiates in the interfollicular zone. *Immunity*. 2011;34(6):947-960.

308. Kitano M, Moriyama S, Ando Y, et al. Bcl6 protein expression shapes pre-germinal center B cell dynamics and follicular helper T cell heterogeneity. *Immunity*. 2011;34(6):961-972.

309. Ochiai K, Maienschein-Cline M, Simonetti G, et al. Transcriptional regulation of germinal center B and plasma cell fates by dynamical control of IRF4. *Immunity*. 2013;38(5):918-929.

310. Ci W, Polo JM, Cerchietti L, et al. The BCL6 transcriptional program features repression of multiple oncogenes in primary B cells and is deregulated in DLBCL. *Blood*. 2009;113(22): 5536-5548.

311. Basso K, Saito M, Sumazin P, et al. Integrated biochemical and computational approach identifies BCL6 direct target genes controlling multiple pathways in normal germinal center B cells. *Blood*. 2010;115(5):975-984.

312. Basso K, Schneider C, Shen Q, et al. BCL6 positively regulates AID and germinal center gene expression via repression of miR-155. *J. Exp. Med*. 2012;209(13):2455-2465.

313. Dent AL, Shaffer AL, Yu X, Allman D, Staudt LM. Control of inflammation, cytokine expression, and germinal center formation by BCL-6. *Science*. 1997;276(5312):589-592.

314. Ye BH, Cattoretti G, Shen Q, et al. The BCL-6 proto-oncogene controls germinal-centre formation and Th2-type inflammation. *Nat. Genet*. 1997;16(2):161-170.

315. Cattoretti G, Pasqualucci L, Ballon G, et al. Deregulated BCL6 expression recapitulates the pathogenesis of human diffuse large B cell lymphomas in mice. *Cancer Cell*. 2005;7(5):445-455.

316. Vallabhapurapu S, Karin M. Regulation and function of NF-kappaB transcription factors in the immune system. *Annu. Rev. Immunol*. 2009;27:693-733.

317. Saito M, Gao J, Basso K, et al. A signaling pathway mediating downregulation of BCL6 in germinal center B cells is blocked by BCL6 gene alterations in B cell lymphoma. *Cancer Cell*. 2007;12(3):280-292.

318. Tavares RM, Turer EE, Liu CL, et al. The ubiquitin modifying enzyme A20 restricts B cell survival and prevents autoimmunity. *Immunity*. 2010;33(2):181-191.

319. Klein U, Casola S, Cattoretti G, et al. Transcription factor IRF4 controls plasma cell differentiation and class-switch recombination. *Nat. Immunol*. 2006;7(7):773-782.

320. Shaffer AL, Yu X, He Y, et al. BCL-6 represses genes that function in lymphocyte differentiation, inflammation, and cell cycle control. *Immunity*. 2000;13(2):199-212.

321. Tunyaplin C, Shaffer AL, Angelin-Duclos CD, et al. Direct repression of prdm1 by Bcl-6 inhibits plasmacytic differentiation. *J. Immunol. Baltim. Md 1950*. 2004;173(2):1158-1165.

322. Delogu A, Schebesta A, Sun Q, et al. Gene repression by Pax5 in B cells is essential for blood cell homeostasis and is reversed in plasma cells. *Immunity*. 2006;24(3):269-281.

323. Nera K-P, Kohonen P, Narvi E, et al. Loss of Pax5 promotes plasma cell differentiation. *Immunity*. 2006;24(3):283-293.

324. Cobaleda C, Schebesta A, Delogu A, Busslinger M. Pax5: the guardian of B cell identity and function. *Nat. Immunol.* 2007;8(5):463-470.

325. Nutt SL, Hodgkin PD, Tarlinton DM, Corcoran LM. The generation of antibody-secreting plasma cells. *Nat. Rev. Immunol.* 2015;15(3):160-171.

326. Sciammas R, Shaffer AL, Schatz JH, et al. Graded expression of interferon regulatory factor-4 coordinates isotype switching with plasma cell differentiation. *Immunity.* 2006;25(2): 225-236.

327. Reimold AM, Iwakoshi NN, Manis J, et al. Plasma cell differentiation requires the transcription factor XBP-1. *Nature.* 2001;412(6844):300-307.

328. Küppers R, Klein U, Hansmann ML, Rajewsky K. Cellular origin of human B-cell lymphomas. *N. Engl. J. Med.* 1999;341(20):1520-1529.

329. Stevenson FK, Sahota SS, Ottensmeier CH, et al. The occurrence and significance of V gene mutations in B cell-derived human malignancy. *Adv. Cancer Res.* 2001;83:81-116.

330. Ramiro AR, Jankovic M, Eisenreich T, et al. AID is required for c-myc/IgH chromosome translocations in vivo. *Cell.* 2004;118(4):431-438.

331. Pasqualucci L, Bhagat G, Jankovic M, et al. AID is required for germinal center-derived lymphomagenesis. *Nat. Genet.* 2008;40(1):108-112.

332. Pasqualucci L, Neumeister P, Goossens T, et al. Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. *Nature.* 2001;412(6844):341-346.

333. Yamane A, Resch W, Kuo N, et al. Deep-sequencing identification of the genomic targets of the cytidine deaminase AID and its cofactor RPA in B lymphocytes. *Nat. Immunol.* 2011;12(1):62-69.

334. Liu M, Duke JL, Richter DJ, et al. Two levels of protection for the B cell genome during somatic hypermutation. *Nature.* 2008;451(7180):841-845.

335. Khodabakhshi AH, Morin RD, Fejes AP, et al. Recurrent targets of aberrant somatic hypermutation in lymphoma. *Oncotarget.* 2012;3(11):1308-1319.

336. Lossos IS, Alizadeh AA, Eisen MB, et al. Ongoing immunoglobulin somatic mutation in germinal center B cell-like but not in activated B cell-like diffuse large cell lymphomas. *Proc. Natl. Acad. Sci. U. S. A.* 2000;97(18):10209-10213.

337. Jardin F, Sahota SS, Ruminy P, et al. Novel Ig V gene features of t(14;18) and t(3;14) de novo diffuse large B-cell lymphoma displaying germinal center-B cell like and non-germinal center-B cell like markers. *Leukemia.* 2006;20(11):2070-2074.

338. Ruminy P, Etancelin P, Couronné L, et al. The isotype of the BCR as a surrogate for the GCB and ABC molecular subtypes in diffuse large B-cell lymphoma. *Leukemia.* 2011;25(4):681-688.

339. Hsu FJ, Levy R. Preferential use of the VH4 Ig gene family by diffuse large-cell lymphoma. *Blood.* 1995;86(8):3072-3082.

340. Küppers R, Rajewsky K, Hansmann ML. Diffuse large cell lymphomas are derived from mature B cells carrying V region genes with a high load of somatic mutation and evidence of selection for antibody expression. *Eur. J. Immunol.* 1997;27(6):1398-1405.

341. Rosenquist R, Lindström A, Holmberg D, Lindh J, Roos G. V(H) gene family utilization in different B-cell lymphoma subgroups. *Eur. J. Haematol.* 1999;62(2):123-128.

342. Lossos IS, Okada CY, Tibshirani R, et al. Molecular analysis of immunoglobulin genes in diffuse large B-cell lymphomas. *Blood.* 2000;95(5):1797-1803.

343. Nakamura N, Kuze T, Hashimoto Y, et al. Analysis of the immunoglobulin heavy chain gene variable region of CD5-positive and -negative diffuse large B cell lymphoma. *Leukemia.* 2001;15(3):452-457.

344. Berkowska MA, Driessen GJA, Bikos V, et al. Human memory B cells originate from three distinct germinal center-dependent and -independent maturation pathways. *Blood.* 2011;118(8):2150-2158.

345. Brezinschek HP, Foster SJ, Brezinschek RI, et al. Analysis of the human VH gene repertoire. Differential effects of selection and somatic hypermutation on human peripheral CD5(+)/IgM+ and CD5(-)/IgM+ B cells. *J. Clin. Invest.* 1997;99(10):2488-2501.

346. Capello D, Martini M, Gloghini A, et al. Molecular analysis of immunoglobulin variable genes in human immunodeficiency virus-related non-Hodgkin's lymphoma reveals implications for disease pathogenesis and histogenesis. *Haematologica.* 2008;93(8):1178-1185.

347. Niemann CU, Wiestner A. B-cell receptor signaling as a driver of lymphoma development and evolution. *Semin. Cancer Biol.* 2013;23(6):410-421.

348. Sutton L-A, Agathangelidis A, Belessi C, et al. Antigen selection in B-cell lymphomas–tracing the evidence. *Semin. Cancer Biol.* 2013;23(6):399-409.

349. Chan CH, Hadlock KG, Foung SK, Levy S. V(H)1-69 gene is preferentially used by hepatitis C virus-associated B cell lymphomas and by normal B cells responding to the E2 viral antigen. *Blood.* 2001;97(4):1023-1026.

350. Stamatopoulos K, Belessi C, Moreno C, et al. Over 20% of patients with chronic lymphocytic leukemia carry stereotyped receptors: Pathogenetic implications and clinical correlations. *Blood.* 2007;109(1):259-270.

351. Murray F, Darzentas N, Hadzidimitriou A, et al. Stereotyped patterns of somatic hypermutation in subsets of patients with chronic lymphocytic leukemia: implications for the role of antigen selection in leukemogenesis. *Blood.* 2008;111(3):1524-1533.

352. Agathangelidis A, Darzentas N, Hadzidimitriou A, et al. Stereotyped B-cell receptors in one-third of chronic lymphocytic leukemia: a molecular classification with implications for targeted therapies. *Blood.* 2012;119(19):4467-4475.

353. Hadzidimitriou A, Agathangelidis A, Darzentas N, et al. Is there a role for antigen selection in mantle cell lymphoma? Immunogenetic support from a series of 807 cases. *Blood.* 2011;118(11):3088-3095.

354. Bikos V, Darzentas N, Hadzidimitriou A, et al. Over 30% of patients with splenic marginal zone lymphoma express the same immunoglobulin heavy variable gene: ontogenetic implications. *Leukemia*. 2012;26(7):1638-1646.

355. Zibellini S, Capello D, Forconi F, et al. Stereotyped patterns of B-cell receptor in splenic marginal zone lymphoma. *Haematologica*. 2010;95(10):1792-1796.

356. Dagklis A, Ponzoni M, Govi S, et al. Immunoglobulin gene repertoire in ocular adnexal lymphomas: hints on the nature of the antigenic stimulation. *Leukemia*. 2012;26(4):814-821.

357. Darzentas N, Hadzidimitriou A, Murray F, et al. A different ontogenesis for chronic lymphocytic leukemia cases carrying stereotyped antigen receptors: molecular and computational evidence. *Leukemia*. 2010;24(1):125-132.

358. Zhu D, McCarthy H, Ottensmeier CH, et al. Acquisition of potential N-glycosylation sites in the immunoglobulin variable region by somatic mutation is a distinctive feature of follicular lymphoma. *Blood*. 2002;99(7):2562-2568.

359. Radcliffe CM, Arnold JN, Suter DM, et al. Human follicular lymphoma cells contain oligomannose glycans in the antigen-binding site of the B-cell receptor. *J. Biol. Chem.* 2007;282(10):7405-7415.

360. Fowler JC, Burgoyne LA, Scott AC, Harding HW. Repetitive deoxyribonucleic acid (DNA) and human genome variation–a concise review relevant to forensic biology. *J. Forensic Sci.* 1988;33(5):1111-1126.

361. Venter JC, Adams MD, Myers EW, et al. The sequence of the human genome. *Science*. 2001;291(5507):1304-1351.

362. Wolfsberg TG, McEntyre J, Schuler GD. Guide to the draft human genome. *Nature*. 2001;409(6822):824-826.

363. Rothman N, Skibola CF, Wang SS, et al. Genetic variation in TNF and IL10 and risk of non-Hodgkin lymphoma: a report from the InterLymph Consortium. *Lancet Oncol.* 2006;7(1):27-38.

364. Skibola CF, Bracci PM, Nieters A, et al. Tumor necrosis factor (TNF) and lymphotoxin-alpha (LTA) polymorphisms and risk of non-Hodgkin lymphoma in the InterLymph Consortium. *Am. J. Epidemiol.* 2010;171(3):267-276.

365. Skibola CF, Bracci PM, Halperin E, et al. Genetic variants at 6p21.33 are associated with susceptibility to follicular lymphoma. *Nat. Genet.* 2009;41(8):873-875.

366. Conde L, Halperin E, Akers NK, et al. Genome-wide association study of follicular lymphoma identifies a risk locus at 6p21.32. *Nat. Genet.* 2010;42(8):661-664.

367. Smedby KE, Foo JN, Skibola CF, et al. GWAS of follicular lymphoma reveals allelic heterogeneity at 6p21.32 and suggests shared genetic susceptibility with diffuse large B-cell lymphoma. *PLoS Genet.* 2011;7(4):e1001378.

368. Cozen W, Li D, Best T, et al. A genome-wide meta-analysis of nodular sclerosing Hodgkin lymphoma identifies risk loci at 6p21.32. *Blood*. 2012;119(2):469-475.

369. Urayama KY, Jarrett RF, Hjalgrim H, et al. Genome-wide association study of classical Hodgkin lymphoma and Epstein-Barr virus status-defined subgroups. *J. Natl. Cancer Inst.* 2012;104(3):240-253.

370. Chaperot L, Delfau-Larue MH, Jacob MC, et al. Differentiation of antitumor-specific cytotoxic T lymphocytes from autologous tumor infiltrating lymphocytes in non-Hodgkin's lymphomas. *Exp. Hematol.* 1999;27(7):1185-1193.

371. Montes-Ares O, Moya-Quiles MR, Montes-Casado M, et al. Human leucocyte antigen-C in B chronic lymphocytic leukaemia. *Br. J. Haematol.* 2006;135(4):517-519.

372. Alcoceba M, Marín L, Balanzategui A, et al. The presence of DRB1*01 allele in multiple myeloma patients is associated with an indolent disease. *Tissue Antigens*. 2008;71(6):548-551.

373. Nathalang O, Tatsumi N, Hino M, et al. HLA class II polymorphism in Thai patients with non-Hodgkin's lymphoma. *Eur. J. Immunogenetics Off. J. Br. Soc. Histocompat. Immunogenetics*. 1999;26(6):389-392.

374. Juszczynski P, Kalinka E, Bienvenu J, et al. Human leukocyte antigens class II and tumor necrosis factor genetic polymorphisms are independent predictors of non-Hodgkin lymphoma outcome. *Blood*. 2002;100(8):3037-3040.

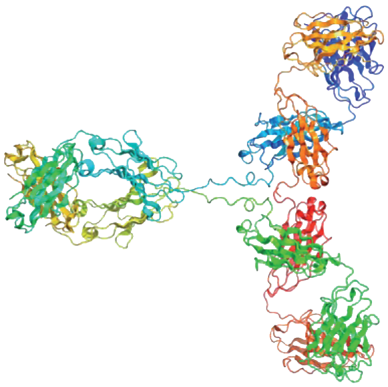
375. Nowak J, Kalinka-Warzocha E, Juszczy ski P, et al. Haplotype-specific pattern of association of human major histocompatibility complex with non-Hodgkin's lymphoma outcome. *Tissue Antigens*. 2008;71(1):16-26.

376. Wang SS, Abdou AM, Morton LM, et al. Human leukocyte antigen class I and II alleles in non-Hodgkin lymphoma etiology. *Blood*. 2010;115(23):4820-4823.

377. Lu Y, Abdou AM, Cerhan JR, et al. Human leukocyte antigen class I and II alleles and overall survival in diffuse large B-cell lymphoma and follicular lymphoma. *ScientificWorldJournal*. 2011;11:2062-2070.

Summary

(Resumen de la Tesis de Inglés - Doctorado Europeo)



1. Introduction

Diffuse large B-cell lymphoma (DLBCL) is the most common aggressive B-cell non-Hodgkin lymphoma (NHL). The combination of the anti-CD20 monoclonal antibody rituximab with anthracycline-based combined chemotherapy (R-CHOP, rituximab, with cyclophosphamide, doxorubicin, vincristine and prednisone) has led to the cure of more than half of patients with DLBCL. Although DLBCL is a curable neoplasm, even in advanced stages, up to one-third of patients have refractory disease or relapse after treatment. The diversity of clinical presentation, as well as the pathological and biological heterogeneity, suggests that DLBCL comprises several disease entities that might ultimately benefit from different therapeutic approaches.

DLBCL is a neoplasm defined as a diffuse proliferation of large neoplastic mature B-cells[1]. Morphological, biological and clinical studies have subdivided DLBCL into morphological variants, molecular and immunophenotypic subgroups. However, a large number of cases remain that may be biologically heterogeneous, which are classified as DLBCL, not otherwise specified (DLBCL NOS). Lymph nodes exhibit a diffuse proliferation of large lymphoid cells that have partially or, more commonly, totally effaced the architecture. Three common variants have been recognized: centroblastic, immunoblastic and anaplastic. Neoplastic cells express pan B-cell markers such as CD19, CD20, CD22 and CD79a, but may lack one or more of these.

DLBCL NOS constitutes 30% of all NHLs. It is more common in the elderly. The median age of appearance is in the seventh decade but it may also occur in children and young adults. It is slightly more common in males than in females. The etiology of DLBCL NOS remains unknown. It usually arises *de novo* (referred to as primary) but can represent progression or transformation (referred to as secondary) of a less aggressive lymphoma, e.g., chronic lymphocytic leukemia (CLL/SLL), follicular lymphoma, marginal zone lymphoma or nodular lymphocyte-predominant Hodgkin lymphoma (NLPHL). Underlying immunodeficiency is a significant risk factor. DLBCL NOS occurring in the setting of immunodeficiency is more often Epstein Barr virus (EBV)-positive.

Patients may present with nodal or extranodal disease, up to 40% being at least confined to extranodal sites[2]. The most common extranodal site is the gastrointestinal tract. Other common sites of extranodal presentation include bone, testis, spleen,

Waldeyer ring, salivary gland, thyroid, liver, kidney and adrenal gland. Bone marrow involvement is reported in 11-27% of cases. The detection rate of minimal involvement may be increased if ancillary studies using, for example, immunohistochemistry or PCR are added to the morphological evaluation of the bone marrow specimen[3]. Bone marrow involvement may be as typical DLBCL (concordant)[4], but is more commonly discordant and linked to the presence of low-grade B-cell lymphoma in the bone marrow.

Patients usually present with a rapidly enlarging tumor mass at single or multiple nodal or extranodal sites. Almost half of the patients have stage I or II disease. In the pre-rituximab era, the long-term remission rate was between 50 and 60%. Although individual prognoses vary widely, the International Prognosis Index (IPI), which is based on clinical parameters, has proved to be highly valuable. However, IPI seems to lose some of its predictive value in patients treated with rituximab[5], which has significantly improved prognosis [6].

Gene expression profile and immunostaining

DLBCL is a clinically and biologically heterogeneous disease. Gene expression profiling (GEP) has enabled two major biological subgroups of DLBCL to be characterized: the germinal center B-cell-like (GCB), which usually has a favorable prognosis, in contrast to the activated B-cell-like (ABC), which is more frequently associated with a poor outcome[7,8]. However, GEP is not used in standard practice because of its complexity and cost. Immunostaining approaches have been proposed as feasible surrogates for GEP because they allow DLBCL to be classified into GCB or non-GCB subtypes[9,10], although discrepant results have been noted[11]. The prognostic value of the immunostaining classification is even more doubtful in the rituximab era because the addition of this drug to conventional chemotherapy has modified the prognostic capacity of already established prognostic factors.

Genetic abnormalities in DLBCL

Genetic abnormalities are other potentially applicable prognostic indicators. Traditional methods have frequently identified *BCL2, BCL6* and *MYC* gene abnormalities in

DLBCL, but only the latter has been consistently associated with poor prognosis in these patients[12–14]. Moreover, some recent studies have revealed that the *MYG-IG* rearrangement alone, and not the *MYC-non-IG* rearrangement, is a negative predictor of survival in DLBCL[15]. Secondly, the role of *MYC* in prognosis is also especially relevant in the 20% of cases with *MYC* breaks that have concurrent *BCL2* and/or *BCL6* abnormalities: the so-called double-hit lymphomas[16,17].

Next-generation sequencing (NGS) has identified many mutated genes in DLBCL that are involved in chromatin remodeling (*MLL2* and *EZH2*), immunoresponse (*B2M*), BCR activation (*CD79B, MYD88*), B-cell differentiation (*BCL6, BLIMP1*) and cell survival (*TP53, TNFRSF14*)[18–21]. The distribution of these mutations differs between DLBCL subtypes. *TP53* mutations occur in 18-30% of DLBCLs[19,21,22] and have been associated with poor prognosis in DLBCL by several groups[23,24], particularly in the small fraction of GC lymphomas that harbor such mutations[25]. Other studies, however, have failed to demonstrate any association between *TP53* mutations and prognosis[26,27]. Furthermore, the *MYD88* L265P mutation, which is found in 12% of DLBCLs[21,28], is preferentially associated with the ABC (29%) subtype rather than the GC subtype (6%)[19,29,30]. However, these frequencies do not help the final distinction between ABC and GC types, so alternatives need to be found.

In recent years, copy number alterations (CNAs) detected by comparative genomic hybridization or high-density SNP arrays have been examined in lymphoproliferative disorders. CNA analysis of DLBCL tumor samples has revealed a number of genomic alterations with frequencies differing between DLBCL subgroups[31,32]. According to these studies, most ABC-DLBCLs have *BCL2* amplifications as well as *CDKN2A* and 6q21–q22 deletions, whereas GCB-DLBCLs frequently harbor *REL* locus amplifications, *PTEN* deletions, apart from the t(14;18), which is present in around 20% of cases. Moreover, ABC DLBCLs frequently present abnormalities characteristically associated with NF- B pathway activation, similar to primary mediastinal B-cell lymphomas[33].

All these findings are challenged because it is difficult to distinguish somatically acquired tumor changes from germinally inherited copy number variations or homozygosity segments. Paired studies with matched tumor and normal DNA samples are superior to those with multiple unpaired samples since they avoid the false discoveries of SNP

array analysis[34], including aneuploidy and copy number neutral-loss of heterozygosity (CNN-LOH), which have been detected in cancer-free individuals[35]. High-resolution CN studies of DLBCL have recently been published[32,36–38]. Unfortunately, studies analyzing paired samples (neoplastic and germinal) are scarce and involve small number of patients. Moreover, molecular inversion probe (MIP) technology has recently been optimized for copy number analysis of highly degraded formalin-fixed, paraffin-embedded (FFPE) samples. The assay performance has been extensively validated with archived FFPE samples[39].

Immunoglobulin gene rearrangements in diffuse large B-cell lymphoma

DLBCL comprises large B-cells harboring a clonal rearrangement of immunoglobulin (IG) genes. The rearrangement of the immunoglobulin heavy chain (IGH) gene occurs during B-cell ontogeny in an ordered fashion, leading to the assembly of distinct variable (IGHV), diversity (IGHD) and joining (IGHJ) genes[40]. During B-cell differentiation, rearrangements start with an IGHD gene joining to an IGHJ gene, forming a partial IGH-D-IGHJ (D-J) rearrangement; a process that can simultaneously occur in both alleles. This is followed by the joining of an IGHV gene to the partial D-J rearrangement, forming a completely functional IGHV-IGHD-IGHJ (V-D-J) rearrangement in one allele, but can also retain a partial non-functional rearrangement in the other allele. Once a rearrangement has become completely functional, the naïve B-cell can react to antigens in the germinal center (GC) of the secondary lymphoid organs, undergoing affinity maturation by the processes of somatic hypermutation (SHM) and class switch recombination.

The molecular features of IGH rearrangements may provide important information about the ontogeny of B-cells. Thus, the intrinsic mutability of IGHV genes has been widely analyzed[41–43]. Since normal B-cells undergo SHM as a response to antigen, classically within GC, the pattern of SHM is thought to provide clues about the mutation machinery and the role of antigen selection. Non-random, potentially antigen-driven SHM has been described in normal B-cells, as well as in chronic lymphocytic leukemia (CLL) and classic hairy cell leukemia (HCL), which has been called canonical SHM[44–48]. However, there are reports suggesting that certain B-cells may accumulate IGHV gene mutations in an alternative T cell-independent fashion[49–51]. The similarities or differences of mutations acquired through an alternative pathway have not yet been defined. Several

models[52–54] have been used for quantifying antigen selection, but results may not be reliably interpretable[43]. Since precise tools are still evolving[54], canonical mutation criteria remain as an available method for establishing whether a set of sequences is affected by similar processes and selections.

Moreover, IGHV gene usage in IGH rearrangements has been analyzed in normal B-cells[55,56], in autoimmune diseases[57] and in many B-cell lymphoproliferative disorders (B-LPDs)[48,58–64]. Certain IGHV genes have been closely associated with some pathogens[65] or linked to autoimmunity[66]. Particular antigens have been identified as possibly being responsible for lymphomagenesis or defined as a secondary event[64]. The involvement of particular antigens in lymphomagenesis may be supported by the identification of stereotyped groups with the closely similar complementarity-determining region 3 (VH CDR3)[47]. The stereotyped B-cell receptors have been studied extensively in CLL[59,67,68], revealing a clinical correlation[69], and more recently, in other lymphomas[61–64].

Although DLBCL is the most prevalent aggressive lymphoma, few studies have characterized the BCR in this entity, and most are based on small series of patients or in selected populations[70–77]. In addition, partial D-J rearrangements, the pattern and distribution of mutations, and the existence of stereotyped sequences in DLBCL are still not well defined[75].

Genetic polymorphisms in DLBCL: the human leukocyte antigen (HLA) specificities

Several genetic polymorphisms have been associated with susceptibility or prognosis in various B-cell non-Hodgkin lymphoma (B-NHL) subtypes[78,79]. In recent years, genome-wide association studies (GWASs) have identified 6p21.3 as a risk region for susceptibility of different lymphomas, such as follicular lymphoma (FL)[79–81] or Hodgkin lymphoma[82,83]. The human leukocyte antigen (HLA) system, located in this region, plays a key role in antitumor immune responses and lymphoma-cell apoptosis[84], and may be essential for neoplasia control. Previous studies have shown a relationship between HLA polymorphisms and susceptibility to certain hematological malignancies such as CLL, multiple myeloma (MM) and acute lymphoblastic leukemia (ALL)[83,85,86]. However, there is little information about the relationship between HLA polymorphisms and susceptibility to developing B-NHL or their outcomes[87–91]. Focusing on DLBCL, some associations be-

tween HLA specificities and this B-NHL subtype have been described, as well as with other genetic polymorphisms. Shorter progression-free survival (PFS) and overall survival (OS) have been observed in DLBCL patients lacking the HLA-DR2 or carrying TNF-308A[88]. In addition, OS is shorter in those patients carrying the C*07-B*08-LTA+252G-TNF-308A[89] or the HLA-C*07:01[91] haplotype. However, these studies considered patients treated before the introduction of the anti-CD20 monoclonal antibody, rituximab. Since, at present, almost all patients with DLBCL receive rituximab combinations as first-line therapy, the current influence of HLA polymorphisms on DLBCL susceptibility and outcome remains to be elucidated.

In summary, through the present doctoral work, we have attempted to broaden the knowledge about the biology and pathogenesis of diffuse large B-cell lymphoma. We have studied HLA polymorphisms and CNAs in non-tumor samples of patients with DLBCL, to evaluate their possible role in the development and prognosis of DLBCL. We describe new biological markers in non-tumor cells that could be involved in the development and prognosis of DLBCL.

We also studied the tumor cells from DLBCL patients, analyzing complete and partial IGH rearrangements in the tumor samples from 165 DLBCL to improve our understanding of the biology of this disease. And the CNAs in another 60 tumor DLBCL samples in order to find out the alterations that could be related with the subgroup of DLBDL showing a bad prognosis.

2. Working Hypothesis

Diffuse large B-cell lymphoma (DLBCL) is an aggressive lymphoma that accounts for about 40% of B-cell non-Hodgkin lymphomas (NHLs). Most of them are classified as DLBCL NOS, according to WHO 2008.

Although multiple genetic studies have been published in DLBCL, its ontogeny and pathogenesis are still not well understood. The molecular diagnosis of DLBCL is based on the common clonal origin of all the tumor cells. The exhaustive characterization of immunoglobulin gene rearrangements could improve our understanding of the biology of this type of lymphoma, in addition to the possible influence of certain antigens in the selection of a tumor clone. Few studies have characterized the BCR in this entity, and most are based on small series of patients or in selected populations. Genome-wide gene expression profile (GEP) studies have allowed DLBCL NOS to be separated into two subtypes according to the cell of origin: DLBCL of the germinal center B-cell (GCB) and activated B-cell (ABC). Although GEP has been in use for over a decade, its cost and methodology are such that it has not been implemented in routine practice for the diagnosis and treatment of our patients. Immunohistochemical methods allow DLBCL NOS to be to classified in non-GCB and GCB, and have been proposed as alternatives to GEP, although the correlation is variable and is of dubious prognostic value in the rituximab era.

In recent years, with the introduction of the NGS techniques, new gene mutations have been discovered that appear in specific histological subtypes. Nevertheless, neither a pathognomonic mutation in DLBCL nor a pattern of gene mutation that could differentiate the two subtypes has been found. Moreover, all these findings, although valuable, have been unable to identify patients with a poor prognosis. The main prognostic factors remain clinical, but the International Prognosis Index (IPI) does not identify the subset of patients who are refractory or who relapse after therapy. The presence of a translocation of MYC and BCL2 and/or BCL6 simultaneously —double- or triple-hit lymphoma— has been described as an adverse prognostic factor in several publications.

In addition, studies of comparative genomic hybridization (CGH) arrays, to identify chromosomal abnormalities, and more recently with single nucleotide polymorphism (SNP) arrays or molecular inversion probe (MIP) arrays can detect deletions, gain or loss of heterozygosity in small regions of the genome up to 100 kb or 4 MB, respectively.

These studies are improving our understanding of DLBCL, describing CNAs that affect genes involved in important signaling pathways that sometimes cannot be identified by other methods. CNA analysis could have an important role, as it is not as expensive as NGS, and can be performed on FFPE tissues. Finally, since CAN analysis has the advantage over immunohistochemistry of being more objective, we plan to use it as an alternative or complementary immunohistochemical technique that could be employed routinely.

The important role of immunity, specifically with respect to the HLA system, is well known in DLBCL. Several alterations have been described in this system that enable tumor cells to escape immune surveillance. Furthermore, there is a lack of information about whether genetic or environmental factors could lead to an increased risk of developing DLBCL.

In this context, the hypotheses of this thesis are:

a. In the first study, our aim was to undertake the detailed immunogenetic analysis of the complete and partial immunoglobulin rearrangements, to increase our knowledge about the pathogenesis, origin and etiology of DLBCL. We also set out to identify the different characteristics of the B-cell receptor (BCR) in the two subtypes of DLBCL NOS, GCB and non-GCB, and the possible existence of stereotyped sequences that may involve antigens implicated in the pathogenesis of DLBCL.

b. In the second study, our hypothesis was to describe CNAs that could be used as an alternative or as complementary to the immunohistochemical classification of DLBCL, thereby validating the existing immunohistochemical algorithms. We also proposed to identify specific CNAs that might be correlated with the prognosis of DLBCL for the purpose of identifying patients with poor prognosis, who it has not been possible to identify so far by clinical scores and biological classifications. It would be important to identify those patients who could benefit from receiving treatment with new molecules in clinical trials.

c. In the third study, the hypothesis was to identify host factors related to the HLA system that affect the development and prognosis of DLBCL.

3. Aims

AIM 1: FIRST PAPER: MOLECULAR CHARACTERIZATION OF IMMUNOGLOBULIN GENE REARRANGEMENTS IN TUMOR CELLS IN DLBCL SAMPLES

To analyze the characteristics of complete and incomplete heavy chain immunoglobulin rearrangements in 165 patients with DLBCL NOS, to improve our understanding of the biology of this disease.

a. To characterize the B-cell receptor (BCR) in DLBCL NOS:

- Immunoglobulin heavy chain (IGH) gene usage.
- Pattern and distribution of somatic mutations.
- To define the possible existence of stereotyped sequences of immunoglobulins in DLBCL and the role of antigenic stimulation in the development of DLBCL.
- To estimate the correlation of IGH characteristics with immunohistochemical subtypes of DLBCL.
- To analyze the possible clinical implications of BCR in DLBCL.
- To compare IGH rearrangements in DLBCL with other B-cell lymphoproliferative disorders (B-LPDs) and normal cells.

b. To analyze incomplete DJ rearrangements to determine frequency, gene usage and mutational status.

AIM 2: SECOND PAPER: COPY NUMBER ALTERATIONS (CNAs) IN PAIRED (NORMAL-TUMOR) SAMPLES OF DLBCL PATIENTS.

- To analyze the CNAs and copy neutral number loss of heterozygosity with (CNN-LOH) in paired tumor and non-tumor samples from 60 DLBCLs: clinical and biological relationships.
- To predict the subtype of DLBCL according to cell of origin (GCB *vs.* non-GCB) and prognostic signatures using specific CNAs.

- To identify patients with a poor prognosis, based on specific CNAs, to provide new targeted therapies to these patients.
- To identify variants in non-tumor cells of DLBCL patients that could have a role in the development of this disease.

AIM 3: THIRD PAPER: ASSESS THE IMPORTANCE OF HOST GENETIC FACTORS (HLA POLYMORPHISMS) IN THE DEVELOPMENT AND EVOLUTION OF THE DISEASE (INCIDENCE AND PROGNOSIS)

To evaluate whether HLA class I (-A, -B and -C) and class II (-DRB1 and -DQB1) polymorphisms are associated with lymphoma incidence or outcome in a large series of DLBCL patients, including 153 cases treated with CHOP-like regimens plus rituximab as first-line therapy.

- Analyze the major histocompatibility complex (MHC) and human leukocyte antigen (HLA) specificities in our series of 250 patients with DLBCL.
- To correlate HLA specificities with the development and prognosis of DLBCL.

4. RESULTS

The results presented in this doctoral thesis have been published in peer-reviewed scientific journals. These papers, and some supplemental material are included in this thesis.

PAPER 1: Molecular Characterization of Immunoglobulin Gene Rearrangements in Diffuse Large B-Cell Lymphoma: Antigen-Driven Origin and IGHV4-34 as a Particular Subgroup of the Non-CGB Subtype.

This work was published in the American Journal Pathology: Molecular Characterization of Immunoglobulin Gene Rearrangements in Diffuse Large B-Cell Lymphoma: Antigen-Driven Origin and IGHV4-34 as a Particular Subgroup of the Non-CGB Subtype. American Journal of Pathology. 2012 Nov;181(5):1879-88.

Background

Pathogenesis of diffuse large B-cell lymphoma (DLBCL) remains partially unexplained. Analysis of the B-cell receptor of the malignant cells could contribute to a better understanding of the DLBCL biology.

Design and methods

In the first block we studied the molecular features of the immunoglobulin heavy chain (IGH) rearrangements in 165 patients diagnosed with DLBCL NOS. Clonal IGH rearrangements were amplified according to the BIOMED-2 protocol and PCR products were sequenced directly. We also analyzed the criteria for stereotyped patterns in all complete IGHV-IGHD-IGHJ (V-D-J) sequences.

Patient characteristics and immunohistochemistry

A total of 165 patients with DLBCL were selected retrospectively on the basis of a diagnosis of *de novo* DLBCL NOS according to the 2008 WHO classification[1]. Median age at diagnosis was 59 years, with a range from 15 to 86 years. Fifty-one percent of the patients were male, and 48% of the total series had an International Prognostic Index score ≥3. To distinguish DLBCL biological subtypes, the Hans algorithm[9] was used as a feasible

surrogate for GEP[9,10], which allowed patients to be separated into GCB DLBCL (38%) and non-GCB DLBCL (62%). A total of 125 patients received treatment with rituximab-based regimens.

Material and methods

Tumor DNA was extracted from fresh or formalin-fixed, paraffin-embedded (FFPE) samples collected at the time of diagnosis. All samples were tested for the amplification of IGH rearrangements according to the BIOMED-2 Concerted Action protocols, in whose standardization our group has participated actively[92]. Under this strategy, complete V-D-J rearrangement amplification was performed by multiplex PCR with a set of family-specific primers of the framework region 1 (FR1) and one IGHJ consensus primer. In the samples in which no amplification was detected for FR1, PCR was carried out for FR2. Partial D-J rearrangements were also amplified. The presence of the monoclonal rearrangement was then confirmed by GeneScan with an ABI 3130 DNA Sequencer (Applied Biosystems, Foster City, CA, USA). PCR products were sequenced directly, including forward and reverse reads, using Big-Dye terminators (Applied Biosystems, Foster City, CA, USA).

Germline IGHV, IGHD and IGHJ genes from complete V-D-J rearrangements were identified using the IMGT/V-QUEST database (http://www.imgt.org). The following information was extracted: IGHV, IGHD and IGHJ gene usage, the percentage of IGHV identity with the closest germline gene, and VH CDR3 length[93] and composition. Partial D-J rearrangements were analyzed using the BLAST database.

Gene usage was compared with previously described patterns in normal B lymphocytes and in other B-LPDs for complete[48,55,58–62] and partial[60,94–96] rearrangements.

Analysis of somatic hypermutation. Percentage identity with the closest IGHV germline

The use of 98% as the percentage identity cut-off value has customarily been employed for the clinical[97] consideration of a patient as mutated or unmutated based on the exclusion of potential genomic polymorphisms[97]. However, it has been shown that most sequence differences in the 98-99.99% group correspond to low-level SHM[98]. We therefore

classified the sequences as previously described: “truly unmutated”, “minimal/borderline mutated”, and “highly mutated”.

Nucleotide characteristics of IGHV mutations in DLBCL were studied in order to establish whether the pattern of distribution was consistent with the canonical SHM process[47]. In addition, to determine whether these characteristics were due to selection or to the mutation machinery, the mutations at redundant wobble bases were analyzed as previously reported for CLL and HCL[47,48]. Wobble bases are the third nucleotide of redundant codons. Moreover, mutations at A:T and G:C pairs were studied to determine whether the mutations were targeted with approximately the same frequency as in normal cells[99]. Finally, recurrent nucleotide and amino acid changes in the most common IGHV genes were analyzed in search of evidence of driven SHM in DLBCL.

We analyzed VH CDR3 characteristics. The criteria of stereotyped rearrangements were analyzed following a previously described pattern-based method[68,100] using the ClustalW/X 2.0 software. Subsets were named as previously published[25,59,68] and were compared with previously published stereotyped patterns[25,35] and with DLBCL sequences from public databases.

Statistical analyses were performed using Graph-Pad 4.0 (GraphPad Software, Inc., San Diego, CA, USA) and SPSS (version 15.0; SPSS Inc., Chicago, IL, USA). Differences were considered to be statistically significant for values of $p<0.05$.

Results

Clonal rearrangements

The overall detection rate of monoclonality by amplifying V-D-J and/or D-J rearrangements was 90% (149/165 patients). D-J rearrangements were seen in 68/165 patients (41%). V-D-J rearrangements amplified in 130 out of 165 patients (79%), of which 109 were successfully sequenced. IGHV gene repertoire analysis revealed that IGHV3 and IGHV4 were the predominant subgroups. *IGHV4-34* (15.5%), *IGHV3-23* (9.7%) and *IGHV4-39* (8.7%) were the most frequent genes. Regarding partial IGHD-IGHJ rearrangements, we observed a specific IGHD gene usage, two of the most frequent IGHD genes in V-D-J rearrangements being completely absent from partial rearrangements, suggesting

that the use of certain IGHD genes may favor the rearrangement process to yield a final complete functional rearrangement.

Somatic hypermutation

A total of 91% of productive complete V-D-J rearrangements carried somatic hypermutations (<98% identity with the closest germline IGHV gene). The median degree of SHM was 10.8% (range 0-26%). As expected from the physiopathological point of view, 88% of the partial D-J rearrangements were unmutated. Overall, point mutations predominated. Interestingly, all but one case with the IGHV4-34 gene (15/16) were "highly mutated".

R-mutations were mainly localized in the CDR1 and CDR2 regions, at the 5' and 3' ends of FR2 and the 3' end of FR3. Analysis of canonical SHM was possible in 96 DLBCL sequences with less than 100% identity with the germline sequence. SHM in DLBCL displayed all the established characteristics of canonical SHM, which persisted in the most common IGHV genes: 1) a higher ratio of replacement (R) versus silent (S) mutations in the CDRs than in the FRs, 2.56 *vs.* 1.42 regions; $p<0.01$; 2) base change bias for transitions: transition to transversion ratio was 1.16, higher than the ratio of 0.5 that would be expected if the frequencies of the types of mutation were random ($p<0.01$); and 3) more than the expected percentage of mutations in the RGYW hot spots (R=A/G, Y=C/T, and W=A/T), 24.9% of the DLBCL mutations were located in RGYW hot spots, a significantly higher percentage (14.0%; $p<0.01$) than expected. The bias for transitions over transversions was also reflected by the wobble base mutations (1.20; $p<0.01$), implying that this bias in DLBCL is due to the same mechanism of SHM. Finally, G:C pairs were preferentially targeted for SHM relative to A:T (ratio 1.86). This observation was especially evident in the IGHV4-34 gene (ratio 2.23).

Recurrent amino acid changes were mainly found among IGHV4-34, IGHV4-39 and IGHV1-18 rearrangements. Although the IGHV4-34 genes were highly mutated, all 10 available IGHV4-34 sequences conserved the specific AVY motif at the end of FR1 codons 24-26, and the W at position 7 was intact in the IGHV4-34 cases. By contrast, the N-glycosylation site in CDR2 of IGHV4-34 was lost in 7/16 (44%) cases. Some of these changes may indicate specific driven SHM in DLBCL.

VH CDR3 composition and stereotyped VH CDR3 sequences

The median VH CDR3 length was 15 amino acids (range 7-26). VH CDR3 was significantly longer in the *IGHV4* family than in the other families (p=0.01), and in "truly unmutated" and borderline cases compared with "highly mutated" cases (p=0.01). Three sequences among the 103 patients in our series met the criteria for stereotyped rearrangements and could be included in different subsets. Two sequences clustered within previously described CLL subsets (subsets 5 and 7D)[59,68]; 35 contained the *IGHV1-69* gene and were "truly unmutated". The other case formed a probably realistic novel provisional subset with a public DLBCL[76] sequence: subset A.

Molecular and immunohistochemical correlations

IGHV usage and percentage of SHM differed between the immunohistochemical subgroups. *IGHV1-69* and *IGHV4-34* rearrangements were more frequent in the non-GCB subtype. A higher percentage of SHM was observed in GCB than in non-GCB cases (p<0.01). All unmutated sequences belonged to the non-germinal center B-cell-like (non-GCB) subtype. In contrast, all *IGHV4-34* cases, despite being highly mutated, were of the non-GCB subtype.

Comparative analysis of IGH rearrangements in DLBCLs with other normal B-cells and B-LPDs

In order to highlight the singular characteristics of DLBCLs, we compared our results with those reported for normal B-cells and other B-LPDs. The *IGHV4-34* gene was significantly over-represented in DLBCL relative to normal B-cells and to other B-LPDs such as CLL, and especially to multiple myeloma (MM) and Waldeström macroglobulinemia (WM). However, it was found in a similar proportion of patients with splenic marginal zone lymphoma (SMZL) and mantle cell lymphoma (MCL)[48,51,55,56,58,59,96,101]. By contrast, some genes frequently used in the rearrangements of normal B-cells and other B-LPDs[13,24-28,48,58,59,61,62,96], such as *IGHV3-30*, were completely absent from the present series (p<0.01). Most complete V-D-J rearrangements were mutated, with a similar percentage of SHM to those previously reported in MM, but higher values than in CLL and WM[58,59,96].

Conclusions

Most of the complete V-D-J rearrangements in our series of DLBCL were highly mutated. Three genes, *IGVH4-34*, *IGVH3-23* and *IGVH4-39*, accounted for one third of the whole cohort. In particular, *IGHV4-34* was highly represented (15.5% overall). Interestingly, all *IGHV4-34* rearrangements and all unmutated sequences belonged to the non-germinal center B-cell-like (non-GCB) subtype. Based on particular features in the sequence of the immunoglobulin genes, we suggest that particular subgroups exist within the non-GCB subtype.

Overall, we found three cases that followed the current criteria for stereotyped heavy chain VH CDR3 sequences, two of which belonged to subsets previously described in CLL. Our data imply that the IGHV gene repertoire is remarkably biased and shares recurrent amino acid changes. To the best of our knowledge, ours is the first description of stereotyped sequences in DLBCL and of "common" stereotypes between CLL and other B-LPDs. These results indicate a role for antigen selection in DLBCL development and also suggest possibilities for future investigations into the biology of DLBCL.

PAPER 2: High-resolution copy number analysis of paired normal-tumor samples from diffuse large B-cell lymphoma

This work has been published online: epublication ahead of print in Annals of Hematology: High-resolution copy number analysis of paired normal-tumor samples from diffuse large B cell lymphoma. Annals of Hematology. Received: 13 August 2015 / Accepted: 6 November 2015. Doi: 10.1007/s00277-015-2552-3.

In the second block we studied the CNAs in 60 patients diagnosed with DLBCL NOS and, for the 40 of them for which data were available, compared the results with those from the normal paired sample.

Materials and methods

Patient characteristics

Sixty patients with a diagnosis of *de novo* DLBCL NOS according to the 2008 World Health Organization Classification[1] were included in the study. Paired constitutional DNA

was available from 40 of these patients. Median age at diagnosis was 59 years (range 25-85 years), 57% were male, and 40% had an International Prognostic Index (IPI) score ≥3. Immunostaining classified them as GC DLBCL (42%) and non-GC DLBCL (58%) according to the Hans algorithm[102]. Most patients were treated with rituximab-based regimens, and risk-adapted therapies were used.

The CN and CNN-LOH changes in the germline of DLBCL patients and in a series of 500 healthy Spanish controls were selected for comparison. These controls were recruited from the database of the Hospital Clínico Universitario de Santiago de Compostela, which has registered healthy individuals who have been examined using the SNP Array 6.0 (Affymetrix Inc., Santa Clara, CA, USA).

Methods

Tumor DNA was extracted from diagnostic samples. Sufficient tumor cells were tested in morphological and immunophenotypic studies. DNA from non-tumor matched samples was isolated from non-infiltrated peripheral blood or non-infiltrated bone marrow.

Interphase fluorescence *in situ* hybridization (FISH) analysis was carried out to detect t(14;18)/IGH-BCL2 and t(8;14)/IGH-MYC translocations, and BCL6 breaks.

Sixty DLBCL cases and paired constitutional DNA from 40 of these patients were successfully analyzed using CN arrays. Thirty-eight cases with available FFPE DNA were hybridized on the MIP assay using the Oncoscan FFPE Express custom service 2.0 (Affymetrix). Twenty-two cases with HMW DNA from frozen tissue and 40 normal DNAs were analyzed using the SNP 6.0 array. CNA and presence of CNN-LOH were evaluated by visual inspection by two different observers using Nexus Biodiscovery version 7 software (Biodiscovery, El Segundo, CA, USA). Only gains and losses with a minimum size of 100 kb and CNN-LOH >5 Mb were considered. Tumor alterations were considered germinal or acquired by comparing them with the corresponding constitutional DNA of each sample. In cases in which normal DNA was not available, alterations were cross-checked with copy number variation databases (http://dgv.tcag.ca/dgv/app/).

We studied the presence of *TP53* mutations and MYD88 L265P[28,103], and IGH rearrangements were analyzed using the BIOMED-2 protocols[92,104]. HLA class I and class II

typing was carried out using standardized methods[105]. Phenotype frequencies and HLA polymorphisms were categorized into previously defined supertypes[106,107].

We used IBM SPSS (version 23.0; IBM Corp., Armonk, NY, USA) for statistical analysis. Appropriate cut-off points for CN alterations were calculated using maximally selected rank statistics (maxstat package, R, version 3.0.1). In order to identify those variables independently related to the GC and non-GC groups, a backward stepwise logistic regression, using the Akaike Information Criterion (AIC) as the stop criterion, was performed (glm function, R, version 3.0.1). Survival analysis was performed and all differences were considered to be statistically significant for values of $p<0.05$.

Results

Genetic abnormalities

The t(14;18)/IGH-BCL2 translocation was present in 25% of the cases studied, most of which were from GC DLBCLs. There were also MYC translocations in 4%, and BCL6 breaks in 29% of cases.

Copy number analysis was successful in all 60 cases. No significant differences were observed between the two CN platforms used, so the results were subsequently analyzed together. Major comparisons were based on the Hans classifier (32 non-GC *vs.* 23 GC DLBCLs).

A total of 867 CNAs were identified (mean 14.5 ± 9.6 SD; median 13, range 0-45). All cases except one had CNAs, but only 63 were recurrent CNAs, including 33 gains and 30 losses. There were also nine recurrent regions of acquired CNN-LOH. We found 59 germinal regions of CNN-LOH. No germinal CNAs were observed. Homozygous deletions of 9p21.3 containing CDKN2A and B were detected in three cases, and other homozygous deletions were noted in single cases, such as 6p21 including HLA class II, 6q21 (PRMD1), 15q21 (B2m). In addition, 48 amplifications were detected in these series. The most common amplification peak, found in 13% of the cases, was at 2p15-p16.1 (REL and BCL11A).

There were two cases with a copy number pattern suggesting chromothripsis[108–110], one of which carried the previously described aberrant 11q pattern[111] and double *TP53* inactivation.

CNAs were evaluated with respect to their GC and non-GC DLBCL classification. Several chromosomal alterations were significantly associated with one of the DLBCL subgroups. The logistic regression model predicting the best discrimination between GC and non-GC DLBCL included losses of 1p36.32 and 15q21.1 and gains of chromosome 7 as predictors for the GC subtype, whereas losses of 8p23.3 and gains of 11q24.3 were more specific to the non-GC DLBCL subtype.

The 6p21.3 region was the most frequent copy number-neutral loss of heterozygosity (CNN-LOH) region in the tumors in our series 8/40 (20%). We also found CNN-LOH regions involving genes of importance in lymphoma, such as TNFRSF14 and TP53. The analysis of the 40 paired normal-tumor DLBCL samples revealed 59 regions of germinal CNN-LOH. Sixteen of the 40 cases had germinal CNN-LOH, although the only recurrent regions were found at 6p21 and 9p24.1 (n=2). Germinal CNN-LOH at 17p13.1 including the TP53 gene was seen in only one case. The incidence of the 59 CNN-LOH regions found in 16 normal samples of DLBCL cases was compared with the CNN-LOH regions found in 152 of the 500 normal controls. Thirty CNN-LOH regions were more frequently seen in DLBCL than in healthy controls ($p<0.05$).

Seventeen cases (28%) showed abnormalities (deletions, CNN-LOH and/or mutations) at the *TP53* gene. There were no cases with concomitant *TP53* deletion and mutation, although there were two cases with concurrent CNN-LOH and *TP53* mutations. Notably, the two cases with double *TP53* inactivation were the most genetically complex cases, exhibiting significantly more CNAs than the other tumors ($p<0.01$). The MYD88 L265P mutations were detected in ten cases (17%). Mutations were present in both DLBCL subtypes. Thirty-five of the 38 cases analyzed carried mutations of the IGHV genes. The most common IGHV rearrangements were *IGHV3-23* and *IGHV4-34*. All patients carrying *IGHV4-34* were of the non-GC subtype. No associations were seen between HLA specificities and genetic abnormalities in this series.

Clinical correlations

The median follow-up of the series was 56 months (range 20-211 months), yielding a 5-year progression-free survival (PFS) and overall survival (OS) of 67% and 74%, respectively, for the whole series. PFS and OS were not associated with any clinical or biological parameter, including LDH, 2m, RCP, age, sex and extranodal sites. Similarly, survival was similar in the two immunohistochemical subgroups. As previously reported by our group, patients with the B44 supertype had a worse prognosis than other patients, although the difference was not statistically significant in this series. There were no differences in B44 supertype frequencies between the two immunohistochemical subtypes, although the B44 supertype was particularly related to a worse prognosis in the GC subtype (p=0.016).

Clinical correlations of genetic abnormalities

Patients with 18p11.32 loss had a lower 5-year PFS (p=0.027). By contrast, 11q gains were related to a better prognosis in this series (p=0.04). Patients with biallelic inactivation of 17p13.1 involving the *TP53* gene demonstrated a very low 5-year PFS compared with single allelic inactivation and wild-type cases (p=0.007), which translated into similar numbers for 5-year OS (p=0.014). Similarly, the two patients with a chromothripsis-like copy number pattern had a very poor prognosis, achieving a median PFS and OS of only 3 and 9 months, respectively. On the other hand, gains of 11q24.3, targeting ETS1 and FLI1 genes, and which are more frequently found in non-GC DLBCLs, were associated with a better prognosis within this subtype (5-year OS 100% versus 67%; p=0.042).

Conclusions

In conclusion, copy number analysis could be used as a complementary method to immunohistochemistry for predicting the GCB/non-GCB subtype of DLBCL and for assessing prognosis in DLBCL patients. We identified double-allele inactivation of *TP53* and loss of *METTL4*, which heavily influence PFS and OS, and germline CNN-LOH as potential lymphoma susceptibility factors.

These findings support the use of this method for evaluating genetic abnormalities associated with DLBCL patients whose ultimate clinical value will need to be validated in large and independent series.

PAPER 3: HLA specificities are related to development and prognosis of diffuse large B-cell lymphoma

This work has been published in Blood: HLA specificities are related to development and prognosis of diffuse large B-cell lymphoma. Blood. 2013 Aug 22;122(8):1448-54.

Patients and methods

We analyzed 250 consecutive patients diagnosed with *de novo* DLBCL according to the 2008 WHO classification[1]. Their median age was 56 years, with a range from 13 to 87 years. Fifty-two percent of the patients were male. The age-adjusted International Prognostic Index (aaIPI) score was ≥2 in 58% of the patients. Immunostaining with antibodies to BCL-6, CD10 and MUM1 could be performed, which allowed patients to be classified into two biological subtypes according to the Hans criteria[102]: germinal center B-cell-like (GCB) (45%) and non-GCB (55%) DLBCL. Most of the patients (68%) received chemotherapy with rituximab as first-line treatment and 48% were included in the Spanish Lymphoma/Autologous Bone Marrow Transplant Study Group (GELTAMO) protocols for *de novo* DLBCL patients.

The control population for HLA-A, HLA-B and HLA-DRB1 loci consisted of 1940 healthy donor individuals from our geographical region. In addition, HLA typing was extended to HLA-C and HLA-DQB1 in 200 of these donors.

This study was approved by the local Ethics Review Committee in accordance with Spanish law. Prior informed consent was obtained from all participants.

DNA from peripheral blood samples was isolated. Low-resolution (two-digit) HLA class I (-A, -B and -C) and class II (-DRB1 and -DQB1) typing was carried out using the polymerase chain reaction reverse sequence-specific oligonucleotide and Luminex® XYP technology (Tepnel Lifecodes Corporation, Stamford, CT, USA), in accordance with the

standards of the European Federation of Immunogenetics (http://www.efiweb.org). Many HLA alleles have overlapping structural and functional features through shared peptide-binding pockets and thus can be clustered into distinct functional groups, known as supertypes. In addition to phenotype frequencies, the HLA polymorphisms were categorized into previously defined supertypes.

Definitions and statistical analyses

Allele frequencies were estimated by an expectation-maximization algorithm using the Arlequin software package, version 3.5.1.2. They were tested to see whether they were in Hardy-Weinberg equilibrium. Allele and phenotype frequencies between populations were compared with the two-sided Fisher's exact test using GraphPad Prism 4.0 (GraphPad Software Inc., San Diego, CA, USA). The strength of associations was estimated by odds ratios (ORs), and their 95% confidence intervals (CIs) were calculated by the Cornfield method (values of $p<0.05$ were considered statistically significant). Probabilities were also Bonferroni-corrected (p_c) to take into account multiple testing.

Individual clinical variables, aaIPI and IPI factors, cell of origin according to the Hans classification, immunohistochemistry, and presence of the various HLA polymorphisms were considered. We used SPSS (version 15.0; SPSS Inc., Chicago, IL, USA) for survival analysis. Since rituximab combinations are the current standard first-line treatment for DLBCL patients, univariate and multivariate analyses were restricted to rituximab-treated patients (n=153).

Results

HLA association with DLBCL development

HLA polymorphism frequencies in DLBCL patients appeared to be in Hardy-Weinberg equilibrium in all cases. Phenotypic frequencies differed significantly between DLBCL patients and healthy control individuals.

The phenotypic frequency of HLA-DRB1*01 was significantly higher in DLBCL patients than in the control population (29% *vs.* 19.5%; p=0.0008, OR=1.69, 95% CI=1.25-2.28, p_c=0.0104). In addition, we observed a lower incidence of the HLA-C*03 phenotype

in DLBCL patients compared with controls (6.4% *vs.* 17.9%; p=0.0005, OR=0.31, 95% CI=0.16-0.62, p_c=0.007).

HLA and clinical outcome in the whole series

With a median follow-up of 81 months (range 11-255 months), the actuarial 5-year PFS and 5-year OS in the whole series were 61% and 78%, respectively.

With respect to HLA polymorphisms, patients with HLA-B*18 specificity had a shorter PFS (35% *vs.* 63% at 5 years; p=0.002) and OS (53% *vs.* 82% at 5 years; p=0.0005) compared with the other DLBCL patients. A trend for shorter OS was observed in HLA-B*44 patients (69% *vs.* 81% at 5 years; p=0.08).

Our data revealed that HLA-B*18 and HLA-B*44 seem to influence the clinical outcome of DLBCL. Interestingly, these two specificities are members of the HLA B44-supertype, which includes the HLA-B*18, HLA-B*37, HLA-B*40, HLA-B*41, HLA-B*44, HLA-B*45 and HLA-B*50 alleles. Consequently, we wanted to analyze the role of the HLA B44-supertype in DLBCL prognosis. Patients carrying any HLA allele of the HLA B44-supertype had a shorter OS (69% *vs.* 86% at 5 years; p=0.004) than those without the supertype.

We analyzed the effect of the treatment with respect to HLA B44-supertype. None of the patients carrying the HLA B44-supertype, irrespective of whether they were treated with rituximab, showed significant differences in PFS or OS. By contrast, those patients carrying other HLA alleles who do not receive rituximab-based regimens had shorter PFS and OS than who did receive it.

HLA and clinical outcome in the rituximab group

With a median follow-up of 56 months (range 11-136 months), the actuarial 5-year PFS and 5-year OS of rituximab-treated patients were 65% and 83%, respectively.

As in the whole series, patients with the HLA-B*18 specificity had a poor outcome. Similarly, the HLA-B*44 specificity was associated with significantly shorter PFS (p=0.06) and OS (p=0.02). When we analyzed the outcome of the patients with respect to the

HLA B44-supertype we found a shorter PFS (p=0.019) and OS (p=0.001) in those patients carrying any allele of the HLA B44-supertype.

To determine whether the results obtained with the HLA B44-supertype directly depended on the HLA-B*18 polymorphism, we assessed the effect of HLA in three groups: group a, patients bearing the HLA-B*18 polymorphism (n=15); group b, patients bearing any allele of the HLA B44-supertype except HLA-B*18 (n=51); and group c, patients bearing "other HLA alleles" (n=77). The three groups had different 5-year PFS and 5-year OS.

We then evaluated the prognostic value of IPI in our series. Patients with high IPI exhibited worse PFS and OS, although the differences were not statistically significant. Since most of the patients in our series were younger than 60 years of age, we analyzed the aaIPI. We found that those patients with aaIPI ≥2 had a worse PFS (p=0.018), but the difference in OS was not statistically significant (p=0.09).

In the multivariate analysis, the presence of HLA-B*18 or of any allele of the HLA B44-supertype except HLA-B*18, and a value of aaIPI ≥2 were the only independent variables with a significant influence on PFS. These three variables were also of independent prognostic value for OS.

In conclusion, our results suggest that some HLA polymorphisms influence the development and outcome of DLBCL, allowing the identification of an extremely good-risk prognostic subgroup independently of the aaIPI. However, these results are preliminary and need to be validated in order to exclude a possible population effect.

5. Discussion

Although DLBCL is the most frequent variant of non-Hodgkin lymphoma, its ontogeny is still not well understood. In this thesis we assess different aspects of tumor and non-tumor cells from patients diagnosed with diffuse large B-cell lymphoma (DLBCL) in order to improve our knowledge of the biology, pathogenesis and ontogeny of this disease.

This work focuses on the biology of DLBCL, not otherwise specified (DLBCL NOS). We studied the tumor samples from DLBCL, analyzing the immunoglobulin heavy chain (IGH) rearrangements in 165 tumor samples and assessing the copy number alterations (CNAs) in 60 additional DLBCL patients. Our aim was to identify changes in the tumor cell that could distinguish patients with a worse prognosis, which currently cannot be selected with clinical indexes. Furthermore, since *Alizadeh et al* identified by gene expression profiling (GEP), two main subtypes of DLBCL, which have variable but acceptable immunohistochemical reproducibility in different series. We wanted to analyze both IGH and CNAs in the two immunohistochemical subgroups classified by the Hans algorithm, in order to improve our knowledge of the pathogenesis and cell of origin of the two subtypes. We could observe differences in the IGHV gene usage and pattern of mutations between the GCB and non-GCB subtypes. In addition, we propose that CNA analysis by SNP/MIP arrays could complement immunohistochemistry to enable the proper classification of DLBCL, and to overcome the limitations of the immunohistochemical techniques.

Forty tumor samples analyzed for CNAs could be compared with the normal paired samples, thereby avoiding false-positives in the tumor samples, and detecting CNN-LOH in the non-tumor samples that could have a role in the development of DLBCL.

On the other hand, the HLA system could be involved in susceptibility and disease control. Thus, we analyzed the HLA system in non-tumor samples in 250 DLBCL patients, to assess the influence of HLA polymorphisms on the development and prognosis of DLBCL.

To perform these analysis we applied the methodology described in the different sections, using fresh and formalin-fixed, paraffin-embedded (FFPE) tissue.

Molecular Characterization of Immunoglobulin Gene Rearrangements in Diffuse Large B-Cell Lymphoma: Antigen-Driven Origin and IGHV4-34 as a Particular Subgroup of the Non-GCB Subtype

We attempted to address the origin of DLBCL by analyzing the features of the B-cell receptor, the hallmark of the B lymphocytes. To this end, we analyzed complete V-D-J rearrangements in 165 untreated DLBCLs, which is the largest series reported to date. We also characterized partial D-J rearrangements.

Our results, based on the characterization of the BCR in DLBCLs, show a biased usage of certain IGHV genes and evidence of canonical SHM. The discovery of recurrent amino acid changes suggests a role for antigen selection in the development of DLBCL.

A strong bias at the level of IGHV gene usage was observed, the *IGHV4-34*, *IGHV3-23* and *IGHV4-39* genes accounting for one third of the cohort. In addition, recurrent amino acid changes were found in certain IGHV genes, which may indicate specific driven SHM in DLBCL. This hypothesis is even more plausible if the presence of some stereotyped sequences is taken into account. We identified three VH CDR3 stereotyped sequences in our series, which were present at a lower frequency than in CLL[68]. Two of them shared their *IGHV1-69* rearrangement with subgroups previously described in CLL[59,68]. It is well known that there is a relationship between CLL and DLBCL, since some cases of DLBCL occur within a CLL background (Richter transformation). However, the cases reported here are *de novo* DLBCL, indicating "common" stereotypes between DLBCL and CLL.

This highly restricted immunoglobulin gene repertoire with stereotyped VH CDR3 and common SHM targeting in DLBCL suggests a role for antigen selection in this entity, through the stimulation of the proliferation of B-cells that express surface IG encoded by certain IGHV genes, at least for some subsets of cases. Some of the antigens could be common to various B-LPDs but could ultimately lead to a different entity, depending on which population they stimulate.

Our findings indicate a different IGHV gene distribution in the two immunohistochemical subtypes. We confirmed, in a larger series, the strong association between DLBCL expressing the *IGHV4-34* and the non-GCB subtype. Moreover, most of the cases

showing *IGHV4-34* rearrangements in our series shared several features, suggesting that DLBCL expressing *IGHV4-34* may constitute a separate subgroup within the non-GCB DLBCL subtype. *IGHV4-34* cases belonged to the non-GCB subtype, were highly mutated but carried shared amino acid changes and conserved the critical W at position 7. This specific FR1 motif is thought to be responsible for binding the N-acetyllactosamine antigenic determinant, implying that these cells may retain the ability to bind to and be activated by superantigens, despite intense SHM activity[112]. This autoreactivity has also been demonstrated by *Young et al*[113], supporting our hypothesis and concluding that IGHV regions used by the BCRs of ABC DLBCL samples vary in their ability to sustain survival of these ABC lines. This raises the possibility of developing a screening procedure to identify patients who might benefit from BCR pathway inhibition.

Several hypotheses have been proposed to explain the origin of clonal cells from non-GCB DLBCLs. *Lossos et al*[114] reported that non-GCB DLBCL must derive from B-cells that have passed through the GC with no possibility of additional SHM. However, *Jardin et al*[115] observed ongoing mutations in non-GCB DLBCL, including *IGHV4-34* cases, which suggests an independent GC mechanism of SHM[115]. Regardless of the GC dependence of these mutations, we observed that SHM of the non-GCB subtype met the canonical criteria (*IGHV4-34* included), which do not support the need to invoke different mutational machinery from that involved in the normal process for IGHV mutations.

In conclusion, based on particular features in the sequence of the immunoglobulin genes, we found that different subgroups exist within the non-GCB subtype. Our results support the idea that DLBCL is characterized by a highly distinctive IG gene repertoire. To the best of our knowledge, we have provided the first description of stereotyped sequences in DLBCL and “common” stereotypes shared by CLL and other B-LPDs. The frequency of these stereotyped sequences and their possible biological and clinical significance in DLBCLs needs further analysis in larger, independent series.

These findings suggest a role for antigen selection in DLBCL development and also raise possibilities for future investigation into the biology of DLBCL.

High-resolution copy number analysis of paired normal-tumor samples from diffuse large B-cell lymphoma

By analyzing copy number alterations of DLBCL samples we aimed to improve our knowledge of the changes in tumor cells compared with the normal cells in DLBCL samples, and to establish a relationship between other biological and clinical variables. In addition, we studied the variations in normal samples, which could have a role in the development of DLBCL.

Our study presents an in-depth CNA analysis of 60 DLBCLs with high-resolution CN arrays. Other studies on this subject have been reported, but most of them did not compare the tumor cells and their normal counterparts in enough patients. Our study is the first to analyze 40 paired tumor-normal DLBCL cases that were mainly treated using rituximab. This strategy allows germinal and tumor-acquired CNA and CNN-LOH regions to be distinguished accurately.

The CNAs identified in our study show quite good concordance with the high-frequency aberrations described in other studies[32,116,117]. We report alterations involving regions of importance in lymphomagenesis that are consistent with previously published data. The most frequently observed amplification in our series targeted the 2p16 region, including REL and BCL11. In parallel, we also describe homozygous losses in relevant genes such as 6p21.32 (HLA II) and 15q21 (B2M), which are required for the immune response.

When we studied the CN profiles of the two immunohistochemical subtypes we found some specific aberrations in both subgroups. Indeed, we conclude that a few CNAs could enable the reliable stratification of patients. Specifically, the predictive model identified losses of 1p36.32 containing TNFRSF14, and 15q21.1 including B2M, as well as gains of chromosome 7, as the most powerful CNAs for classifying the GC subtype. By contrast, 8p23.3 losses including ERICH1 and 11q24.3gains, containing ETS1 and FLI1, were more specific to the non-GC subtype. To our knowledge, this is the first time that a predictor of DLBCL subgroups has been derived from CN data. This analysis and classification based on CN arrays in combination with an immunohistochemical classifier could have advantages over RNA-based GEP[7], mainly because the latter method requires very high-quality material that is not currently available from FFPE samples.

We also describe here germinal CNN-LOH in patients diagnosed with DLBCL. This has not been reported before, probably because other methods are not sufficiently sensitive to detect them. The most frequent region of germinal CN-LOH was 6p21 (HLA region), which was seen in two cases. We also found one case with germinal CNN-LOH of 17p including TP53; this CNN-LOH would have been considered to have been acquired in the tumor if we had not analyzed the normal paired sample. The comparison of our patients with the normal Spanish population demonstrated that 6p22.3-21.33 was significantly more frequent in DLBCL than in control cases. These results strongly suggest that the 6p22.3-21.33 region is involved in DLBCL development. This is consistent with the idea that the HLA system, which is essential in immunological surveillance, is related to lymphoma susceptibility.

The prognostic analysis of our series did not identify a clinical or immunohistochemical variable that reliably identified high-risk patients. This may have been due to the small sample size of our series, and because most patients were treated with rituximab, and risk-adapted therapies were used. However, a small subset of very high-risk patients was identified with a few genetic alterations such as biallelic TP53 inactivation and 18p11.32 losses including the *METTL4* gene. In addition, 11q24.3 gains including the *ETS1* and *FLI1* genes identified a good prognostic subgroup of DLBCL patients, suggesting a tumor suppressor role for these genes.

In summary, SNP/MIP-assay array analysis could be used as a complementary method to immunohistochemistry for predicting the GC/non-GC subtype of DLBCL and identifying high-risk DLBCL patients. Some germline CNN-LOH may play a role as potential lymphoma susceptibility factors. These findings support the use of this method for evaluating genetic abnormalities associated with DLBCL patients. Its ultimate clinical value will need to be validated in large and independent series.

HLA specificities are related to development and prognosis of diffuse large B-cell lymphoma

The HLA system plays an essential role in immunological surveillance. The different abilities of certain alleles to present tumor antigens or the deregulation of allelic expression efficiently could result in the tumor cells escaping T- and Natural Killer (NK)-cell recognition. In addition, HLA as an antigen presenting system could have a role in the

development and/or control of some non-hematological and hematological neoplasias, including lymphoma[87–91]. However, specifically in DLBCL, this feature has not been extensively studied. Previous research including survival analyses was limited and performed before the rituximab era, so its conclusions cannot be applied to the current standard of care[89,91].

Accepting this premise, we analyzed its role in diffuse large B-cell lymphoma patients. We studied the influence of the HLA-A, -B, -C, -DRB1 and -DQB1 polymorphisms on the development of DLBCL in 250 patients, the largest series reported to date. In addition, we describe for the first time the role of HLA polymorphisms in the prognosis of DLBCL patients receiving rituximab-based regimens as first-line treatment. Our results show that the presence of the HLA-DRB1*01 polymorphism and the absence of the HLA-C*03 polymorphism are associated with susceptibility to DLBCL, while the HLA B44-supertype, especially with the presence of HLA-B*18 and HLA-B*44 polymorphisms, have an independent influence on survival.

Our findings indicate a higher HLA-DRB1*01 frequency in DLBCL patients. It is of note that a higher incidence of HLA-DRB1*01 has been described in other B-cell lymphoproliferative disorders such as follicular lymphoma[80,90] and smoldering multiple myeloma[86], indicating a possible role for HLA-DRB1*01 in controlling the tumor B-cell clone. Likewise, a higher incidence of this specificity was found in some autoimmune disorders such as rheumatoid arthritis[118,119] and Crohn's disease[118,120], which have also been linked to higher susceptibility to DLBCL[121].

We also found a lower frequency of HLA-C*03 in DLBCL. The HLA-C*03 belongs to the HLA-C group 1 (HLA-C1) and is a ligand of the activating Natural Killer Ig-like receptor (KIR) KIR2DS2 gene, and the inhibitory KIR2DL2 and KIR2DL3 genes[122]. Some studies have related the combination of HLA and KIR genes to susceptibility to or outcome of various infectious and autoimmune disorders[123–126]. Considering these data we hypothesize that those patients carrying HLA-C*03 could develop more efficient antitumor activity by NK cells, probably influenced by their combination with KIR genes. Further analysis of the KIR genotypes combined with HLA typing of DLBCL patients should be done to address this question.

One of the important outcomes of our study is the description of the influence of the HLA B44-supertype (and especially the HLA-B*18 polymorphism), on the prognosis of DLBCL patients. This influence is particularly evident in those patients receiving rituximab-based regimens, and is independent of aaIPI. Remarkably, the HLA B44-supertype and the HLA-B*18 polymorphism can discriminate three risk-prognostic subgroups within patients with an aaIPI >2, where those bearing the HLA-B*18 have extremely poor PFS and OS (0% and 35%, respectively). It is also interesting to note that the presence of the HLA B44-supertype characterizes a group of patients who do not respond to rituximab, and in which novel strategies should be tested.

HLA-B*44 is more easily lost than other HLA polymorphisms in some types of solid tumor[127], which indicates that this allele may have a central role in the mechanisms of immune escape. An alternative explanation for these differences in prognosis is that the HLA B44-supertype does not efficiently present common tumor peptides to the cytolytic T lymphocytes, leading to a poor response. Our data indicate that tumor cells from DLBCL patients with the HLA B44-supertype allele may easily escape immune surveillance, resulting in poor disease control.

In conclusion, our results suggest a role for HLA in DLBCL susceptibility and prognosis, and are consistent with the increasing evidence of the influence of 6p21 on B-NHL development and, in particular, on DLBCL. These findings could be helpful in the future classification, treatment and monitoring of DLBCL patients.

6. Conclusions

FIRST PAPER

1. Diffuse large B-cell lymphomas use variable region of the immunoglobulin heavy chain genes with specific characteristics that may help us distinguish them from other B-cell lymphoproliferative disorders with respect to type (families and VH genes) and other characteristics: the canonical somatic hypermutation process, the degree of mutation, recurrent changes in certain amino acids, and the length of the clonotypic CDR3 region. These findings indicate that, at least in a subset of diffuse large B-cell lymphomas, VH genes play a significant role in the origin and development of these lymphomas, possibly by an antigenic selection mechanism.

2. Stereotypical sequences can be found in diffuse large B-cell lymphomas, as in other B-cell lymphoproliferative disorders, albeit with a lower incidence. This finding, described here for the first time, implies that, at least in some cases, specific antigens exist whose stimulus could be a possible developmental mechanism for these lymphomas.

3. We identify some diffuse large B-cell lymphomas in which the use of a particular *VH* gene is correlated with clinical, biological and histological features. Thus, unmutated forms and lymphomas carrying the *IGHV4-34* gene are of the non-germinal center subtype. These findings suggest that there are particular subgroups within histological subtypes of these lymphomas that need to be validated in independent series in order to confirm their differential characteristics.

SECOND PAPER

1. The study of gene copy number identifies recurrent gains and deletions, and can help us understand the pathogenesis and behavior of diffuse large B-cell lymphomas. However, it is essential to analyze simultaneously tumor and normal samples from the same patient in order to avoid confusion between real changes and polymorphisms, and to identify real losses of heterozygosity in

which the copy number is normal in non-tumor samples. Our study therefore challenges the findings of studies performed without analysis of paired tumor and non-tumor samples.

2. The study of copy number abnormalities helps identify histological subtypes of diffuse large B-cell lymphoma. *β2M* losses and chromosome 7 gains are the most specific changes in the germinal center cell lymphoma, while loss of *ERICH1* and gains of *ETS1* and *FLI1* are typical of the non- germinal center subtype.

3. The 6p21 region, which contains the HLA genes, is a target region for loss of heterozygosity in diffuse large B-cell lymphoma tumor samples. This suggests that the HLA system plays an important role in the pathogenesis and development of these lymphomas through an immune escape mechanism.

4. Our study shows that the analysis of numerical abnormalities enables the prognosis of diffuse large B-cell lymphomas to be refined. Thus, the double inactivation of the *TP53* gene (loss of heterozygosity and mutation) and the loss of *METTL4* are both adverse prognostic factors in these lymphomas; by contrast, *ETS1* and *FLI1* gains are associated with a better prognosis.

THIRD PAPER

1. The HLA system, which is essential to the immune antitumor mechanisms, may play a role in the pathogenesis of diffuse large B-cell lymphoma. The frequency of some HLA specificities differs significantly between patients with these lymphomas and the normal healthy population. Thus, the HLA-DRB1*01 antigen is more frequent in patients with lymphoma, while the HLA-C*03 antigen is more frequent in the healthy population.

2. In diffuse large B-cell lymphoma patients treated with R-CHOP, the presence of the HLA-B44 supertype and, especially with the HLA-B*18 specificity, is associated with a shorter progression-free survival and poorer overall survival. This leads us to hypothesize that tumor cells with these specificities can escape immune recognition more easily.

7. REFERENCES

1. Swerdlow SH, Cancer IA for R on. WHO classification of tumours of haematopoietic and lymphoid tissues: Vol. 2: International Agency for Research on Cancer. WHO; 2008.
2. Harris NL, Jaffe ES, Stein H, et al. A revised European-American classification of lymphoid neoplasms: a proposal from the International Lymphoma Study Group. *Blood*. 1994;84(5):1361-1392.
3. Talaulikar D, Dahlstrom JE, Shadbolt B, et al. Occult bone marrow involvement in patients with diffuse large B-cell lymphoma: results of a pilot study. *Pathology (Phila.)*. 2007;39(6):580-585.
4. Chung R, Lai R, Wei P, et al. Concordant but not discordant bone marrow involvement in diffuse large B-cell lymphoma predicts a poor clinical outcome independent of the International Prognostic Index. *Blood*. 2007;110(4):1278-1282.
5. Sehn LH, Berry B, Chhanabhai M, et al. The revised International Prognostic Index (R-IPI) is a better predictor of outcome than the standard IPI for patients with diffuse large B-cell lymphoma treated with R-CHOP. *Blood*. 2007;109(5):1857-1861.
6. Coiffier B. Rituximab therapy in malignant lymphoma. *Oncogene*. 2007;26(25):3603-3613.
7. Alizadeh AA, Eisen MB, Davis RE, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature*. 2000;403(6769):503-511.
8. Lenz G, Wright G, Dave SS, et al. Stromal gene signatures in large-B-cell lymphomas. *N. Engl. J. Med*. 2008;359(22):2313-2323.
9. Hans CP, Weisenburger DD, Greiner TC, et al. Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. *Blood*. 2004;103(1):275-282.
10. Choi WWL, Weisenburger DD, Greiner TC, et al. A new immunostain algorithm classifies diffuse large B-cell lymphoma into molecular subtypes with high accuracy. *Clin. Cancer Res*. 2009;15(17):5494-5502.
11. Meyer PN, Fu K, Greiner TC, et al. Immunohistochemical methods for predicting cell of origin and survival in patients with diffuse large B-cell lymphoma treated with rituximab. *J. Clin. Oncol*. 2011;29(2):200-207.
12. Savage KJ, Johnson NA, Ben-Neriah S, et al. MYC gene rearrangements are associated with a poor prognosis in diffuse large B-cell lymphoma patients treated with R-CHOP chemotherapy. *Blood*. 2009;114(17):3533-3537.
13. Barrans S, Crouch S, Smith A, et al. Rearrangement of MYC is associated with poor prognosis in patients with diffuse large B-cell lymphoma treated in the era of rituximab. *J. Clin. Oncol*. 2010;28(20):3360-3365.
14. Akyurek N, Uner A, Benekli M, Barista I. Prognostic significance of MYC, BCL2, and BCL6 rearrangements in patients with diffuse large B-cell lymphoma treated with cyclophosphamide, doxorubicin, vincristine, and prednisone plus rituximab. *Cancer*. 2012;118(17): 4173-4183.

15. Copie-Bergman C, Cuillière-Dartigues P, Baia M, et al. MYC-IG rearrangements are negative predictors of survival in DLBCL patients treated with immunochemotherapy: a GELA/LYSA study. *Blood*. 2015;126(22):2466-2474.

16. Ueda C, Nishikori M, Kitawaki T, Uchiyama T, Ohno H. Coexistent rearrangements of c-MYC, BCL2, and BCL6 genes in a diffuse large B-cell lymphoma. *Int. J. Hematol.* 2004;79(1):52-54.

17. Hummel M, Bentink S, Berger H, et al. A biologic definition of Burkitt's lymphoma from transcriptional and genomic profiling. *N. Engl. J. Med.* 2006;354(23):2419-2430.

18. Pasqualucci L, Dominguez-Sola D, Chiarenza A, et al. Inactivating mutations of acetyltransferase genes in B-cell lymphoma. *Nature*. 2011;471(7337):189-195.

19. Pasqualucci L, Trifonov V, Fabbri G, et al. Analysis of the coding genome of diffuse large B-cell lymphoma. *Nat. Genet.* 2011;43(9):830-837.

20. Lohr JG, Stojanov P, Lawrence MS, et al. Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. *Proc. Natl. Acad. Sci. U. S. A.* 2012;109(10):3879-3884.

21. Morin RD, Mungall K, Pleasance E, et al. Mutational and structural analysis of diffuse large B-cell lymphoma using whole-genome sequencing. *Blood*. 2013;122(7):1256-1265.

22. Lossos IS, Morgensztern D. Prognostic biomarkers in diffuse large B-cell lymphoma. *J. Clin. Oncol.* 2006;24(6):995-1007.

23. Ichikawa A, Kinoshita T, Watanabe T, et al. Mutations of the p53 gene as a prognostic factor in aggressive B-cell lymphoma. *N. Engl. J. Med.* 1997;337(8):529-534.

24. Young KH, Weisenburger DD, Dave BJ, et al. Mutations in the DNA-binding codons of TP53, which are associated with decreased expression of TRAILreceptor-2, predict for poor survival in diffuse large B-cell lymphoma. *Blood*. 2007;110(13):4396-4405.

25. Zainuddin N, Berglund M, Wanders A, et al. TP53 mutations predict for poor survival in de novo diffuse large B-cell lymphoma of germinal center subtype. *Leuk. Res.* 2009;33(1): 60-66.

26. Osada M, Ishioka C, Ichinohasama R, et al. Influence of p53 mutation on pathological grade, but not prognosis of non-Hodgkin's lymphoma. *Anticancer. Drug Des.* 1999;14(2): 107-114.

27. Barrans SL, Carter I, Owen RG, et al. Germinal center phenotype and bcl-2 expression combined with the International Prognostic Index improves patient risk stratification in diffuse large B-cell lymphoma. *Blood*. 2002;99(4):1136-1143.

28. Jiménez C, Sebastián E, Chillón MC, et al. MYD88 L265P is a marker highly characteristic of, but not restricted to, Waldenström's macroglobulinemia. *Leukemia*. 2013;27(8):1722-1728.

29. Ngo VN, Young RM, Schmitz R, et al. Oncogenically active MYD88 mutations in human lymphoma. *Nature*. 2011;470(7332):115-119.

30. Zhang J, Grubor V, Love CL, et al. Genetic heterogeneity of diffuse large B-cell lymphoma. *Proc. Natl. Acad. Sci. U. S. A.* 2013;110(4):1398-1403.

31. Bea S, Zettl A, Wright G, et al. Diffuse large B-cell lymphoma subgroups have distinct genetic profiles that influence tumor biology and improve gene-expression-based survival prediction. *Blood.* 2005;106(9):3183-3190.

32. Scholtysik R, Kreuz M, Hummel M, et al. Characterization of genomic imbalances in diffuse large B-cell lymphoma by detailed SNP-chip analysis. *Int. J. Cancer.* 2015;136(5):1033-1042.

33. Lenz G, Staudt LM. Aggressive lymphomas. *N. Engl. J. Med.* 2010;362(15):1417-1429.

34. Heinrichs S, Li C, Look AT. SNP array analysis in hematologic malignancies: avoiding false discoveries. *Blood.* 2010;115(21):4157-4161.

35. Jacobs KB, Yeager M, Zhou W, et al. Detectable clonal mosaicism and its relationship to aging and cancer. *Nat. Genet.* 2012;44(6):651-658.

36. Scandurra M, Mian M, Greiner TC, et al. Genomic lesions associated with a different clinical outcome in diffuse large B-Cell lymphoma treated with R-CHOP-21. *Br. J. Haematol.* 2010;151(3):221-231.

37. Green MR, Aya-Bonilla C, Gandhi MK, et al. Integrative genomic profiling reveals conserved genetic mechanisms for tumorigenesis in common entities of non-Hodgkin's lymphoma. *Genes Chromosomes Cancer.* 2011;50(5):313-326.

38. Monti S, Chapuy B, Takeyama K, et al. Integrative analysis reveals an outcome-associated and targetable pattern of p53 and cell cycle deregulation in diffuse large B cell lymphoma. *Cancer Cell.* 2012;22(3):359-372.

39. Wang Y, Carlton VEH, Karlin-Neumann G, et al. High quality copy number and genotype data from FFPE samples using Molecular Inversion Probe (MIP) microarrays. *BMC Med. Genomics.* 2009;2:8.

40. Tonegawa S. Somatic generation of antibody diversity. *Nature.* 1983;302(5909):575-581.

41. Betz AG, Milstein C, González-Fernández A, et al. Elements regulating somatic hypermutation of an immunoglobulin kappa gene: critical role for the intron enhancer/matrix attachment region. *Cell.* 1994;77(2):239-248.

42. Wagner SD, Milstein C, Neuberger MS. Codon bias targets mutation. *Nature.* 1995;376 (6543):732.

43. Dunn-Walters DK, Dogan A, Boursier L, MacDonald CM, Spencer J. Base-specific sequences that bias somatic hypermutation deduced by analysis of out-of-frame human IgVH genes. *J. Immunol.* 1998;160(5):2360-2364.

44. Smith DS, Creadon G, Jena PK, et al. Di- and trinucleotide target preferences of somatic mutagenesis in normal and autoreactive B cells. *J. Immunol.* 1996;156(7):2642-2652.

45. Spencer J, Dunn M, Dunn-Walters DK. Characteristics of sequences around individual nucleotide substitutions in IgVH genes suggest different GC and AT mutators. *J. Immunol.* 1999;162(11):6596-6601.

46. Shapiro GS, Aviszus K, Murphy J, Wysocki LJ. Evolution of Ig DNA sequence to target specific base positions within codons for somatic hypermutation. *J. Immunol.* 2002;168(5):2302-2306.

47. Messmer BT, Albesiano E, Messmer D, Chiorazzi N. The pattern and distribution of immunoglobulin VH gene mutations in chronic lymphocytic leukemia B cells are consistent with the canonical somatic hypermutation process. *Blood.* 2004;103(9):3490-3495.

48. Arons E, Roth L, Sapolsky J, et al. Evidence of canonical somatic hypermutation in hairy cell leukemia. *Blood.* 2011;117(18):4844-4851.

49. Weller S, Faili A, Garcia C, et al. CD40-CD40L independent Ig gene hypermutation suggests a second B cell diversification pathway in humans. *Proc. Natl. Acad. Sci. U. S. A.* 2001;98(3): 1166-1170.

50. Chiorazzi N, Ferrarini M. B cell chronic lymphocytic leukemia: lessons learned from studies of the B cell antigen receptor. *Annu. Rev. Immunol.* 2003;21:841-894.

51. Berkowska MA, Driessen GJA, Bikos V, et al. Human memory B cells originate from three distinct germinal center-dependent and -independent maturation pathways. *Blood.* 2011;118(8):2150-2158.

52. Chang B, Casali P. The CDR1 sequences of a major proportion of human germline Ig VH genes are inherently susceptible to amino acid replacement. *Immunol. Today.* 1994;15(8): 367-373.

53. Lossos IS, Tibshirani R, Narasimhan B, Levy R. The inference of antigen selection on Ig genes. *J. Immunol.* 2000;165(9):5122-5126.

54. Uduman M, Yaari G, Hershberg U, et al. Detecting selection in immunoglobulin sequences. *Nucleic Acids Res.* 2011;39(Web Server issue):W499-504.

55. Brezinschek HP, Foster SJ, Brezinschek RI, et al. Analysis of the human VH gene repertoire. Differential effects of selection and somatic hypermutation on human peripheral CD5(+)/IgM+ and CD5(-)/IgM+ B cells. *J. Clin. Invest.* 1997;99(10):2488-2501.

56. Wu Y-C, Kipling D, Leong HS, et al. High-throughput immunoglobulin repertoire analysis distinguishes between human IgM memory and switched memory B-cell populations. *Blood.* 2010;116(7):1070-1078.

57. Collins AM, Sewell WA, Edwards MR. Immunoglobulin gene rearrangement, repertoire diversity, and the allergic response. *Pharmacol. Ther.* 2003;100(2):157-170.

58. González D, van der Burg M, García-Sanz R, et al. Immunoglobulin gene rearrangements and the pathogenesis of multiple myeloma. *Blood.* 2007;110(9):3112-3121.

59. Stamatopoulos K, Belessi C, Moreno C, et al. Over 20% of patients with chronic lymphocytic leukemia carry stereotyped receptors: Pathogenetic implications and clinical correlations. *Blood.* 2007;109(1):259-270.

60. Martín-Jiménez P, García-Sanz R, Balanzategui A, et al. Molecular characterization of heavy chain immunoglobulin gene rearrangements in Waldenström's macroglobulinemia and IgM monoclonal gammopathy of undetermined significance. *Haematologica.* 2007;92(5):635-642.

61. Hadzidimitriou A, Agathangelidis A, Darzentas N, et al. Is there a role for antigen selection in mantle cell lymphoma? Immunogenetic support from a series of 807 cases. *Blood*. 2011;118(11):3088-3095.

62. Bikos V, Darzentas N, Hadzidimitriou A, et al. Over 30% of patients with splenic marginal zone lymphoma express the same immunoglobulin heavy variable gene: ontogenetic implications. *Leukemia*. 2012;26(7):1638-1646.

63. Zibellini S, Capello D, Forconi F, et al. Stereotyped patterns of B-cell receptor in splenic marginal zone lymphoma. *Haematologica*. 2010;95(10):1792-1796.

64. Dagklis A, Ponzoni M, Govi S, et al. Immunoglobulin gene repertoire in ocular adnexal lymphomas: hints on the nature of the antigenic stimulation. *Leukemia*. 2012;26(4):814-821.

65. Chan CH, Hadlock KG, Foung SK, Levy S. V(H)1-69 gene is preferentially used by hepatitis C virus-associated B cell lymphomas and by normal B cells responding to the E2 viral antigen. *Blood*. 2001;97(4):1023-1026.

66. Cappione AJ, Pugh-Bernard AE, Anolik JH, Sanz I. Lupus IgG VH4.34 antibodies bind to a 220-kDa glycoform of CD45/B220 on the surface of human B lymphocytes. *J. Immunol.* 2004;172(7):4298-4307.

67. Murray F, Darzentas N, Hadzidimitriou A, et al. Stereotyped patterns of somatic hypermutation in subsets of patients with chronic lymphocytic leukemia: implications for the role of antigen selection in leukemogenesis. *Blood*. 2008;111(3):1524-1533.

68. Agathangelidis A, Darzentas N, Hadzidimitriou A, et al. Stereotyped B-cell receptors in one-third of chronic lymphocytic leukemia: a molecular classification with implications for targeted therapies. *Blood*. 2012;119(19):4467-4475.

69. Tobin G, Thunberg U, Johnson A, et al. Chronic lymphocytic leukemias utilizing the VH3-21 gene display highly restricted Vlambda2-14 gene use and homologous CDR3s: implicating recognition of a common antigen epitope. *Blood*. 2003;101(12):4952-4957.

70. Hsu FJ, Levy R. Preferential use of the VH4 Ig gene family by diffuse large-cell lymphoma. *Blood*. 1995;86(8):3072-3082.

71. Küppers R, Rajewsky K, Hansmann ML. Diffuse large cell lymphomas are derived from mature B cells carrying V region genes with a high load of somatic mutation and evidence of selection for antibody expression. *Eur. J. Immunol.* 1997;27(6):1398-1405.

72. Rosenquist R, Lindström A, Holmberg D, Lindh J, Roos G. V(H) gene family utilization in different B-cell lymphoma subgroups. *Eur. J. Haematol.* 1999;62(2):123-128.

73. Lossos IS, Okada CY, Tibshirani R, et al. Molecular analysis of immunoglobulin genes in diffuse large B-cell lymphomas. *Blood*. 2000;95(5):1797-1803.

74. Nakamura N, Kuze T, Hashimoto Y, et al. Analysis of the immunoglobulin heavy chain gene variable region of CD5-positive and -negative diffuse large B cell lymphoma. *Leukemia*. 2001;15(3):452-457.

75. Capello D, Martini M, Gloghini A, et al. Molecular analysis of immunoglobulin variable genes in human immunodeficiency virus-related non-Hodgkin's lymphoma reveals implications for disease pathogenesis and histogenesis. *Haematologica*. 2008;93(8):1178-1185.

76. Ruminy P, Etancelin P, Couronné L, et al. The isotype of the BCR as a surrogate for the GCB and ABC molecular subtypes in diffuse large B-cell lymphoma. *Leukemia*. 2011;25 (4):681-688.

77. Catherwood MA, Gonzalez D, Patton C, et al. Improved clonality assessment in germinal centre/post-germinal centre non-Hodgkin's lymphomas with high rates of somatic hypermutation. *J. Clin. Pathol.* 2007;60(5):524-528.

78. Rothman N, Skibola CF, Wang SS, et al. Genetic variation in TNF and IL10 and risk of non-Hodgkin lymphoma: a report from the InterLymph Consortium. *Lancet Oncol.* 2006;7(1):27-38.

79. Skibola CF, Bracci PM, Halperin E, et al. Genetic variants at 6p21.33 are associated with susceptibility to follicular lymphoma. *Nat. Genet.* 2009;41(8):873-875.

80. Conde L, Halperin E, Akers NK, et al. Genome-wide association study of follicular lymphoma identifies a risk locus at 6p21.32. *Nat. Genet.* 2010;42(8):661-664.

81. Smedby KE, Foo JN, Skibola CF, et al. GWAS of follicular lymphoma reveals allelic heterogeneity at 6p21.32 and suggests shared genetic susceptibility with diffuse large B-cell lymphoma. *PLoS Genet.* 2011;7(4):e1001378.

82. Cozen W, Li D, Best T, et al. A genome-wide meta-analysis of nodular sclerosing Hodgkin lymphoma identifies risk loci at 6p21.32. *Blood*. 2012;119(2):469-475.

83. Urayama KY, Jarrett RF, Hjalgrim H, et al. Genome-wide association study of classical Hodgkin lymphoma and Epstein-Barr virus status-defined subgroups. *J. Natl. Cancer Inst.* 2012;104(3):240-253.

84. Chaperot L, Delfau-Larue MH, Jacob MC, et al. Differentiation of antitumor-specific cytotoxic T lymphocytes from autologous tumor infiltrating lymphocytes in non-Hodgkin's lymphomas. *Exp. Hematol.* 1999;27(7):1185-1193.

85. Montes-Ares O, Moya-Quiles MR, Montes-Casado M, et al. Human leucocyte antigen-C in B chronic lymphocytic leukaemia. *Br. J. Haematol.* 2006;135(4):517-519.

86. Alcoceba M, Marín L, Balanzategui A, et al. The presence of DRB1*01 allele in multiple myeloma patients is associated with an indolent disease. *Tissue Antigens*. 2008;71(6):548-551.

87. Nathalang O, Tatsumi N, Hino M, et al. HLA class II polymorphism in Thai patients with non-Hodgkin's lymphoma. *Eur. J. Immunogenetics* 1999;26(6):389-392.

88. Juszczynski P, Kalinka E, Bienvenu J, et al. Human leukocyte antigens class II and tumor necrosis factor genetic polymorphisms are independent predictors of non-Hodgkin lymphoma outcome. *Blood*. 2002;100(8):3037-3040.

89. Nowak J, Kalinka-Warzocha E, Juszczy ski P, et al. Haplotype-specific pattern of association of human major histocompatibility complex with non-Hodgkin's lymphoma outcome. *Tissue Antigens*. 2008;71(1):16-26.

90. Wang SS, Abdou AM, Morton LM, et al. Human leukocyte antigen class I and II alleles in non-Hodgkin lymphoma etiology. *Blood*. 2010;115(23):4820-4823.

91. Lu Y, Abdou AM, Cerhan JR, et al. Human leukocyte antigen class I and II alleles and overall survival in diffuse large B-cell lymphoma and follicular lymphoma. *ScientificWorldJournal*. 2011;11:2062-2070.

92. van Dongen JJM, Langerak AW, Brüggemann M, et al. Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936. *Leukemia*. 2003;17(12):2257-2317.

93. Lefranc M-P, Pommié C, Ruiz M, et al. IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains. *Dev. Comp. Immunol.* 2003;27(1):55-77.

94. Tsakou E, Agathangelidis A, Agathagelidis A, et al. Partial versus productive immunoglobulin heavy locus rearrangements in chronic lymphocytic leukemia: implications for B-cell receptor stereotypy. *Mol. Med.* 2012;18:138-145.

95. González D, González M, Balanzategui A, et al. Molecular characteristics and gene segment usage in IGH gene rearrangements in multiple myeloma. *Haematologica*. 2005;90(7):906-913.

96. Martín-Jiménez P, García-Sanz R, González D, et al. Molecular characterization of complete and incomplete immunoglobulin heavy chain gene rearrangements in hairy cell leukemia. *Clin. Lymphoma Myeloma*. 2007;7(9):573-579.

97. Hamblin TJ, Davis Z, Gardiner A, Oscier DG, Stevenson FK. Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. *Blood*. 1999;94(6):1848-1854.

98. Davis ZA, Orchard JA, Corcoran MM, Oscier DG. Divergence from the germ-line sequence in unmutated chronic lymphocytic leukemia is due to somatic mutation rather than polymorphisms. *Blood*. 2003;102(8):3075.

99. Di Noia JM, Neuberger MS. Molecular mechanisms of antibody somatic hypermutation. *Annu. Rev. Biochem*. 2007;76:1-22.

100. Darzentas N, Hadzidimitriou A, Murray F, et al. A different ontogenesis for chronic lymphocytic leukemia cases carrying stereotyped antigen receptors: molecular and computational evidence. *Leukemia*. 2010;24(1):125-132.

101. Klein U, Rajewsky K, Küppers R. Human immunoglobulin (Ig)M+IgD+ peripheral blood B cells expressing the CD27 cell surface antigen carry somatically mutated variable region genes: CD27 as a general marker for somatically mutated (memory) B cells. *J. Exp. Med.* 1998;188(9):1679-1689.

102. Hans CP, Weisenburger DD, Greiner TC, et al. Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. *Blood*. 2004;103(1):275-282.

103. Gonzalez D, Martinez P, Wade R, et al. Mutational status of the TP53 gene as a predictor of response and survival in patients with chronic lymphocytic leukemia: results from the LRF CLL4 trial. *J. Clin. Oncol.* 2011;29(16):2223-2229.

104. Sebastián E, Alcoceba M, Balanzategui A, et al. Molecular characterization of immunoglobulin gene rearrangements in diffuse large B-cell lymphoma: antigen-driven origin and IGHV4-34 as a particular subgroup of the non-GCB subtype. *Am. J. Pathol.* 2012;181(5): 1879-1888.

105. Alcoceba M, Sebastián E, Marín L, et al. HLA specificities are related to development and prognosis of diffuse large B-cell lymphoma. *Blood*. 2013;122(8):1448-1454.

106. Sidney J, Peters B, Frahm N, Brander C, Sette A. HLA class I supertypes: a revised and updated classification. *BMC Immunol.* 2008;9:1.

107. Greenbaum J, Sidney J, Chung J, et al. Functional classification of class II human leukocyte antigen (HLA) molecules reveals seven different supertypes and a surprising degree of repertoire sharing across supertypes. *Immunogenetics*. 2011;63(6):325-335.

108. Rausch T, Jones DTW, Zapatka M, et al. Genome sequencing of pediatric medulloblastoma links catastrophic DNA rearrangements with TP53 mutations. *Cell*. 2012;148(1-2):59-71.

109. Edelmann J, Holzmann K, Miller F, et al. High-resolution genomic profiling of chronic lymphocytic leukemia reveals new recurrent genomic alterations. *Blood*. 2012;120(24):4783-4794.

110. Zhang C-Z, Leibowitz ML, Pellman D. Chromothripsis and beyond: rapid genome evolution from complex chromosomal rearrangements. *Genes Dev.* 2013;27(23):2513-2530.

111. Salaverria I, Martin-Guerrero I, Wagener R, et al. A recurrent 11q aberration pattern characterizes a subset of MYC-negative high-grade B-cell lymphomas resembling Burkitt lymphoma. *Blood*. 2014;123(8):1187-1198.

112. Li Y, Spellerberg MB, Stevenson FK, Capra JD, Potter KN. The I binding specificity of human VH 4-34 (VH 4-21) encoded antibodies is determined by both VH framework region 1 and complementarity determining region 3. *J. Mol. Biol.* 1996;256(3):577-589.

113. Young RM, Wu T, Schmitz R, et al. Survival of human lymphoma cells requires B-cell receptor engagement by self-antigens. *Proc. Natl. Acad. Sci. U. S. A.* 2015;112(44):13447-13454.

114. Lossos IS, Alizadeh AA, Eisen MB, et al. Ongoing immunoglobulin somatic mutation in germinal center B cell-like but not in activated B cell-like diffuse large cell lymphomas. *Proc. Natl. Acad. Sci. U. S. A.* 2000;97(18):10209-10213.

115. Jardin F, Sahota SS, Ruminy P, et al. Novel Ig V gene features of t(14;18) and t(3;14) de novo diffuse large B-cell lymphoma displaying germinal center-B cell like and non-germinal center-B cell like markers. *Leukemia*. 2006;20(11):2070-2074.

116. Monti S, Chapuy B, Takeyama K, et al. Integrative analysis reveals an outcome-associated and targetable pattern of p53 and cell cycle deregulation in diffuse large B cell lymphoma. *Cancer Cell.* 2012;22(3):359-372.

117. Kreisel F, Kulkarni S, Kerns RT, et al. High resolution array comparative genomic hybridization identifies copy number alterations in diffuse large B-cell lymphoma that predict response to immuno-chemotherapy. *Cancer Genet.* 2011;204(3):129-137.

118. International MHC and Autoimmunity Genetics Network, Rioux JD, Goyette P, et al. Mapping of multiple susceptibility variants within the MHC region for 7 immune-mediated diseases. *Proc. Natl. Acad. Sci. U. S. A.* 2009;106(44):18680-18685.

119. Holoshitz J. The rheumatoid arthritis HLA-DRB1 shared epitope. *Curr. Opin. Rheumatol.* 2010;22(3):293-298.

120. Newman B, Silverberg MS, Gu X, et al. CARD15 and HLA DRB1 alleles influence susceptibility and disease localization in Crohn's disease. *Am. J. Gastroenterol.* 2004;99(2):306-315.

121. Ekström Smedby K, Vajdic CM, Falster M, et al. Autoimmune disorders and risk of non-Hodgkin lymphoma subtypes: a pooled analysis within the InterLymph Consortium. *Blood.* 2008;111(8):4029-4038.

122. Colonna M, Brooks EG, Falco M, Ferrara GB, Strominger JL. Generation of allospecific natural killer cells by stimulation across a polymorphism of HLA-C. *Science.* 1993;260(5111):1121-1124.

123. Yen JH, Moore BE, Nakajima T, et al. Major histocompatibility complex class I-recognizing receptors are disease risk genes in rheumatoid arthritis. *J. Exp. Med.* 2001;193(10):1159-1167.

124. Martin MP, Nelson G, Lee J-H, et al. Cutting edge: susceptibility to psoriatic arthritis: influence of activating killer Ig-like receptor genes in the absence of specific HLA-C alleles. *J. Immunol.* 2002;169(6):2818-2822.

125. Martin MP, Gao X, Lee J-H, et al. Epistatic interaction between KIR3DS1 and HLA-B delays the progression to AIDS. *Nat. Genet.* 2002;31(4):429-434.

126. Khakoo SI, Thio CL, Martin MP, et al. HLA and NK cell inhibitory receptor genes in resolving hepatitis C virus infection. *Science.* 2004;305(5685):872-874.

127. Cabrera T, Maleno I, Lopez-Nevot MA, et al. High frequency of HLA-B44 allelic losses in human solid tumors. *Hum. Immunol.* 2003;64(10):941-950.