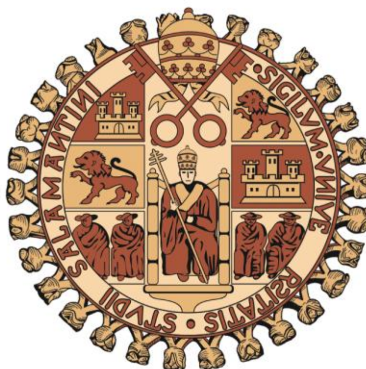
UNIVERSIDAD DE SALAMANCA

DEPARTAMENTO DE MEDICINA

HEMATOLOGÍA



VNIVERSIDAD DE SALAMANCA

TESIS DOCTORAL

IDENTIFICACIÓN DE NUEVOS SUBGRUPOS CITOGENÉTICOS EN LEUCEMIA LINFÁTICA CRÓNICA MEDIANTE EL ANÁLISIS COMBINADO DEL GENOMA, EXOMA Y TRANSCRIPTOMA

ANA EUGENIA RODRÍGUEZ VICENTE

2012

D. Jesús María Hernández Rivas, Doctor en Medicina, Profesor Titular del Departamento de Medicina de la Universidad de Salamanca y Médico Adjunto del Servicio de Hematología del Hospital Clínico Universitario de Salamanca.

CERTIFICA:

Que el trabajo realizado bajo mi dirección por D.ª Ana Eugenia Rodríguez Vicente titulado "IDENTIFICACIÓN DE NUEVOS SUBGRUPOS CITOGENÉTICOS EN LEUCEMIA LINFÁTICA CRÓNICA MEDIANTE EL ANÁLISIS COMBINADO DEL GENOMA, EXOMA Y TRANSCRIPTOMA", reúne las condiciones de originalidad requeridas para optar al grado de Doctor por la Universidad de Salamanca.

Y para que así conste, firmo la siguiente certificación en Salamanca, a uno de Junio de dos mil doce.

Fdo. Dr. Jesús Mª Hernández Rivas

Indice

Artículos científicos

Abstract

Introducción

Leucemia linfática crónica

Biología y fisiopatología

Factores pronósticos

Alteraciones citogenéticas

Perfil de expression génico

Cinética de la LLC

microRNAs

Técnicas de análisis genómico global

La tecnología de microarrays

Técnicas de secuenciación masiva

Hipótesis

Objetivos

Resultados

- **Artículo I**: El análisis combinado del genoma y del perfil de expresión de los enfermos con LLC ha permitido la identificación de una nueva ganancia recurrente en 20q13.
- **Artículos II y III**: Los enfermos con LLC y deleción de 13q presentan diferentes características clínicas y biológicas.
- Análisis genético de la LLC por secuenciación masiva mediante captura de secuencias.
- **Artículo IV**: El polimorfismo rs2307842 en la región 3´UTR de HSP90B1 provoca su sobreexpresión en los linfocitos B de los enfermos con LLC.

Discusión general

Conclusiones

Bibliografía

Esta tesis doctoral se basa en los siguientes artículos científicos, referenciados en el texto como se indica a continuación:

I. Rodríguez AE, Robledo C, García JL, González M, Gutiérrez NC, Hernández JA, Sandoval V, García de Coca A, Recio I, Risueño A, Martín-Núñez G, García E, Fisac R, Conde J, de Las Rivas J, Hernández JM. **Identification of a novel recurrent gain on 20q13 in chronic lymphocytic leukemia by array CGH and gene expression profiling.** Ann Oncol. 2012 Jan 6. [Epub ahead of print].

II. Hernández JA*, Rodríguez AE*, González M, Benito R, Fontanillo C, Sandoval V, Romero M, Martín-Núñez G, de Coca AG, Fisac R, Galende J, Recio I, Ortuño F, García JL, de las Rivas J, Gutiérrez NC, San Miguel JF, Hernández JM. **A high number of losses in 13q14 chromosome band is associated with a worse outcome and biological differences in patients with B-cell chronic lymphoid leukemia**. Haematologica. 2009 Mar;94(3):364-71.

*shared first authorship

III. Rodríguez AE, Hernández JA, Benito R, Gutiérrez NC, García JL, Hernández-Sánchez M, Risueño A, Sarasquete ME, Fermiñán E, Fisac R, García de Coca A, Martín-Núñez G, de las Heras N, Recio I, Gutiérrez O, de las Rivas J, González M, Hernández-Rivas JM. **Molecular characterization of chronic lymphocytic leukemia patients with a high number of losses in 13q14.** Submitted.

IV. Rodríguez AE, Quwaider D, Benito R, Misiewicz-Krzeminska I, Hernández-Sánchez M, Sarasquete ME, Gutiérrez NC, Hernández-Rivas JM. **A common polymorphism at miRNA-223 binding site in 3´untranslated region deregulates HSP90B1 expression in chronic lymphocytic leukemia.** Submitted.

Chronic lymphocytic leukemia (CLL) is a heterogeneous disease with marked variability in the clinical course. The disease is also characterized by the presence of a genetical heterogeneity that is becoming apparent through studies of immunoglobulin heavy chain gene, chromosomal aberrations, microRNA deregulation and genetic abnormalities identified by whole genome sequencing. Furthermore, a strong relationship between specific genetic aberrations and the clinical course of CLL is observed.

Recent advances in genomics have transformed the research on hematologic malignancies by improving molecular approaches to gene networks. The recognition of novel molecular variables identified by the use of high-throughput molecular analytical techniques could contribute to a better knowledge of the pathogenesis of the disease and to the development of more accurate biological predictive factors. Thus microarrays have emerged as powerful tools for increasing the potential of standard methods through genome-wide biological studies. Array-CGH analysis has shown great promise as a tool for the analysis of genetic alterations in complex cancer genomes providing, in a single experiment, a general view of genomic gains and losses. Therefore the use of aCGH may detect new genetic lesions in CLL. Microarray expression profiling of CLL has been used to define the patterns of gene expression related to different clinical outcomes and chromosomal abnormalities. The global views of gene expression, when coupled with available knowledge about gene function, also provide the basis for formulating testable hypothesis about specific cellular pathways of CLL. Moreover, recently, the application of next generation sequencing technology has identified previously unknown recurrent mutations in CLL.

The general aim of this thesis was to analyze and to characterize the heterogeneity of CLL patients by applying high-throughput genomic technologies to gain knowledge in the molecular characteristics of CLL and their possible influence on the disease. First we analyzed the genomic imbalances in CLL cytogenetic subsets from both genomic and gene expression perspectives to identify new recurrent alterations. Overall, the results confirmed FISH studies for the regions frequently involved in CLL and also defined a new recurrent gain on chromosome 20q, in 19% (13/67) of the CLL patients. Genomic analysis identified the minimum region of gain on 20q13.12 of ~2.31 Mb. This gain is associated with great genomic complexity and results in an overexpression of the genes located on this loci. Our

results suggest that the diversity of genomic aberrations in CLL is much greater than previously suggested.

We then focused our attention on CLL patients displaying 13q deletion. This cytogenetic subgroup is classically associated with the better outcome in CLL, although this statement is now under discussion and 13q- CLL patients could be a clinically heterogeneous group. The present study demonstrated that the number of malignant cells carrying this genetic lesion strongly correlates with the disease outcome. Thus patients with a high proportion (>80%) of 13q- cells had both a shorter overall survival and a shorter time to first therapy than those of patients with <80% 13q- cells. To better understand the molecular mechanisms responsible for this different outcome, a molecular characterization of these patients by gene expression profiling and microRNAs studies was carried out. The results demonstrated that 13q- patients are also a biologically heterogeneous group, in which a high number of 13q- cells could involve the deregulation of relevant cellular pathways, such as apoptosis, BCR and NFkB signaling and miRNA regulation. Moreover, our results suggest that some of the biological characteristics of CLL patients with higher percentages of 13q- cells are similar to those of high-risk cytogenetic subgroups (17p- and 11q-), since they share the deregulation of several key signaling pathways, which highlights the fact that CLL patients carrying higher percentages of 13q- cells have more aggressive clinical course. Whether a high load of 13q- cells has an independent prognostic value in CLL or it is a marker of progression remains unsolved.

Finally we performed a mutational study in CLL patients by Next-Generation sequencing techniques. In this study, we detected a common 4 bp deletion SNP (rs2307842) in 25% of CLL patients, which disrupts the binding site for miR-223 in HSP90B1 3´UTR. We demonstrated that this SNP alters the regulation of HSP90B1 expression in CLL, leading to its overexpression only in B lymphocytes. Indirectly, this work provides novel information about how the downregulation of miR-223 can be determining the poor outcome of CLL patients, maybe through upregulation of HSP90B1 expression. Thus, our results could be a first step towards elucidation of miR-223 and HSP90B1 roles in CLL and further investigations in the field are needed.

In conclusion our results show that the combination of global genome and transcriptome analysis, such as microarrays and next-generation sequencing technologies, represents a

useful strategy for the study and characterization of patients with CLL. It allows the identification of novel genetic alterations, even those present in a low number of patients and supports the concept of CLL as a heterogeneous disease, both from the clinical and the molecular point of view.

Leucemia linfática crónica

La leucemia linfática crónica (LLC) es la leucemia más frecuente en el mundo occidental y representa el 30% de las neoplasias de tipo B. La edad media de los pacientes en el momento del diagnóstico es de 65 años, con <15% por debajo de los 50, y suele ser más frecuente en varones.[1,2]

El diagnóstico de la LLC se basa en un cuadro clínico compatible (linfocitosis periférica igual o superior a $5x10^9$/L, con o sin adenopatías) y en la demostración del origen clonal de los linfocitos de la sangre periférica mediante el estudio inmunofenotípico realizado por citometría de flujo,[3] que permite establecer el diagnóstico diferencial con otros síndromes linfoproliferativos B leucémicos, como la tricoleucemia y los linfomas no hodgkinianos B con expresión periférica (principalmente folicular, manto y marginal).

La LLC se caracteriza por una heterogeneidad clínica manifiesta, fundamentalmente debida a las características genéticas de sus células.[1,4-6] Presenta un curso clínico muy variable: algunos enfermos viven pocos años, mientras que en otros la supervivencia es de décadas o incluso tienen una expectativa de vida normal. Esta heterogeneidad clínica ha planteado si la LLC la conforman varias entidades o es realmente solo una con diferente comportamiento. A este respecto, los estudios genéticos, moleculares e inmunofenotípicos de los últimos años han supuesto un avance muy importante en el estudio biológico de la enfermedad. A pesar de estos avances, la patogénesis y las causas del variable comportamiento biológico de la LLC aún necesitan una definición más precisa.

Biología de la LLC

Tradicionalmente se ha considerado que la LLC era una enfermedad clínicamente heterogénea con un origen celular homogéneo, debida a la acumulación de linfocitos B maduros por una apoptosis defectuosa. Sin embargo, este concepto ha cambiado, y actualmente la LLC se considera una enfermedad tanto de proliferación como de acumulación de células B clonales, que pueden diferir en la activación, en el estado de maduración o en el subgrupo celular.[4,6] Además existen diferencias en la morfología, el inmunofenotipo, la citogenética y las características moleculares. Esta heterogeneidad se traduce en una evolución clínica y respuesta al tratamiento variables. En la Tabla 1 se recoge una visión de las modificaciones cronológicas que se han producido en la interpretación de los conceptos básicos de la LLC.

Table 1. Comparison of historical and current views of CLL biology

Historical view		Current view
CLL is a clinically heterogeneous disease with a homogeneous celllar origin	→	CLL is a clinically heterogeneous disease originating from B lymphocytes that may differ in activation, maturity or cellular subgroup
CLL is a disease derived from naive B lymphocytes	→	CLL is a disease derived from antigen-experienced B lymphocytes that differ in the level of immunoglobulin variable gene mutations
Leukemic-cell accumulation occurs because of an inherent apoptotic defect involving the entire mass of leukemic	→	Leukemic cell accumulation occurs because of survival signals from the external enviroment
CLL is a disease of lymphocyte accumulation	→	CLL is a disease of lymphocyte accumulation with a higher associated level of proliferation than was previously recognized
Prognostic markers identify patients at various risk levels (low, intermediate or high in the Rai staging categories, and A, B or C in the Binet categories) with an acknowledged heterogeneity in clinical outcomes among patients in the low- and intermediate-risk categories	→	New molecular biomarkers are used in both diagnosis and prognosis to better assess patients

A nivel morfológico, en la LLC predominan los linfocitos de pequeño tamaño y aspecto maduro (aunque una minoría pueden ser prolinfocitos), con escaso citoplasma y núcleos redondos con cromatina condensada y nucléolos evidentes. Estos linfocitos B se desarrollan en la médula ósea pasando por diferentes etapas de maduración, cada una de ellas caracterizada por el estado de recombinación de los genes de las cadenas pesadas y ligeras de las inmunoglobulinas y de la expresión de ciertos marcadores de superficie.[7] Las inmunoglobulinas son proteínas heterodiméricas compuestas de dos cadenas pesadas (H) y dos ligeras (L). Cada cadena H y L contienen dominios V que definen los sitios específicos de unión al antígeno del anticuerpo, y un dominio constante, el cual tiene función efectora sin cambio de especificidad antigénica. Uno de los parámetros genéticos moleculares más importantes para establecer el pronóstico de los pacientes con LLC es el estado mutacional de la región variable (V) del gen de inmunoglobulinas.[8,9] Inicialmente se consideró que la LLC se originaba a partir de células B *naive* o del centro pregerminal. Sin embargo, en la mitad de los casos como mínimo, la célula original es un linfocito B de memoria que pasa a través del centro germinal: en aproximadamente la mitad de los casos con LLC se observan mutaciones somáticas de la región VH del gen de la cadena pesada de las Igs y este hecho predice la evolución de la enfermedad. La definición del patrón mutado *vs* no mutado reside en un punto de corte, definido arbitrariamente, por una homología mayor del 98% (no mutados) respecto al gen de la línea germinal. Los pacientes con LLC y patrón no mutado presentan un curso desfavorable con progresión más rápida y supervivencia claramente menor. Además se ha observado que, independientemente del estado mutacional, el reordenamiento de algunas regiones específicas VH se asocia a características clínicas y geográficas diferentes[10]

Los linfocitos de la LLC se caracterizan por presentar un inmunofenotipo definido por tres características principales: (1) comparten antígenos B (expresión de los antígenos CD19, CD20, CD22 y reactividad habitual con CD23) con fuerte positividad del antígeno T CD5, en ausencia de otros marcadores T; (2) expresan sólo una cadena ligera de las inmunoglobulinas (κ o λ) como marcador de clonalidad y (3) tienen una baja densidad de inmunoglobulinas de superficie (SIg). Estos elementos son adecuados para un diagnóstico preciso de la LLC y muy importantes para establecer el diagnóstico diferencial con otros síndromes linfoproliferativos (SLP) crónicos.

Factores pronósticos

Debido a la variabilidad clínica evolutiva de los pacientes con LLC,[11,12] el estudio de los factores pronósticos tiene gran importancia para diferenciar grupos de riesgo. Los sistemas de estadificación clínica de Rai y Binet constituyen los índices más utilizados en la práctica para establecer el pronóstico. Sin embargo, no pueden predecir el curso individual de los pacientes en estadios iniciales. Por eso en los últimos años se han definido diversos factores pronósticos basados en estudios biológicos y moleculares, que se podrían clasificar como “clásicos” y “biológicos” (Tabla 2). Los factores pronósticos clásicos suelen ser aquellos que se extraen de la clínica, el examen físico y las pruebas analíticas, mientras que los factores biológicos suelen referirse a los aspectos moleculares de la enfermedad.

Tabla 2. Factores pronósticos en la LLC.

Classical prognostic markers
Clinical stage
Blood lymphocyte count
Lymphocytes morphology in peripheral blood
Blood lymphocyte doubling time
Bone marrow infiltration degree (aspirate/biopsy)
Biological prognostic markers
Extensively studied
Serum markers
IGVH mutational status
V3-21 gene usage
Cytogenetics
CD38 expression
ZAP-70 expression
Requiring further studies

Chromosomal translocations
CLLU1 expression
microRNA signature
TCL-1 gene
Bcl2/Bax ratio
MDR1/MDR3 genes
Activation-induce cytidine deaminase (AID) mRNA
Lipoprotein lipase A expression
ADAM29 expression
VEGF
Thrombopoietin
Telomere lenght and telomerase activity
CD49d
CD69
FCRL
Treatment related
Response to therapy (minimal residual disease status after therapy)

Alteraciones citogenéticas

En la LLC se produce una inestabilidad genética que facilita la presencia de alteraciones cromosómicas, aunque los mecanismos por los que se producen se desconocen[13]. Nuestro conocimiento de las alteraciones genéticas en la LLC se basó inicialmente en los estudios de citogenética convencional, utilizando TPA como mitógeno. Estos estudios son poco rentables en las LLC debido al bajo índice mitótico de sus células. Con esta técnica se puede detectar una clonalidad del 40-50%, y las alteraciones observadas son, fundamentalmente, pérdidas en 13q y la trisomía del cromosoma 12.[14] Sin embargo, algunos estudios recientes sugieren que el uso de oligonucleótidos como factores estimulantes puede aumentar el número de metafases clonales en la LLC.[15-18] Así, después de la estimulación de los linfocitos de la LLC con oligonucleótidos e IL2, el porcentaje de clonalidad es comparable al conseguido en los estudios de FISH.

A pesar de ello, en la actualidad, el estudio mediante citogenética convencional en la LLC está en entredicho porque los estudios de FISH permiten demostrar la presencia de clonalidad en la mayoría de los enfermos con LLC (80%) y definir subgrupos con diferente pronóstico.[19] A finales del año 2000, el grupo de la Universidad de Ülm publicó la asociación de diversas alteraciones citogenéticas determinadas mediante FISH con la supervivencia[19] (Figura 1).

Utilizando un panel de 8 sondas de FISH, se observó que el 82% de los pacientes presentaba alteraciones cromosómicas, algunas de las cuales tenían relevancia pronóstica.[20,21] Los cambios más frecuentes (el 55% de los casos) correspondían a la pérdida de 13q14, seguidas de la pérdida de 11q22-23 (18%), la trisomía del cromosoma 12 (16%), la pérdida de 17p13 (7%) y la pérdida de 6q21 (6%). Sólo el 18 % de los enfermos no tenían alteraciones por FISH y un tercio de los pacientes tenían dos o más alteraciones citogenéticas. Estas diferencias genéticas son las responsables de la gran heterogeneidad clínica observada en los pacientes con LLC y permiten definir grupos de pacientes con pronóstico desfavorable (17p- y 11q-), intermedio (+12, citogenética normal) y favorable (13q-), que presentan medianas de supervivencia significativamente diferentes (32, 79, 114, 111 y 133 meses, respectivamente). Además, los enfermos con 17p- y 11q-, que suelen asociarse a casos con patrón no mutado, presentaban el menor intervalo de tiempo hasta el primer tratamiento. Algunas de las alteraciones citogenéticas se relacionan con características de la enfermedad: los pacientes con 11q- o con trisomía del cromosoma 12 suelen tener grandes masas adenopáticas,[22,23] mientras que los que presentan 17p- no suelen responder a los tratamientos habituales, como la fludarabina.[24]

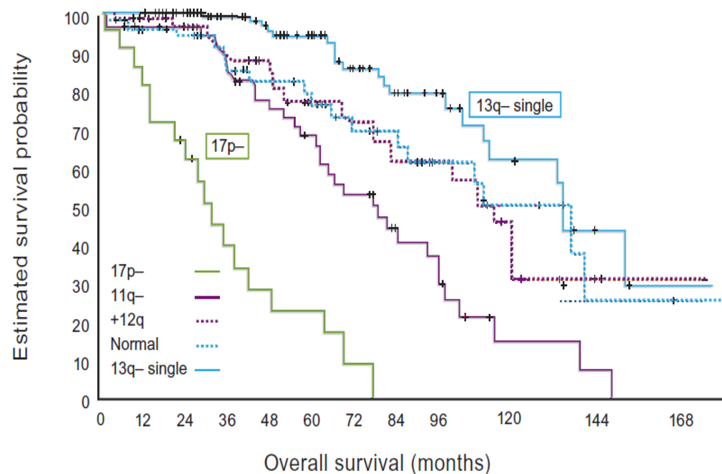


Figura 1. Probabilidad de supervivencia desde la fecha del diagnóstico de 325 pacientes con LLC según la categoría citogenética (Dohner et al[19])

A pesar de que varios grupos han analizado el valor pronóstico de algunas de las alteraciones citogenéticas que se observan en la LLC[25-27] y de que otros han reproducido los resultados del grupo de Döhner *et al,*[21,28-34] existen pocas series que permitan conocer qué ocurre con algunas alteraciones menos frecuentes en la LLC. Además en algunas series con un largo

seguimiento se observa, más allá de los 10 años de seguimiento, un entrecruzamiento de las curvas de supervivencia y del tiempo hasta la administración del primer tratamiento de los pacientes con LLC y pérdidas en 13q y de los que no tienen alteraciones citogenéticas. Esto podría indicar que, a largo plazo, los enfermos con cariotipo normal tendrían un mejor curso clínico y que el grupo de enfermos con 13q- podría ser heterogéneo a nivel clínico y pronóstico.[4,19,35]

Los estudios de FISH y las técnicas moleculares con sondas específicas de locus, y más recientemente los arrays de SNPs, han servido para delimitar las regiones genómicas que se encuentran perdidas.

La deleción en 13q14, la más frecuente en LLC, ha sido extensamente estudiada. El tamaño de la deleción es variable, con puntos de ruptura heterogéneos. Algunos autores defienden la existencia de una mínima región delecionada (MDR) que comprendería el gen DLEU2, el cluster miR-15/16 y el primer exón del gen DLEU1,[36-39] mientras que el marcador de las pérdidas de mayor tamaño sería el gen RB1, localizado en 13q14.1-q14.2[40-42] (Figura 2). Se han identificado varios genes comprendidos tanto en la MDR como en deleciones de mayor tamaño, candidatos a ser supresores tumorales, pero hasta ahora no se ha podido demostrar su inactivación.[36,43-45] Sí hay evidencias de la implicación de los dos microRNAs (miR-15a y miR-16-1) localizados en este locus, que afectan a más del 65% de los casos de LLC y que está asociada a una sobreexpresión concomitante de Bcl-2.[39]

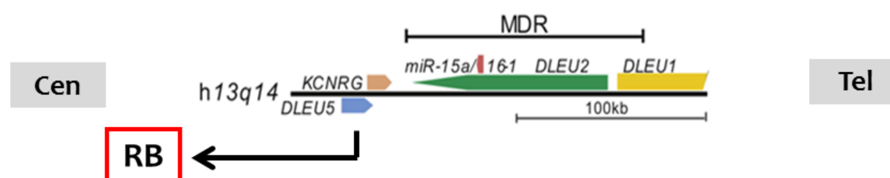


Figura 2. Tamaño de la deleción de 13q.

La importancia clínica del tamaño de la deleción en 13q no está bien establecida aunque varios estudios recientes defienden la asociación entre una pérdida de mayor tamaño y características clínicas adversas[42] o la presencia de una pérdida de mayor tamaño y la presencia de características adversas y de menor tiempo hasta el tratamiento.[40] En itros estudios, sin embargo, no se ha podido demostrar que existan diferencias ni en la clínica ni en la evolución de los pacientes de LLC según el tamaño de la deleción.[41] En esta tesis doctoral se

defiende el valor del tamaño del clon tumoral 13q- en el valor pronóstico.[35] Dos grupos han confirmado estos resultados de manera independiente.[46,47] También se han intentado integrar ambas teorías, de tal manera que los pacientes con un mayor porcentaje de células 13q- tendrían pérdidas de mayor tamaño que implicarían la MDR y RB1.[47]

La región comúnmente delecionada en la del(11q) se sitúa entre 11q22.3 y 11q23.1, donde se localiza el gen *ATM* y también presenta variabilidad en su tamaño. El gen *ATM* codifica una proteína quinasa que coordina la respuesta celular ante rupturas de la doble cadena de ADN. La inactivación de *ATM* en LLC puede ocurrir por pérdida o por mutación, que es generalmente somática, aunque también puede estar presente en la línea germinal, sugiriendo que existe una predisposición a padecer LLC en los mutantes heterocigotos. El papel que pueden desempeñar otros genes en la del(11q) todavía se desconoce, aunque se ha descrito que las alteraciones de *ATM* (deleción, mutación o ambas) son fenómenos independientes y que ambos afectan al pronóstico[48] y están asociados con un perfil de expresión génico característico[49], con solo cuatro genes comunes: *PCDH9, RXRA* (sobreexpresados), *BACE2 y TMPRSS6* (infraexpresados)[50].

La anomalía en la trisomía 12 parece que está vinculada a la banda 12q13-q22, un segmento que está duplicado en la LLC. Los genes *MDM2* y Ciclina D2 se encuentran localizados en esta región genómica. Sin embargo, el significado de la sobreexpresión de éstos en LLC no ha sido totalmente aclarado.

La deleción de 17p13 siempre afecta a *P53*, pero generalmente comprende casi todo el brazo corto del cromosoma 17, por lo que se postula la posible implicación de otros genes regulatorios en este subgrupo de pacientes. En el 80-90% de los casos está asociada a mutaciones inactivantes en el alelo homólogo.[51] La pérdida de la función de esta proteína se relaciona con estadio avanzado, resistencia al tratamiento con análogos de las purinas y alquilantes y peor pronóstico.[52,53] Además, las alteraciones de *P53* se correlacionan con una mayor complejidad genética.

Además de estas alteraciones, en algunas LLC se observan reordenamientos de *IGH*, de los que el más frecuente es la t(14;18).[19] Otras alteraciones citogenéticas son la deleción de 6q,[20,21,54] la trisomía parcial o total del cromosoma 3[25,55] y diversas traslocaciones.[52,54,56-58] Estos defectos son menos frecuentes (<10%), los genes implicados aun no se han identificado y su valor pronóstico es limitado.

Genómica

Los estudios de la expresión génica mediante microarrays de ARN han permitido mejorar el proceso diagnóstico y la clasificación de los tumores hematológicos. En el caso de la LLC, estos análisis revelan la existencia de una firma molecular única, a pesar de la heterogeneidad clínica de la enfermedad. Este perfil característico diferencia la LLC del resto de enfermedades linfoproliferativas tipo B y parece estar más próximo al de linfocitos B de memoria que al de células B "naive", CD5+, y las células del centro germinal.[59] Esta firma genética incluye las alteraciones de genes característicos de la LLC como *CD5, IL2RA* (*CDC25*) y *BCL-2*, pero también otros cuya implicación se desconocía en esta enfermedad, como *WNT3* y *ROR1*. Muchos de los genes sobreexpresados están implicados en rutas de transducción de señales. Así, *CDC25* es un intercambiador de nucleótidos de guanina en la ruta de *RAS*, mientras que *EPAC* desempeña un importante papel en la señalización de la vía de *RAP* y se relaciona con oncogénesis. Los receptores de superficie que se sobreexpresan en la LLC (por ejemplo, el receptor para *ROR1* o para tromboxano A2) son dianas terapéuticas potenciales, mientras que los genes infraexpresados están involucrados en la progresión del ciclo celular y en la replicación del ADN y metabolismo (*CDC2*, ciclina B, timidina quinasa, topoisomerasa IIa, dihidrofolato reductasa), lo que sugiere un fenotipo quiescente.

El perfil de expresión de los linfocitos B clonales en la LLC también la diferencia de otras enfermedades hematológicas como el mieloma múltiple o la macroglobulinemia de Waldeström,[60] así como de la linfocitosis B monoclonal.[61]

Aunque se ha demostrado que, independientemente de la presencia o ausencia de mutación en IgVH, todos los pacientes con LLC comparten un perfil génico común, las diferencias clínicas entre estos dos subgrupos sugieren que deben existir diferencias a nivel biológico. Por eso se han realizado estudios comparativos atendiendo al estado mutacional. En 2001, Rosenwald *et al* demostraron la asociación entre la expresión genética y el patrón mutacional.[62] En éste y en otros estudios se ha demostrado que los dos subtipos de LLC se segregan claramente en la expresión de un grupo reducido de genes como ciclina D2, *HPRT*, *GFI-1*, adenilato kinasa 2, *ILK*, CTP-sintetasa o *FGFR1*.[49,62,63] Destacaba la sobreexpresión de ZAP70 en los pacientes no mutados, lo que sugería que podía ser un marcador sustituto y más asequible que la estimación del estado mutacional de *IGVH*, hecho que se corroboró mediante determinación de este marcador por citometría de flujo.[64,65] También se han descrito diferencias significativas en la expresión génica según el sexo, lo que podría sugerir que las diferencias en las firmas moleculares referentes al estado mutacional de *IGVH* podrían estar relacionadas con el sexo

del paciente.[49,66] Existe además una correlación en la expresión de genes implicados en la señalización por *BCR* en los casos no mutados. Este hallazgo sugiere que la estimulación de los linfocitos a través de *BCR* puede ser importante en la patogénesis de la LLC, por lo que esta ruta y la regulación de *NFkB* pueden ser determinantes de la supervivencia celular en la LLC.[67]

Al estudiar el perfil de expresión de los diferentes subgrupos citogenéticos, se observa que los genes expresados de un modo más significativo se localizan en las correspondientes regiones cromosómicas aberrantes, lo que indica que la existencia de un efecto de dosis génica puede ejercer un papel patogénico en la LLC.[49] En este trabajo de tesis doctoral hemos combinado el análisis del genoma mediante CGH arrays y del transcriptoma mediante microarrays de oligonucleótidos corroborando este efecto de dosis génica en nuestra serie de pacientes. Además hemos demostrados que el perfil de expresión puede estar condicionado no sólo por la alteración citogenética sino también por el porcentaje de células afectadas, como es el caso de los enfermos con LLC y deleción de 13q.[35]

La presencia de mutaciones en diversos genes, diferentes a las inmunoglobulinas, también puede determinar cambios en el perfil de expresión. Éste es el caso de las LLC con alteraciones en *ATM* o en *TP53*. Las alteraciones en *ATM* condicionan cambios en la expresión génica, y los genes diferencialmente expresados son diferentes dependiendo del mecanismo inactivante (deleción o mutación), aunque en ambos casos el resultado final sea el mismo. El mismo fenómeno se observa en el grupo de enfermos con alteraciones inactivantes de TP53, que se correlacionan con un pronóstico desfavorable y resistencia al tratamiento. Tanto los casos con pérdida en 17p como las mutaciones en TP53 se asocian con un perfil de expresión diferencial, aunque en el caso de las mutaciones este perfil es menos característico que en la del(17p), posiblemente debido al efecto de dosis génica observado en este segundo grupo.[68]

La aplicación de las técnicas de secuenciación masiva ha permitido identificar nuevas alteraciones recurrentes en genes como *NOTCH1, XPO1, MYD88* y *KLHL6* en el 12.2%, 2.9%, 2.4% y 1.8% de las LLCs, respectivamente. En algunos casos, estas mutaciones se han asociado a un perfil de expresión característico, como las mutaciones de *NOTCH1*. Así, en los pacientes mutados se observa una sobreexpresión de los genes implicados en la ruta de NOTCH, como *NOTCH1, NOTCH2, NOTCH4, CREBBP, ADAM17* o *NCOR2*, entre otros. Estos pacientes se asocian, además, con un pronóstico desfavorable.

Cinética de la LLC

La idea tradicional de que la LLC es una enfermedad en la que los linfocitos B se acumulan debido a un fallo en los mecanismos de apoptosis causado por un defecto genético, con una mínima proliferación del clon leucémico[69] ha ido cambiando a la luz de los nuevos conocimientos de la biología de la LLC. Desde hace tiempo se sabe que los pacientes en los queel número absoluto de linfocitos se duplica en menos de un año generalmente tienen un curso clínico peor que aquellos que tienen una acumulación más lenta, incluso en estadios tempranos de la enfermedad.[70] Apoyando esta observación, estudios recientes que utilizan técnicas más sofisticadas, sugieren que la caracterización inicial de la biología de la LLC necesita ser reevaluada.[71] Con un ensayo de marcaje in vivo ha sido posible evaluar la velocidad de proliferación de las células de LLC, y se ha documentado que una pequeña fracción del clon leucémico (aproximadamente 0.1 – 1.78%) está dividiéndose de manera activa, lo que generaría aproximadamente entre $1x10^9$ y $1x10^{12}$ nuevas células por día en los pacientes estudiados. Por lo tanto, hay una dinámica entre muerte y división celular similar a la que tiene lugar en otros tipos de leucemias y linfomas. La LLC es una enfermedad tanto de proliferación como de acumulación de células B, de tal manera que en los pacientes con un número de linfocitos estable existiría un balance homeostático entre ambos procesos, que estaría desequilibrado en los pacientes cuyo número de linfocitos aumenta de manera sostenida, duplicándose en menos de un año, y generando así una enfermedad más agresiva.[72]

Las células de LLC se dividen en los centros de proliferación (PC) en los nódulos linfáticos y en el bazo. En estas estructuras, los linfocitos de LLC se comunican con las células del micromedioambiente, como las células T y las células del estroma. La interacción con estas células proporciona las señales necesarias para la supervivencia y expansión de las células de LLC, promovidas por moléculas como IL-4 y CD40L, cuyo efecto ha sido documentado in vitro.[73,74] Por lo tanto la acumulación de las células leucémicas podría estar relacionada con la generación de señales de supervivencia producidas por el medio ambiente externo a través de diversos receptores (receptor de la célula B, receptores de citocinas…).[59,75,76] Cuando se cultivan in vitro, los linfocitos de LLC sufren apoptosis, lo que respaldaría el papel esencial del microambiente celular en la supervivencia de estas células.[77]

La señalización por BCR es clave en el comportamiento del linfocito B. El estudio de la estructura y función del BCR utilizado por las células leucémicas en su relación con el entorno celular ha ayudado a redefinir algunos aspectos de la LLC, y cada vez son mayores las evidencias de que la señalización via BCR desempeña un papel fundamental en el desarrollo de

la LLC e influye en su heterogéneo comportamiento clínico. Se ha demostrado que las LLCs de peor pronóstico presentan una señalización vía BCR por IgM aumentada, produciéndose una mayor fosforilación de proteínas intracelulares y de los fenómenos celulares relacionados con la activación celular.[78] Por lo tanto, la señalización por BCR influye en la supervivencia clonal y promueve la expansión,[72,79] impactando negativamente en el pronóstico.

La regulación de la apoptosis es un mecanismo fisiológico complejo, y en las células de la LLC puede estar alterada por diversas causas, como las alteraciones genéticas (defectos en p53 y ATM) o los cambios en la expresión de moléculas reguladoras, como algunos miembros de la familia de Bcl-2. Las células de la LLC expresan una gran variedad de proteínas de la familia bcl-2 con un perfil que favorece la inhibición de la apoptosis, observándose niveles altos de esta proteina en más del 85 % de los casos estudiados. Este hecho, junto con la interacción con las células del microambiente, explicaría la larga supervivencia y la subsecuente acumulación de las células de la LLC.[20] Las rutas de PI3K/Akt y NFkB desempeñan también un papel importante en la supervivencia de las células B, induciendo la expresión de proteínas antiapoptóticas como Mcl-1 y XIAP, e inactivando blancos celulares involucrados en la inducción de apoptosis, tales como BAD y Caspasa 9. El control de la apoptosis en LLC es un mecanismo complejo y, debido a la influencia que ejerce el micromedioambiente en la supervivencia de los linfocitos de LLC, difícil de evaluar *in vitro*.

microRNAs

Los microRNAs (miRNAs) son una familia de RNAs pequeños, de unos 21-25 nucleótidos, que inhiben la expresión génica por la unión a la región 3´UTR de su gen diana, donde se encuentra la secuencia complementaria, modulando la expresión del gen a nivel post-transcripcional. Estas moléculas de RNA no codificante están implicadas en la regulación de la expresión de genes relacionados con la proliferación celular, la diferenciación, el ciclo celular, la apoptosis y el metabolismo y con patologías como la diabetes y el cáncer**.**

MiRNAs y cáncer

La primera relación que se estableció entre miRNAs y cáncer fue en enfermos con LLC, en los que se observó una pérdida del cluster miR-15/16-1, localizado en 13q14.3, concretamente entre los exones 4 y 5 del gen DLEU2,[80] así como una sobreexpresión de los miRNA-21 y miRNA-155. Esta observación condujo a indagar en la asociación entre las localizaciones de los miRNAs y las regiones genómicas involucradas en cáncer. Así, la mitad de los miRNAs

conocidos están localizados dentro o cerca de sitios frágiles, en regiones de pérdida de heterocigosidad, en regiones de amplificación génica o en puntos de ruptura habitualmente asociados a tumores.[81]

El análisis del perfil de expresión de miRNAs en pacientes con cáncer mediante microarrays ha demostrado que los miRNAs se expresan de manera diferente en muestras normales frente a tumorales.[82] Cada vez son mayores las evidencias que sugieren que la expresión anormal de los miRNAs es una característica común de los procesos neoplásicos.[83] Los miRNAs desempeñan un papel directo en la oncogénesis, pudiendo actuar como oncogenes o como genes supresores de tumores, y su expresión aberrante se ha asociado con muchos tipos de cánceres incluyendo tumores sólidos y hematológicos. Los patrones de expresión de los miRNAs pueden distinguir cánceres según su diagnóstico y estadio de desarrollo, de forma mucho más precisa que los estudios tradicionales de expresión génica. Por lo tanto una mejor comprensión de los mecanismos que controlan la expresión de los miRNAs en el cáncer y sus consecuencias funcionales puede mejorar la clasificación, el pronóstico y el tratamiento del cáncer.

Los miRNAs en la patogénesis de la LLC

Uno de los primeros indicios de que los miRNAs estaban asociados al cáncer fue la observación de que los miR-15a y miR-16-1 se encontraban en la región perdida en 13q14 en los enfermos con LLC. Debido a la pérdida de esta región, ambos miRNAs se infraexpresan en LLC. Su papel como supresores tumorales fue evidente al demostrarse que *BCL-2*, gen antiapoptótico sobreexpresado en LLC, es una de sus dianas, por lo que niveles bajos o inexistentes de miR-15a y miR-16-1 permiten la expresión de *BCL-2*.[84] Desde entonces, se han realizadomuchos estudios que confirman la relación entre los miRNAs y otros parámetros biológicos de relevancia en LLC.

La expresión de miRNAs en los linfocitos B de LLC se ha estudiado principalmente mediante microarrays de miRNAs. De esta manera se han identificado miRNAs cuya expresión diferencia a los linfocitos B de LLC de las células normales CD5+ y las asemeja a células B activadas.[85-88] Además se ha estudiado la expresión de miRNAs en relación a parámetros clínicos establecidos, como el estado mutacional, la expresión de ZAP-70 y LPL o las alteraciones citogenéticas (Tabla 3). De esta manera se han descrito **firmas de microRNAs** con valor pronóstico. Aunque ninguna de estas firmas es exactamente igual, sí presentan algunos miRNAs comúnmente desregulados.

Tabla 3. MiRNA profiling studies of CLL

Table 2. MiRNA profiling studies of CLL

MiRNA signatures	Prognostic marker	Study	Year
miR-15a, miR-195, miR-221, miR-23b, miR-155, miR-223, miR-29a-2, miR-24-1, miR-29b-2, miR-146, miR-16-1, miR-16-2, and *miR-29c*	Unmutated IgV_H, high ZAP-70	Calin et al[21]	2005
miR-29c and *miR-223*	Unmutated IgV_H	Fulci et al[80]	2007
miR-181a, let-7a, miR-30d, miR-155, and *miR-29*	Unmutated IgV_H	Marton et al[60]	2008
miR-223 and *miR-29c*	ZAP-70 and LPL	Stamatopoulos et al[81]	2009

LPL indicates lipoprotein lipase.

Los miRNAs también son importantes en LLC porque regulan la expresión de muchos genes claves en la patogénesis de la enfermedad (Figura 3). Así, miR-29 y miR-181 regulan el oncogen *TCL1*, que está sobreexpresado en aproximadamente el 25-35% de los pacientes con LLC y se ha asociado con un fenotipo agresivo en LLC.[89,90] La sobreexpresión de miR-29 reduce los niveles de Mcl-1, con función antiapoptótica, provocando la apoptosis en las células de LLC.[91-93] miR15a y miR-16-1 regulan la expresión de *BCL-2* y de varios factores con actividad pro-proliferativa, como *CCND2, CCND3, CDK4* y *CDK6*.[39,94]. MiR-34a está directamente regulado por p53, y a su vez regula a Bcl-2, y recientemente se ha descrito la relación que existe entre la señalización por BCR en los linfocitos B de la LLC y miR-155.[95]

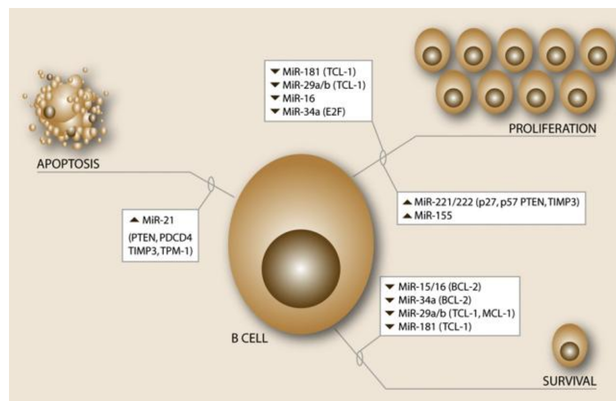


Figura 3. miRNAs implicados en apoptosis, supervivencia y proliferación en la LLC. Las flechas indican el patrón de expresión en la LLC (sobre o infraexpresado) y los genes entre paréntesis, las dianas validadas.[96]

Técnicas de análisis genómico global

A mediados de los años 50 Joe Hin Tjio y Albert Levan establecieron el número y la estructura los cromosomas humanos.[97] En 1953 Watson y Crick definieron la estructura de la molécula de ADN,[98] dando el primer paso en el estudio del genoma humano. Desde entonces se ha producido un progreso espectacular en el conocimiento sobre el material genético humano y los grandes adelantos tecnológicos surgidos en las dos últimas décadas han dado lugar a un espectacular avance en el conocimiento biológico y médico.

Las nuevas tecnologías aplicadas al análisis genético ofrecen grandes ventajas como la rapidez, especificidad, sensibilidad y posibilidad de automatización. Sin duda la mayor ventaja de las técnicas de análisis genómico global frente a los métodos tradicionales es la posibilidad de analizar simultáneamente miles de genes en un único experimento. Se están aplicando en la actualidad, entre otros, a la identificación de perfiles genéticos y dianas terapéuticas, detección de mutaciones y polimorfismos, secuenciación, seguimiento de terapias, farmacogenómica y diagnóstico molecular.

Microarrays de expresión

El conjunto de genes que se expresan o se transcriben a partir del ADN genómico y que constituyen el transcriptoma o perfil de expresión es un importante determinante del fenotipo y la función celular. Los microarrays de expresión permiten el análisis simultáneo de miles de genes en un único experimento.

El estudio del perfil genético propio de cada tipo de neoplasia ha adquirido cada vez mayor relevancia, tanto desde el punto de vista de la biología molecular para el entendimiento de la etiopatogenia de los procesos neoplásicos, como desde el punto de vista clínico. De esta manera, se han podido correlacionar los perfiles de expresión de algunos tumores y el pronóstico de la enfermedad, clasificar mejor neoplasias ya conocidas, definir nuevas entidades indistinguibles con las tecnologías empleadas hasta el momento e incluso identificar dianas diagnósticas y terapéuticas que ya tienen una gran utilidad práctica en el manejo de pacientes con neoplasias hematológicas.[60,64,99-110] Sin duda esta tecnología ha permitido un salto cuantitativo (se analizan miles de genes) y cualitativo en la investigación, al aportar una visión molecular global de los procesos patológicos.

Microarrays genómicos

Los microarrays genómicos son herramientas de análisis genético que permiten un análisis global de las ganancias y pérdidas de ADN. Esta técnica se basa en los mismos principios que la CGH convencional, pero en vez de cromosomas se utilizan sondas de ADN, sobre las que se hibrida el ADN del tumor.[111-113] Pueden incluir sondas de varios tamaños, que van desde 100-200 kb (BACs y PACs) a sondas de menor tamaño, pero mayor especificidad como los oligonucleótidos (25-80 bases), que pueden cubrir regiones cromosómicas concretas o todo el genoma completo.[114] Las ventajas y las limitaciones de los arrays genómicos dependen de la plataforma elegida. Por ello, tanto la cantidad de clones depositados en el array como la manera de amplificarlos es crucial para la sensibilidad final y la calidad del análisis. El uso de clones de grandes insertos como son los BAC y PAC proporciona una intensidad de señal suficiente para detectar cambios en una sola copia y las alteraciones pueden ser inmediatamente relacionadas con los marcadores genéticos.[115]

Esta tecnología ha resultado de gran utilidad para delimitar áreas de ganancia y pérdida genómica en enfermedades de origen genético, hereditario y tumoral, así como para caracterizar los perfiles genómicos de varios procesos tumorales con una mayor resolución que la que se conseguía con las técnicas disponibles hasta el momento, como la CGH convencional.[105,116-123]

En este trabajo de tesis doctoral hemos aplicado para el análisis genético de la LLC un BAC-array de resolución 1Mb diseñado y fabricado en nuestro laboratorio, que contiene 3.528 clones, 176 de los cuales contenían genes relacionados con cáncer (oncogenes o genes supresores de tumores). El resto de los clones están distribuidos a lo largo del genoma humano, espaciados a intervalos de aproximadamente 1 Mb. Además hemos usado microarrays de oligonucleótidos para confirmar los resultados observados por el BAC array

Técnicas de secuenciación masiva

Desde que se describió la secuenciación del ADN en 1977, las mejoras en las técnicas y equipos, así como en la bioinformática necesaria para el análisis, han permitido la automatización y han mejorado el coste de este tipo de análisis genéticos y su utilidad en la práctica médica. En los últimos años han surgido en el panorama de la genómica médica diversos métodos de secuenciación masiva en paralelo, denominados genéricamente **“next-generation sequencing”** o secuenciación de segunda generación, que permiten la

secuenciación de grandes fragmentos de ADN de manera rápida y asequible. Este nuevo tipo de secuenciación generación permite buscar simultáneamente mutaciones en cientos de loci para enfermedades genéticamente heterogéneas como el cáncer. Además, la secuenciación masiva permitirá abordar de manera más comprensiva disciplinas tales como la farmacogenética y la epigenética, integrando datos globales que permitan interpretar interacciones génicas y mecanismos epigenéticos de regulación en la expresión.

Las principales plataformas de secuenciación masiva disponibles en la actualidad son tres: Illumina Genome Analyzer, Applied biosystems SOLID Sequencer y Roche GS-FLX 454 Genome Sequencer. La tecnología de Illumina utiliza la “secuenciación por síntesis”, en la que la incorporación de un nucleótido marcado con fluorescencia y protegido en la cadena naciente impide que ésta siga creciendo. Tras detectar la señal fluorescente, se elimina el grupo protector y se puede incorporar otro nucleótido marcado, con lo que se empieza de nuevo el ciclo. Así se generan lecturas sencillas de 75 bp con un rendimiento de 17 GB de secuencia en 7 días. El secuenciador SOLID de Applied se fundamenta en la “secuenciación por unión”. Esta tecnología secuencia por ligación de octámeros marcados de secuencia conocida a la cadena de ADN, con la posterior detección de la señal fluorescente emitida tras cada unión. De ésta manera se secuencian fragmentos de 50 bp con un rendimiento de 10-15 GB en 3-7 días. La tecnología utilizada por Roche y su GS-FLX Genome Sequencer se basa por un lado en la PCR en emulsión, y por otro, en la “pirosecuenciación”. Las etapas/pasos seguidos en esta tecnología están resumidos en la Figura 4.

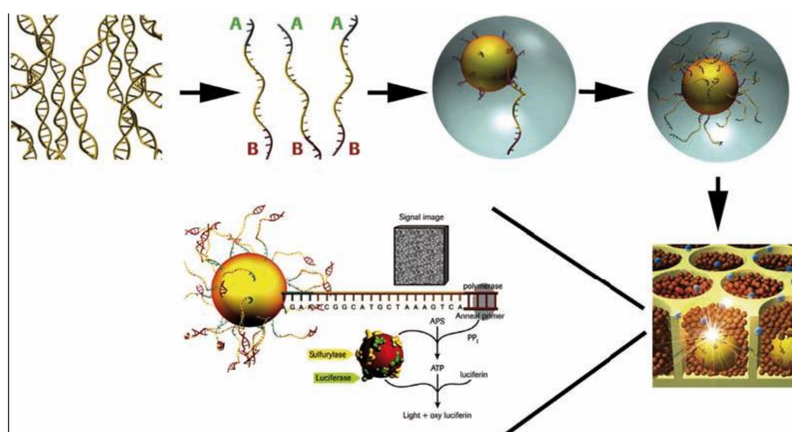


Figura 4. Método de “pirosecuenciación” del sistema GS-FLX Genome Sequencer de Roche

1- Preparación de la librería de DNA

La preparación de la librería de DNA consiste en el fraccionamiento del DNA genómico (gDNA) en pequeños fragmentos (de 300 a 500 pb) que son posteriormente pulidos (extremos romos) y ligados a adaptadores (A y B). Estos adaptadores proporcionan las secuencias de hibridación para la posterior amplificación y secuenciación de los fragmentos de la librería. El adaptador B está biotinilado en su extremo 5' y esto permite la inmovilización de la librería mediante las perlas recubiertas de estreptavidina.

2- PCR en emulsión

Cada esfera contiene una única molécula de DNA de cadena sencilla de la librería. La esfera unida a la librería se emulsiona con los reactivos de amplificación en una micela de agua y aceite, de tal manera que cada esfera queda englobada en su propio microreactor dentro del cual ocurre la amplificación mediante PCR. El resultado será una esfera inmovilizada que contiene fragmentos de DNA amplificados clonalmente.

3- Secuenciación

Las esferas unidas a los fragmentos de ADN se distribuyen en capas dentro de la placa de secuenciación junto con unas esferas enzimáticas que contienen luciferasa y sulfurilasa. Cada pocillo de la placa contiene una única esfera con miles de copias del mismo fragmento de ADN. Durante el flujo de nucleótidos, cada una de los cientos de miles de perlas con millones de copias de DNA se secuencian en paralelo. Si un nucleótido es complementario a la cadena molde en algún pocillo, la polimerasa extiende la hebra existente de DNA mediante la adición de nucleótido(s). La adición de uno (o más) nucleótido(s) resulta en una reacción que genera una señal de luz que es recogida por la cámara CCD del equipo. La intensidad de la señal es proporcional al número de nucleótidos incorporados en un solo flujo de nucleótidos. Esto genera una imagen que se analizan e interpreta en forma de secuencias de nucleótidos.

La ventaja de esta metodología es que genera secuencias de 400- 500 bp, generando 400-600 Mb de datos en 10 horas, aunque el coste en reactivos es superior al de otras plataformas y puede confundirse en las secuencias homopoliméricas.

Paralelamente, el **desarrollo de protocolos de captura selectiva del ADN** permite realizar la resecuenciación parcial de regiones del genoma (exones, genes o regiones de interés) a unos costes que hacen actualmente posible el estudio de un mayor número de muestras. Los protocolos existentes para la captura selectiva de regiones de interés se basan mayoritariamente en sistemas de hibridación (Albert et al., 2007; Hodges et al., 2007), en arrays (NimbleGen Array Capture, Febit HybSelect) o en solución (NimbleGen, Agilent

SureSelect), mediante miles de oligonucleótidos que cubren adecuadamente las regiones seleccionadas y capturan regiones de hasta 5-10 Mb. Aunque también se han descrito sistemas basados en PCR múltiple (Porreca et al., 2007; Tewhey et al., 2009) su aplicación ha sido menor debido al elevado número de PCRs necesarias para capturar regiones de un cierto tamaño. En este trabajo de tesis doctoral hemos utilizado los arrays de captura de secuencia de Nimblegen, que permiten capturar y enriquecer regiones génicas de interés, contiguas o no en el genoma, con una elevada sensibilidad y especificidad. Estas regiones son posteriormente amplificadas y secuenciadas mediante tecnologías de alto rendimiento (FLX 454 en nuestro caso). De esta manera, el coste de la secuenciación se reduce considerablemente y técnicamente el proceso también es menos costoso.

La combinación de técnicas de análisis global del genoma y del transcriptoma aplicadas en este trabajo de tesis doctoral como son los microarrays y la secuenciación masiva, constituye una estrategia de gran utilidad para el estudio y la caracterización de una enfermedad tan heterogénea como es la LLC y puede permitir la identificación de nuevas alteraciones genéticas, incluso presentes en un pequeño porcentaje de enfermos.

Hipótesis

La leucemia linfática crónica (LLC) es una hemopatía con un comportamiento clínico variable. El desarrollo de la citogenética ha demostrado que la LLC es una enfermedad muy heterogénea también desde el punto de vista genético. Por lo tanto, las nuevas técnicas de análisis, que facilitan el estudio del genoma de manera global, pueden constituir una herramienta de gran utilidad para el estudio de las alteraciones genéticas en LLC.

Los microarrays genómicos permiten analizar, en un solo experimento, la presencia de ganancias y pérdidas de cualquier segmento del genoma y, gracias a su gran resolución, delimitar regiones génicas implicadas en los procesos neoplásicos que podrían pasar inadvertidas mediante otras técnicas de citogenética molecular. Los microarrays se han aplicado en el estudio global de la LLC para profundizar en el estudio de las alteraciones previamente definidas mediante FISH e identificar nuevas alteraciones recurrentes. Sin embargo, pocos estudios han investigado la presencia de nuevas alteraciones en la expresión génica en relación con los grupos citogenéticos de la LLC.

Hasta el momento se estimaba que los pacientes con pérdida de 13q14, la alteración citogenética más frecuente en LLC, son los que presentan un mejor pronóstico, incluso superior a aquellos en los que no se evidencian aberraciones citogenéticas determinadas mediante FISH. Sin embargo, nuestro grupo ha demostrado en un trabajo reciente que los pacientes con pérdidas de 13q constituyen un grupo heterogéneo, en el que el porcentaje de células con esta alteración condiciona un comportamiento clínico diferente. Aunque estos resultados han sido corroborados por otros grupos, todavía no se han definido los mecanismos biológicos responsables de esta diversidad clínica.

En la LLC se produce una inestabilidad genética que facilita la presencia de alteraciones cromosómicas, aunque los mecanismos por los que se producen se desconocen[13]. Estos datos han sido confirmados recientemente, mediante estudios de secuenciación completa del genoma humano. En los trabajos publicados por el Consorcio Español de LLC se han identificado mutaciones recurrentes en los genes *NOTCH1*, *XPO1*, *MYD88* y *SF3B1*, con evidencias que relacionan estas alteraciones con la patogenia de la LLC y, en algunos casos,

pueden ayudar a definir el pronóstico. Las técnicas de secuenciación masiva constituyen, por tanto, una herramienta muy útil para descifrar el genoma de pacientes con LLC en busca de nuevas claves sobre la enfermedad.

Conocemos la secuencia de los genes pero, en la mayoría de los casos, su función está poco estudiada, por lo que son necesarios estudios funcionales complementarios que nos ayuden a comprender cómo las alteraciones a nivel genómico (ganancias, pérdidas, mutaciones o polimorfismos) están afectando a los procesos celulares. Además, se desconocen muchos de los genes afectados en la LLC, por lo que la aplicación de técnicas de análisis masivo como los microarrays o la ultrasecuenciación puede resultar de gran utilidad para definir la presencia de nuevos genes y rutas alteradas en LLC.

Por todo ello, en el presente trabajo de tesis doctoral hemos analizado una serie amplia de pacientes combinando las técnicas más recientes de análisis masivo (microarrays y secuenciación) combinadas con análisis funcionales, para profundizar en el conocimiento de la LLC, con especial atención al grupo de enfermos con pérdida de 13q. Nuestra hipótesis se basa en que la LLC es una enfermedad heterogénea, con varios subtipos genéticos y que, dentro de cada uno de ellos, es posible que existan diferencias en cuanto a la biología, que puedan condicionar las variaciones observadas en la práctica clínica. El análisis combinado y exhaustivo del genoma y del trascriptoma de la célula tumoral, debería confirmar la presencia de diferencias en estos tipos genéticos.

Objetivos

Objetivo general

Profundizar en el conocimiento de los mecanismos fisiopatológicos de la LLC en los distintos subgrupos citogenéticos mediante un análisis combinado del genoma y transcriptoma de los linfocitos B clonales.

Objetivos particulares

1. Determinar la existencia de nuevas regiones genéticas alteradas en los enfermos con LLC mediante la aplicación de microarrays genómicos.

2. Analizar, mediante microarrays de exones, los perfiles de expresión de los diferentes subgrupos citogenéticos de LLC: FISH normal, 13q-, +12, 11q- y 17p- así como de las nuevas alteraciones genéticas identificadas.

3. Caracterizar el subgrupo de enfermos con pérdida de 13q mediante estudios de expresión génica y de microRNAs.

4. Determinar la presencia de mutaciones en genes que intervienen en la patogénesis de la LLC mediante técnicas de secuenciación masiva de ADN.

5. Realizar estudios funcionales en los genes que presenten una expresión génica alterada o mutaciones, para definir su implicación en la patogénesis de la LLC.

Paper I

Rodríguez AE, Robledo C, García JL, González M, Gutiérrez NC, Hernández JA, Sandoval V, García de Coca A, Recio I, Risueño A, Martín-Núñez G, García E, Fisac R, Conde J, de Las Rivas J, Hernández JM. **Identification of a novel recurrent gain on 20q13 in chronic lymphocytic leukemia by array CGH and gene expression profiling.** Ann Oncol. 2012 Jan 6

Annals of Oncology Advance Access published January 6, 2012

original article

Annals of Oncology
doi:10.1093/annonc/mdr579

Identification of a novel recurrent gain on 20q13 in chronic lymphocytic leukemia by array CGH and gene expression profiling

A. E. Rodríguez[1], C. Robledo[1], J. L. García[2], M. González[3], N. C. Gutiérrez[3], J. A. Hernández[4], V. Sandoval[5], A. García de Coca[6], I. Recio[7], A. Risueño[8], G. Martín-Núñez[9], E. García[10], R. Fisac[11], J. Conde[12], J. de las Rivas[8] & J. M. Hernández[1,3]*

[1]IBMCC, Centro de Investigación del Cáncer, Universidad de Salamanca-CSIC, Salamanca; [2]Instituto de Estudios de Ciencias de la Salud de Castilla y León (IECSCYL)–HUSAL, Castill y León; [3]Department of Hematology, Hospital Clínico Universitario de Salamanca, Salamanca; [4]Department of Hematology, Hospital Infanta Leonor, Madrid; [5]Department of Hematology, Hospital Virgen Blanca, León; [6]Department of Hematology, Hospital Clínico Universitario, Valladolid; [7]Department of Hematology, Hospital Nuestra Señora de Sonsoles, Ávila; [8]Bioinformatics and Functional Genomics, Centro de Investigación del Cáncer, Universidad de Salamanca-CSIC, Salamanca; [9]Department of Hematology, Hospital Virgen del Puerto, Plasencia; [10]Genomics and Proteomics Unit, Centro de Investigación del Cáncer, Universidad de Salamanca-CSIC, Salamanca; [11]Department of Hematology, Hospital General de Segovia, Segovia; [12]Department of Hematology, Hospital del Río Hortega, Valladolid, Spain

Received 3 August 2011; revised 11 November 2011; accepted 16 November 2011

Downloaded from http://annonc.oxfordjournals.org/ by guest on January 8, 2012

Background: The presence of genetic changes is a hallmark of chronic lymphocytic leukemia (CLL). The most common cytogenetic abnormalities with independent prognostic significance in CLL are 13q14, *ATM* and *TP53* deletions and trisomy 12. However, CLL displays a great genetic and biological heterogeneity. The aim of this study was to analyze the genomic imbalances in CLL cytogenetic subsets from both genomic and gene expression perspectives to identify new recurrent alterations.

Patients and methods: The genomic imbalances and expression levels of 67 patients were analyzed. The novel recurrent abnormalities detected with bacterial artificial chromosome array were confirmed by FISH and oligonucleotide microarrays. In all cases, gene expression profiling was assessed.

Results: Copy number alterations were identified in 75% of cases. Overall, the results confirmed FISH studies for the regions frequently involved in CLL and also defined a new recurrent gain on chromosome 20q13.12, in 19% (13/67) of the CLL patients. Oligonucleotide expression correlated with the regions of loss or gain of genomic material, suggesting that the changes in gene expression are related to alterations in copy number.

Conclusion: Our study demonstrates the presence of a recurrent gain in 20q13.12 associated with overexpression of the genes located in this region, in CLL cytogenetic subgroups.

Key words: CLL, cytogenetic aberrations, gene expression profile, genomic arrays

introduction

Chronic lymphocytic leukemia (CLL) is the most common leukemia in the western world and is characterized by a highly variable clinical course with survival times ranging from months to decades despite a remarkable phenotypic homogeneity [1, 2]. This clinical heterogeneity reflects its biological diversity [3]. Our understanding of the biology of CLL has helped to identify several markers of prognostic significance, delineating CLL into several distinct diseases. These markers include cytogenetic abnormalities, the mutational status of the immunoglobulin heavy chain variable (*IGHV*) and ZAP-70, CD38 and CD49d expression [2, 4–6]. Conventional cytogenetic analyses have revealed chromosomal aberrations in 40%–50% of patients, but detection of abnormalities is limited by the low mitotic activity of CLL cells. By contrast, interphase FISH (iFISH) has identified chromosomal changes in ∼80% of patients with CLL, the presence of specific chromosomal abnormalities being a prognostic indicator of disease progression and survival [2, 7]. Thus, half of the CLL patients carry deletions of 13q, which is correlated with an indolent disease course in patients with this abnormality as their sole aberration. In contrast, deletions of 11q and 17p (which cover the *ATM* and *TP53* genes, respectively) have a poorer outcome. Furthermore, trisomy 12 is related to an intermediate prognosis, whereas deletion of 6q has been identified as a recurrent CLL progression marker [8]. In addition, great genomic complexity has been associated with worse survival and is also closely related to markers of poor prognosis [9–11].

*Correspondence to: Prof. J. M. Hernández, Hematology Unit, Department of Medicine, Hospital Universitario de Salamanca, Paseo San Vicente 58, 37007 Salamanca, Spain. Tel: +34-923-291-100; Fax: +34-923-294-624; E-mail: jmhr@usal.es

© The Author 2012. Published by Oxford University Press on behalf of the European Society for Medical Oncology. All rights reserved. For permissions, please email: journals.permissions@oup.com

Considering the great heterogeneity of CLL from both genetic and prognostic points of view, microarray technology is a powerful tool for the analysis of genetic alterations in CLL. Thus, comparative genomic hybridization using high-density arrays, array comparative genomic hybridization (aCGH), allows high-resolution genome-wide scan for detection of copy number alterations in a single hybridization and aCGH using bacterial artificial chromosome (BAC) clones has been widely applied in the analysis of hematological malignancies [12–15]. Regarding oligonucleotide microarrays, the study of the gene expression profile (GEP) in CLL has given us insights into the molecular mechanisms involved in its pathogenesis by analyzing the impact of genomic aberrations on the expression of genes located on the corresponding loci [16–18].

Although the application of microarray technology in CLL has provided additional knowledge of the known recurrent aberrations as well as enabling novel aberrations, such as gain of 2p and deletion of 22q to be identified [19–23], to date, few studies have investigated genomic aberrations specifically in relation to CLL cytogenetic subsets. Therefore, the aim of this study was not only to screen and identify new genomic events in CLL patients but also to compare the prevalence of these genomic aberrations in cytogenetic CLL subsets. Furthermore, our data revealed an association between altered transcription levels and genomic imbalances in the genetic subsets of CLL, indicating that gene dosage might have pathogenic effects in CLL and delineate a new gained region, on 20q13, in CLL patients.

methods

patients

Peripheral blood samples from 67 patients with CLL were analyzed. The diagnoses were confirmed by standardized clinical, morphological and immunological data according to the World Health Organization classification and the criteria of the Working Group of the National Cancer Institute [24]. FISH studies and *IGH* mutational status were determined in all patients. The study protocol was approved by the local ethical committees and prior written informed consent was obtained from the patients. All patients were untreated and most of them were studied at the moment of diagnosis (Table 1). The main characteristics of the 67 CLL patients included in the study are reported in Table 2.

FISH analysis

Interphase FISH was carried on all the samples using commercially available probes for the following regions: 13q14, 12q13, 11q22/*ATM*, 17p13/*TP53* and 14q32/*IGH* (Abbott Co., Downers Grove, IL) using the previously described methods [25].

To confirm the gains and losses assessed by aCGH, FISH analysis was done using Vysis LSI ZNF217, the commercially available probe for 20q13.2 (Abbott Co.) and the BAC clones dJ1028D15–dJ781B1, mapping to 20q13.12, as previously described [25]. The clones were located in the same region of gain as detected by aCGH and were selected from the aCGH BAC clone library (Wellcome Trust Sanger Institute), whereas the commercial probe was located in 20q13.2 (breast tumor amplicon). DNA from the BAC clones was isolated, labeled and hybridized, as previously described [26]. The changes were validated in fixed cells from the same diagnostic samples as used for aCGH ($n = 20$).

FISH analysis was carried out on 400 interphase cells using standard fluorescent microscopy.

mutation status of IGVH genes

IGVH genes were amplified and sequenced according to the ERIC recommendations on *IGHV* gene mutational status analysis in CLL [27].

array comparative genomic hybridization

BAC arrays. DNA samples were analyzed using a BAC array containing 3523 sequence-validated BACs covering the genome with a mean resolution of 1 Mb, as previously described [26].

Table 1. Status of disease in the total series ($n = 67$) and in +20q CLL patients ($n = 13$)

	CLL patients ($n = 67$) n (%)	+20q CLL patients ($n = 13$) n/CLL (%)
At diagnosis	50 (75)	13/50 (26)
Progressive	17 (25)	4/17 (24)

CLL, chronic lymphocytic leukemia.

Table 2. Clinical and molecular characteristics of the CLL patients

Characteristics	(%)
Median age in years (range)	68 (35–90)
Male/Female (ratio)	73/23 (2.7)
White blood cells, range/ml	39 000 (7600–175 000)
Lymphocytes/ml (range)	32 000 (5000–160 000)
Hemoglobin, g/dl (range)	13.6 (7.1–16.3)
Platelet count/ml (range)	167 000 (59 000–306 000)
LDH	
Normal	82
High	12
β_2-microglobulin	
Normal	52
High	48
Status of the disease	
At diagnosis	71.6
Progressive	28.4
Binet stage	
A	66
B	26
C	9
ZAP-70 expression	
Positive	44
Negative	56
CD38 expression	
Positive	26
Negative	74
IgVH mutational status	
Mutated	41
Unmutated	59
Interphase FISH analysis	
Normal karyotype	22
13q deletion	37
Trisomy 12	16
11q deletion	7
17p deletion	9
IGH translocation	9
20q13.12 gain	19

CLL, chronic lymphocytic leukemia; LDH, lactate dehydrogenase.

Downloaded from http://annonc.oxfordjournals.org/ by guest on January 8, 2012

oligonucleotide microarrays. In order to confirm the results of the BAC aCGH analysis, a subset of 35 patients were analyzed using a NimbleGen Human CGH 4×72K Whole Genome v2.0 array (Roche Diagnostics, Mannheim, Germany).

The complete description of BAC and oligonucleotide microarrays experiments is available as supplementary Material (available at *Annals of Oncology* online).

GEP analysis

RNA isolation, labeling and microarray hybridization were carried out, as previously reported [28]. The GEP was analyzed in all cases with Human Genome U133A microarray (Affymetrix, Santa Clara, CA). Data analysis is available as supplementary material (available at *Annals of Oncology* online).

comparative analysis of CGH and expression arrays

In order to achieve a comparative analysis of the copy number changes, from the CGH arrays, and the gene changes, from the expression arrays, for the same patients, we select the patients who showed significant gains in the aCGH data and their corresponding expression data. We normalized the expression dataset using the R package GeneMapper [29] that allows an accurate assignment to ENSEMBL genes (instead to Affymetrix probesets) including their location in the genome. Following this, we selected the three regions in chromosome 20q where the gains detected by aCGH were significative. For such regions, we calculate, in the corresponding samples, the mean and median expression signal based on the genes included. On these expression numbers, we carried out a statistical one-tail *t*-test (using R) to check if there was a significant correlation between the aCGH gain observed in the 20q regions and the overexpression of the genes included in such regions.

statistical analysis

Two-tailed Chi-square and Fisher's exact tests were used to analyze the associations between variables. For all tests, values of $P < 0.05$ were considered to indicate statistical significance. The calculations were carried out using SPSS 17.0 for Windows (SPSS Inc., Chicago, IL).

results

FISH and mutational status

FISH analyses revealed that 25 of the 67 cases analyzed (37%) carried the 13q14 deletion and that this was the only abnormality in 20 patients (30%). Overall, the 11q22.3 and 17p13.1 deletions were present in 5 (7%) and 6 (9%) patients, respectively, while trisomy 12 was present in 11 (16%) and t(14q32) in 9 (13%) patients. The remaining 21 (31%) patients did not show aberrations by FISH. To better characterize the 20q13.12 gain, FISH analysis was also carried out in a validation series of 58 patients: 17% patients showed this alteration in ≥4.5% cells (ranging from 4.5% to 12%). In relation to mutational status, 49% of cases had unmutated *IgVH* gene.

aCGH showed recurrent genomic imbalances in CLL

Fifty of 67 patients (75%) displayed genomic changes with aCGH. In addition to the regions detected by FISH abnormalities, aCGH enabled the presence of novel recurrent genomic imbalances to be demonstrated. In order to rule out previously described single nucleotide polymorphisms, the minimal regions of overlap for all the recurrent lesions were compared with the frequencies of known copy number variations. A total of 443 altered chromosomal regions were found, of which 237 (53%) were deletions. The median number of changes per patient was five (range 0–14). The most commonly recurring alterations (observed in >5% of cases), their boundaries and frequencies are shown in Table 3. Losses in 13q14.2–q14.3 (21%), 11q13.3 (16%), 17p13.2–p13.1 (10%), 11q22.3–q23.1 (9%) and 5q13.3–q14.1, 5q31.1 and 7q22 (6% each) as well as gains in 1q21.3–q22 (22%), 11q13.3 (21%), 16q23.2–q24.2 (21%), 6p21.31–p21.1 (19%) and 10q22.3 (7.5%)

Downloaded from http://annonc.oxfordjournals.org/ by guest on January 8, 2012

Table 3. Recurrent aberrations identified by aCGH in CLL

Chromosome	Cytoband	Start position	End position	Size	Frequency of gains	Frequency of losses	Number of genes covered by the aCGH
1	q21.3–q22	151208615	155645185	4.44	22		31
1	q31.1–q31.2	189804292	192389810	2.59	13		1
5	q13.3–q14.1	74766678	77235911	2.47		6	18
5	q31.1	131568885	132631297	1.06		6	21
6	p21.31–p21.1	33722842	43897461	10.17	19		146
7	q22.1	98853924	101547956	2.69		6	78
10	q22.3	80565766	81502183	0.94	7.5		2
11	q13.1	64268965	65470124	1.2	22		51
11	q13.3–q13.4	69974549	73001497	3.03	21	16	37
11	q22.3–q23.1	108518932	112964950	3.73		9	24
12	p13.33–q24.33	1	133851895	133.9	15		
13	q14.2–q14.3	48863579	54941189	6.08		21	27
16	q23.2–q24.2	80972437	88692209	7.72	21		56
17	p13.2–p13.1	4571828	7483888	2.91		10	104
17	q25.3	75993754	80470659	4.48	19		105
18	q21.2	49844734	50270501	0.43	9	7.5	5
20	q13.12	42188467	44495323	2.31	19		52

Positions and sizes are expressed in base pairs. Bp locations according to GRCh37, February 2009 (hg 19).
aCGH, array comparative genomic hybridization; CLL, chronic lymphocytic leukemia.

were the most frequent changes revealed by aCGH (Table 3). Interestingly, a critical segment of gain was delineated on chromosome 20q in 13 patients (19%). The analysis identified a minimal region of gain on 20q13.12 of ~2.31 Mb involving three clones at linear positions (42 188 467–44 495 323), as shown in Figure 1. Most of these cases (75%) were studied at the time of diagnosis (Table 1). Changes detected in diagnostic and progression groups are shown in Table 4.

genomic abnormalities in the cytogenetic subgroups of CLL

Overall, correlation between FISH and aCGH was observed for +12, 11q- or 17p- cases (100%, 91% and 83%, respectively). However, this was not the case for 13q- subgroup. Interestingly, most of these cases displayed <30% of 13q deletions and in the 14 cases with a deletion of 13q revealed by aCGH, losses were located in 13q14.2–q21.1, with heterogeneous breakpoints. It should be noted that the 21 CLL samples showing no aberrations with FISH also appeared to be normal for the CLL FISH regions when analyzed by aCGH. However, novel recurrent alterations by aCGH were detected in this group of CLL patients (Table 5).

The analysis of the relationship between the recurrent abnormalities revealed by FISH and the presence of novel chromosomal imbalances detected by aCGH showed a significant association between the loss of 13q and the loss of 5q13.3–q14.1 and 5q31.1 ($P < 0.05$). No other additional abnormalities were observed in any of the cytogenetic subgroups.

In order to assess the genomic complexity in the cytogenetic CLL subsets, a comparison between the number of genetic changes ascertained by aCGH and the FISH CLL subgroup was carried out. As the median of changes per patient was five (range 0–14), we defined two groups to analyze the number of changes with respect to the FISH categories: ≤5 (low genomic complexity) and >5 (high genomic complexity) (Table 6). Interestingly, an association between the presence of a large number of changes detected by aCGH and ATM deletion ($P = 0.026$) by FISH was observed. The presence of gains on 20q12.13 ($P = 0.002$) was also associated with a high frequency of changes as revealed by aCGH (Table 6).

oligonucleotide and FISH studies validated the changes observed by aCGH

Oligonucleotide aCGH was carried out in 35 cases to confirm the BAC array results. Genomic patterns of gains and losses representative of the probe sizes (~150 kb) were compared with those obtained by BAC aCGH and found to be 100% concordant.

FISH experiments were carried out on 20 patients to confirm the gains on 20q13.12 observed with aCGH (supplemental Figure S1, available at *Annals of Oncology* online). All but one of the cases (95%) was concordant with the aCGH results. The median of cells showing this aberration was 20% (range 16%–25%). In addition, the cases were analyzed with a probe covering 51 992 266–52 410 801 bp (Vysis LSI ZNF217, breast tumor amplicon at 20q13.2). The results failed to show any

Downloaded from http://annonc.oxfordjournals.org/ by guest on January 8, 2012

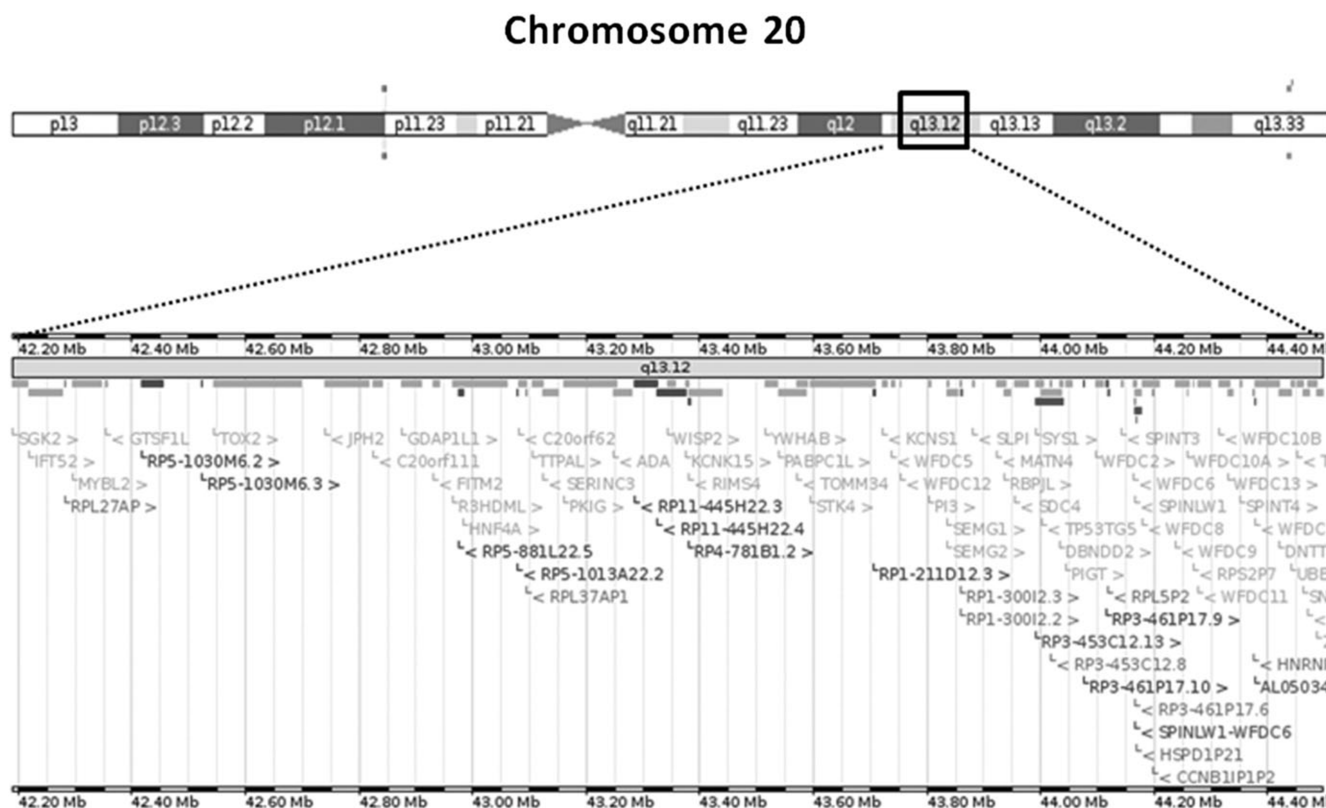


Figure 1. Integration of annotated genomic sequence with array comparative genomic hybridization data: common region of gain (CRG) on 20q (42188467–44495323 bp) showing the candidate genes (GRCh37, February 2009, hg 19).

Table 4. Characteristics of the CLL series (IgVH mutational status, number of aberrations, FISH subgroup and frequency of recurrent alterations detected by aCGH) according to the status of disease (at diagnosis versus progression)

Characteristics	Status of disease	
	At diagnosis (%)	Progression (%)
IgVH mutational status		
Mutated	63.6	46.7
Unmutated	36.4	53.3
Number of aberrations		
≤5	64	52.9
>5	36	47.1
FISH subgroup		
Normal FISH	38[a]	11.8
13q deletion	40	29.4
Trisomy 12	8	41.2[a]
17p deletion	6	17.6
11q deletion	6	11.8
t(14q32)	16	5.9
Recurrent alteration by aCGH		
Gains		
1q21.3–q22	24	17.6
1q31.2	8	29.4
6p21.31–p21.1	18	23.5
10q22.3	6	11.8
11q13.1	24	17.6
11q13.3	20	23.5
12	6	41.2[a]
16q23.2–q24.2	22	17.6
17q25.3	20	17.6
18q21.2	6	17.6
20q13.12	20	17.6
Losses		
5q13.3–q14.1	8	0
5q31.1	8	0
7q22	4	11.8
11q13.3	16	17.6
11q22.3–q23.1	8	11.8
13q14.2–q14.3	24	11.8
17p13.2–p13.1	10	11.8
18q21.2	8	5.9

[a]Statistically significant associations ($P < 0.05$).
aCGH, array comparative genomic hybridization; CLL, chronic lymphocytic leukemia.

involvement of this region, delineating the commonly gained region at 20q between 42 188 467–44 495 323 bp (2.31 Mb) (Figure 1).

gene expression profile confirmed the dosage effect of aCGH changes

In order to assess the relevance of the genomic imbalances in gene expression, a gene expression profile study was carried out. For this purpose, we grouped the cases by aCGH findings. The group of patients displaying trisomy 12 showed deregulation of 89 genes when compared with the rest of patients. A total of 76 of the 89 genes were overexpressed in relation to the other patients and 56% of them were located on chromosome 12.

It should be noted that overexpression of the 52 genes located in 20q13.12 (Figure 1), the 20q region gained by aCGH, was also observed ($P = 0.01$). Among these genes, we found well-known protein-coding cancer-related genes (supplemental Table S1, available at *Annals of Oncology* online) such as *PI3* (elafin), *SLPI* (secretory leukocyte peptidase inhibitor) and *WFDC2* [whey acidic protein (WAP) four-disulfide core domain 2], members of the WAP family; *PIGT* (phosphatidylinosotol glycan anchor biosynthesis, class T), a component of the glycosylphosphatidylinositol (GPI) glycan transamidase complex; *HNF4A* (hepatocyte nuclear factor 4, alpha) and *YWHAB* (tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, beta polypeptide), members of the SMAD and Ras signal transduction pathways, respectively. In addition, *ADA* (adenosine deaminase), a regulator of B-cell proliferation, overexpression was also present in CLL cases with gains on 20q.

Moreover, in patients with the 17p13 deletion, a significant proportion (83%) of the differentially underexpressed genes clustered in this region ($P < 0.05$). Among the downregulated genes were *GPS2* (G protein pathway suppressor 2)/*AMF1*, *SGSM2* (small G protein signaling modulator 2), *DRG2* (developmentally regulated GTP-binding protein 2), *SAT2* (spermidine/spermine N1-acetyltransferase family member 2) and *C17orf49* (chromosome 17 open reading frame 49). This gene dosage effect was also observed in CLL patients showing 11q-. Thus, all the genes located in the minimal region of deletion observed with aCGH on 11q22.3–q23.2 (108518932–112964950 bp) were downregulated when compared with the rest of patients ($P < 0.01$).

discussion

The presence of cytogenetic abnormalities is a hallmark of CLL. Indeed, these abnormalities have been associated with the prognosis or progression of the disease and for this reason the genetic changes have been extensively studied in CLL [30, 31]. The present study integrates genomic and gene expression profile analyses in a cohort of 67 CLL patients. Overall, the results enable us to detect hitherto undescribed recurrent alterations in CLL, such as gains on chromosome 20 and confirm the dosage effect, not only for the common cytogenetic abnormalities but also for this new genetic abnormality.

The present study found genomic copy number changes in 75% of the CLL patients. Our findings are similar to those previously reported in this disease [19, 22, 23, 32]. Detection rates of genomic alterations involving loci known to be associated with CLL occurred at expected frequencies [33] and overall, correlation between FISH and aCGH was observed except in the 13q- subgroup. Both FISH and aCGH revealed that 13q- was an heterogeneous group in size of the deletion and percentage of cells displaying the abnormality. Interestingly, when aCGH failed to demonstrate the presence of 13q deletion, FISH data revealed that most of these cases had <30%. This could justify, at least in part, the lack of correlation between both the techniques. We also confirmed that deletions are more abundant than gains in CLL: deletions in

Downloaded from http://annonc.oxfordjournals.org/ by guest on January 8, 2012

Table 5. Correlation between the most frequent chromosomal imbalances identified by aCGH and the CLL cytogenetic subgroups

Aberration/%	Cytogenetic/FISH subgroup						Total (%)
	13q14.3	Trisomy 12	11q22.3	17p13.1	t(14q32)	Normal FISH	
1q21.3–q22							
Gain	28	27	20		11	24	22
1q31.1–q31.2							
Gain	8	27	40	17	22		13
5q13.3–q14.1							
Loss	16[a]			17	11		6
5q31.1							
Loss	16[a]			17	11		6
6p21.31–p21.1							
Gain	24	27	20		22	14	19
7q22.1							
Loss	12	9	20		11		6
10q22.3							
Gain	16	9	20		11		7.5
11q13.1							
Gain	28	18	20	33	22	19	22
11q13.3–q13.4							
Loss	12	18	60[a]		22	14	16
Gain	28	18		17	22	19	21
11q22.3–q23.1							
Loss		18	100			5	9
12							
Gain		91					15
13q14.2–q14.3							
Loss	56			17	11		21
16q23.2–q24.2							
Gain	24	27	20	17	22	19	21
17p13.2–p13.1							
Loss	16			83			10
17q25.3							
Gain	24	9	20	17	22	14	19
18q21.2							
Loss	16			33			7.5
Gain	12	18		17		9	12
20q13.12							
Gain	24	27	20		22	14	19

Results are expressed as percentages.
[a]Statistically significant associations ($P < 0.05$).
aCGH, array comparative genomic hybridization; CLL, chronic lymphocytic leukemia.

chromosomes 5, 7, 11, 13, 17 and 18 and gains in chromosomes 1, 6, 10, 11, 12, 16, 17, 18 and 20 were present in this series. Regarding other recently reported alterations, we observed gain on 2p [34] in one case.

Our study identifies a previously undescribed recurrent region of gain in CLL, located on 20q13 in 19% of CLL patients. This frequency is similar to other well-characterized abnormalities in CLL (+12, 11q- and 17p-). It should be noted that gains in 20q were not associated with any other cytogenetic abnormality, although no patients with loss on 17p displayed 20q gains. The presence of 20q gains was not associated with mutational status either. Abnormalities of chromosome 20 are frequently observed aberrations in cancer [35–37]. In addition, the presence of gains on 20q has been associated with aggressive tumor behavior and poor clinical prognosis [38]. By contrast, deletions of the long arm of chromosome 20 are a common chromosomal abnormality associated with myeloid malignancies and are rarely seen in lymphoid malignancies [39]. A detailed analysis of 20q gains in cancer revealed that the size and location of the alteration are both variable. A region of gain at 20q13 was identified in CGH studies in human breast tumors [40]. The region has been analyzed at higher resolution, enabling three independently amplified regions to be characterized, with 20q13.2 being the most common region of gain in breast cancer. In the present study, FISH studies identified a minimal region of gain on 20q13.12 of ~2.31 Mb. This region is located close to the 20q breast cancer amplicon but is not included in it.

The gain on 20q13 in CLL could be relevant to the pathogenesis and evolution of CLL because 11 protein-coding cancer-related genes have been identified in this region (supplemental Table S1, available at *Annals of Oncology* online). It should be noted that all of these genes were upregulated in

Downloaded from http://annonc.oxfordjournals.org/ by guest on January 8, 2012

Table 6. Number of changes per patient in FISH groups and 20q13.12 cases

FISH	Number of aberrations (% cases)		Median of changes
	≤5	>5	
13q14.3	10 (40)	15 (60)	6
Trisomy 12	5 (45)	6 (55)	3
11q22.3	0	5[a] (100)	9
17p13.1	3 (50)	3 (50)	4
t(14q32)	5 (56)	4 (44)	3
Normal FISH	20[a] (95)	1 (5)	3
20q13.12 gain	3 (23)	10[a] (77)	7
Total (%)	4	33	

[a]Statistically significant associations ($P < 0.05$).

the CLL patients showing 20q gains in comparison with the other CLL cases. Thus, *PIGT*, *PI3*, *SLPI* and *WFDC2* could be potential candidate genes since they have been previously related to progression or tumor invasion. Phosphatidylinositol glycan (PIG) class T (*PIG-T*) is a component of the GPI transamidase complex and is amplified and overexpressed in human breast cancer cell lines and primary tumors [41]. Previous studies suggested that activation of the GPI transamidase complex could be a molecular mechanism underlying the progression of various human cancers [41, 42]. Interestingly, *GIP-S*, another GPI subunit, is located on 17p13.2, a region frequently deleted in cancer and in CLL. Therefore, further studies of these genes and their biological effects of all GPI transamidase complex subunits could be relevant in CLL. *PI3*, *SLPI* and *WFDC2* are members of the WAP family, a group of genes coding for proteins with a WAP motif. All of them have been identified as molecular markers for cancer and are clustered on chromosome 20q12–13.1. These genes are amplified and upregulated in several cancers [43]. The expression levels of all these genes were significantly higher in CLL cases with gains on 20q. Therefore, we suggest that 20q13.12 overexpressed genes may also be important in the evolution of CLL and warrant detailed study.

The present study also revealed a gene dosage effect in other chromosomal regions. Thus, CLL patients with trisomy 12 overexpressed genes located on chromosome 12, while patients with losses on 17p underexpressed genes located on 17p, as previously reported [16–18].

Gains in 20q13 in CLL did not occur as a single aberration because all CLL patients with gains in this region also had additional genetic changes. In fact, gains on 20q were associated with genomic complexity (Table 6). It is of note that genomic complexity has a significant impact on cancer prognosis and a number of studies have described the presence of several genomic changes as being predictors of disease progression and chemosensitivity in CLL [9, 44]. A significantly high level of genomic complexity in patients with loss on 11q was also observed. However, the CLL patients with losses on 17p did not have a large number of genomic alterations. This observation may indicate that the poor prognosis of patients with CLL exhibiting loss on 17p is unrelated to their genomic complexity [9]. The presence of a large number of genomic alterations in 20q13-gain patients suggest that this new genetic entity could be associated with a more advanced disease in CLL, as has been suggested in non-Hodgkin's lymphomas [45]. Genomic instability could therefore be another molecular feature of CLL progression, as has recently been suggested [46]. In order better to assess the clinical value of gain on 20q, a prospective study in a large series of CLL patients needs to be carried out.

Our results failed to demonstrate the presence of recurrent secondary genetic imbalances in the cytogenetic subgroups. In fact, only the group of patients with losses in 13q showed an association with losses in 5q13.3–q14.1 and 5q31. These changes had not been previously reported and could be examined further in subsequent studies.

In summary, our results demonstrated that submicroscopic 20q13.12 gains are common in CLL and confirmed that these gains result in an overexpression of the genes located on 20q13 [Figure 1, supplemental Table S1 (available at *Annals of Oncology* online)]. Furthermore, 20q gain is associated with great genomic complexity. These results suggest that the diversity of genomic aberrations in CLL is much greater than previously suggested. Further studies are needed to assess the prognostic significance of these alterations and how the genes located in these loci could contribute to the pathogenesis of CLL.

acknowledgements

We thank N Carter and H Fiegler (Sanger Center, Cambridge, UK) for providing us with the BACs library. We thank Irene Rodríguez, Sara González, Teresa Prieto, Mª Ángeles Ramos, Almudena Martín, Ana Díaz, Ana Simón, María del Pozo and Vanesa Gutiérrez of the Centro de Investigación del Cáncer, Salamanca, Spain, for their technical assistance.

funding

This work was partially supported by grants from the Spanish Fondo de Investigaciones Sanitarias (02/1041 and FIS 09/01543); Fondo Social Caja de Burgos de Investigación Clínica, Proyectos de investigación del SACYL (106/A/06) and by the 'Acción Transversal del Cáncer' project, through an agreement between the Instituto de Salud Carlos III (ISCIII), Spanish Ministry of Science and Innovation and the Cancer Research Foundation of Salamanca University and the Redes de Investigación RTIIC (FIS). AER is fully supported by an 'Ayuda predoctoral FIS de formación en investigación' by the Spanish Fondo de Investigaciones Sanitarias.

disclosure

The authors declare no conflicts of interest.

references

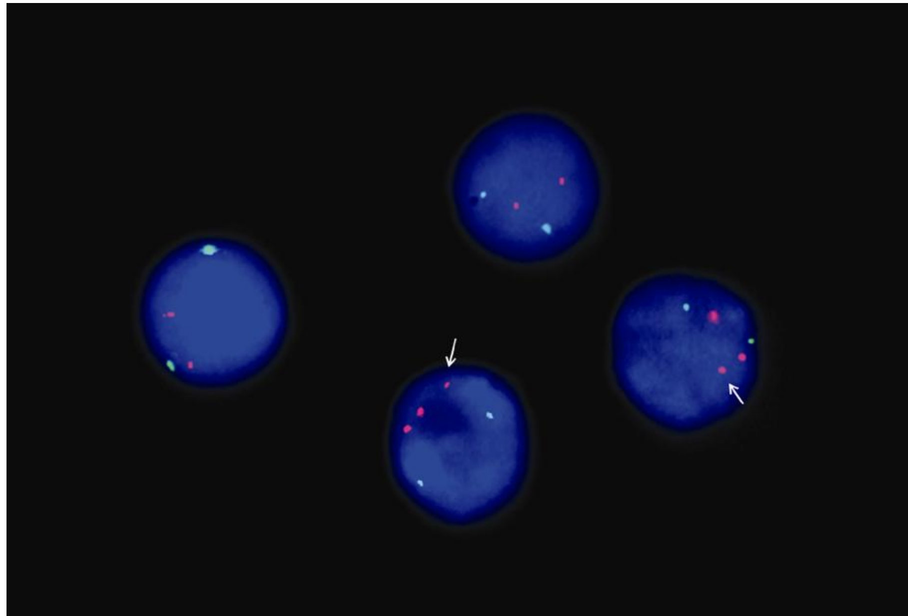
1. Chiorazzi N, Rai KR, Ferrarini M. Chronic lymphocytic leukemia. N Engl J Med 2005; 352: 804–815.
2. Dohner H, Stilgenbauer S, Benner A et al. Genomic aberrations and survival in chronic lymphocytic leukemia. N Engl J Med 2000; 343: 1910–1916.
3. Rozman C, Montserrat E. Chronic lymphocytic leukemia. N Engl J Med 1995; 333: 1052–1057.
4. Damle RN, Wasil T, Fais F et al. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. Blood 1999; 94: 1840–1847.

Downloaded from http://annonc.oxfordjournals.org/ by guest on January 8, 2012

5. Orchard JA, Ibbotson RE, Davis Z et al. ZAP-70 expression and prognosis in chronic lymphocytic leukaemia. Lancet 2004; 363: 105–111.
6. Rassenti LZ, Huynh L, Toy TL et al. ZAP-70 compared with immunoglobulin heavy-chain gene mutation status as a predictor of disease progression in chronic lymphocytic leukemia. N Engl J Med 2004; 351: 893–901.
7. Hernandez JA, Rodriguez AE, Gonzalez M et al. A high number of losses in 13q14 chromosome band is associated with a worse outcome and biological differences in patients with B-cell chronic lymphoid leukemia. Haematologica 2009; 94: 364–371.
8. Finn WG, Kay NE, Kroft SH et al. Secondary abnormalities of chromosome 6q in B-cell chronic lymphocytic leukemia: a sequential study of karyotypic instability in 51 patients. Am J Hematol 1998; 59: 223–229.
9. Kujawski L, Ouillette P, Erba H et al. Genomic complexity identifies patients with aggressive chronic lymphocytic leukemia. Blood 2008; 112: 1993–2003.
10. Kipps TJ. Genomic complexity in chronic lymphocytic leukemia. Blood 2008; 112: 1550.
11. Grubor V, Krasnitz A, Troge JE et al. Novel genomic alterations and clonal evolution in chronic lymphocytic leukemia revealed by representational oligonucleotide microarray analysis (ROMA). Blood 2009; 113: 1294–1303.
12. de Leeuw RJ, Davies JJ, Rosenwald A et al. Comprehensive whole genome array CGH profiling of mantle cell lymphoma model genomes. Hum Mol Genet 2004; 13: 1827–1837.
13. Kohlhammer H, Schwaenen C, Wessendorf S et al. Genomic DNA-chip hybridization in t(11;14)-positive mantle cell lymphomas shows a high frequency of aberrations and allows a refined characterization of consensus regions. Blood 2004; 104: 795–801.
14. Rubio-Moscardo F, Climent J, Siebert R et al. Mantle-cell lymphoma genotypes identified with CGH to BAC microarrays define a leukemic subgroup of disease and predict patient outcome. Blood 2005; 105: 4445–4454.
15. Tyybakinoja A, Saarinen-Pihkala U, Elonen E et al. Amplified, lost, and fused genes in 11q23-25 amplicon in acute myeloid leukemia, an array-CGH study. Genes Chromosomes Cancer 2006; 45: 257–264.
16. Haslinger C, Schweifer N, Stilgenbauer S et al. Microarray gene expression profiling of B-cell chronic lymphocytic leukemia subgroups defined by genomic aberrations and VH mutation status. J Clin Oncol 2004; 22: 3937–3949.
17. Porpaczy E, Bilban M, Heinze G et al. Gene expression signature of chronic lymphocytic leukaemia with trisomy 12. Eur J Clin Invest 2009; 39: 568–575.
18. Dickinson JD, Joshi A, Iqbal J et al. Genomic abnormalities in chronic lymphocytic leukemia influence gene expression by a gene dosage effect. Int J Mol Med 2006; 17: 769–778.
19. Pfeifer D, Pantic M, Skatulla I et al. Genome-wide analysis of DNA copy number changes and LOH in CLL using high-density SNP arrays. Blood 2007; 109: 1202–1210.
20. Gunn SR, Bolla AR, Barron LL et al. Array CGH analysis of chronic lymphocytic leukemia reveals frequent cryptic monoallelic and biallelic deletions of chromosome 22q11 that include the PRAME gene. Leuk Res 2009; 33: 1276–1281.
21. Patel A, Kang SH, Lennon PA et al. Validation of a targeted DNA microarray for the clinical evaluation of recurrent abnormalities in chronic lymphocytic leukemia. Am J Hematol 2008; 83: 540–546.
22. Schwaenen C, Nessling M, Wessendorf S et al. Automated array-based genomic profiling in chronic lymphocytic leukemia: development of a clinical tool and discovery of recurrent genomic alterations. Proc Natl Acad Sci U S A 2004; 101: 1039–1044.
23. Tyybakinoja A, Vilpo J, Knuutila S. High-resolution oligonucleotide array-CGH pinpoints genes involved in cryptic losses in chronic lymphocytic leukemia. Cytogenet Genome Res 2007; 118: 8–12.
24. Binet JL, Caligaris-Cappio F, Catovsky D et al. Perspectives on the use of new diagnostic tools in the treatment of chronic lymphocytic leukemia. Blood 2006; 107: 859–861.
25. Gonzalez MB, Hernandez JM, Garcia JL et al. The value of fluorescence in situ hybridization for the detection of 11q in multiple myeloma. Haematologica 2004; 89: 1213–1218.
26. Robledo C, Garcia JL, Caballero D et al. Array comparative genomic hybridization identifies genetic regions associated with outcome in aggressive diffuse large B-cell lymphomas. Cancer 2009; 115: 3728–3737.
27. Ghia P, Stamatopoulos K, Belessi C et al. ERIC recommendations on IGHV gene mutational status analysis in chronic lymphocytic leukemia. Leukemia 2007; 21: 1–3.
28. Gutierrez NC, Lopez-Perez R, Hernandez JM et al. Gene expression profile reveals deregulation of genes with relevant functions in the different subclasses of acute myeloid leukemia. Leukemia 2005; 19: 402–409.
29. Risueno A, Fontanillo C, Dinger ME et al. GATExplorer: genomic and transcriptomic explorer; mapping expression probes to gene loci, transcripts, exons and ncRNAs. BMC Bioinformatics 2010; 11: 221.
30. Di Bernardo MC, Crowther-Swanepoel D, Broderick P et al. A genome-wide association study identifies six susceptibility loci for chronic lymphocytic leukemia. Nat Genet 2008; 40: 1204–1210.
31. Crowther-Swanepoel D, Broderick P, Di Bernardo MC et al. Common variants at 2q37.3, 8q24.21, 15q21.3 and 16q24.1 influence chronic lymphocytic leukemia risk. Nat Genet 2010; 42: 132–136.
32. Ouillette P, Erba H, Kujawski L et al. Integrated genomic profiling of chronic lymphocytic leukemia identifies subtypes of deletion 13q14. Cancer Res 2008; 68: 1012–1021.
33. Gunn SR, Mohammed MS, Gorre ME et al. Whole-genome scanning by array comparative genomic hybridization as a clinical tool for risk assessment in chronic lymphocytic leukemia. J Mol Diagn 2008; 10: 442–451.
34. Jarosova M, Urbankova H, Plachy R et al. Gain of chromosome 2p in chronic lymphocytic leukemia: significant heterogeneity and a new recurrent dicentric rearrangement. Leuk Lymphoma 2010; 51: 304–313.
35. Yang SH, Seo MY, Jeong HJ et al. Gene copy number change events at chromosome 20 and their association with recurrence in gastric cancer patients. Clin Cancer Res 2005; 11: 612–620.
36. Zhu H, Lam DC, Han KC et al. High resolution analysis of genomic aberrations by metaphase and array comparative genomic hybridization identifies candidate tumour genes in lung cancer cell lines. Cancer Lett 2007; 245: 303–314.
37. Lassmann S, Weis R, Makowiec F et al. Array CGH identifies distinct DNA copy number profiles of oncogenes and tumor suppressor genes in chromosomal- and microsatellite-unstable sporadic colorectal carcinomas. J Mol Med (Berl) 2007; 85: 293–304.
38. Bar-Shira A, Pinthus JH, Rozovsky U et al. Multiple genes in human 20q13 chromosomal region are involved in an advanced prostate cancer xenograft. Cancer Res 2002; 62: 6803–6807.
39. Bench AJ, Nacheva EP, Hood TL et al. Chromosome 20 deletions in myeloid malignancies: reduction of the common deleted region, generation of a PAC/BAC contig and identification of candidate genes. UK Cancer Cytogenetics Group (UKCCG). Oncogene 2000; 19: 3902–3913.
40. Tanner MM, Tirkkonen M, Kallioniemi A et al. Independent amplification and frequent co-amplification of three nonsyntenic regions on the long arm of chromosome 20 in human breast cancer. Cancer Res 1996; 56: 3441–3445.
41. Wu G, Guo Z, Chatterjee A et al. Overexpression of glycosylphosphatidylinositol (GPI) transamidase subunits phosphatidylinositol glycan class T and/or GPI anchor attachment 1 induces tumorigenesis and contributes to invasion in human breast cancer. Cancer Res 2006; 66: 9829–9836.
42. Scotto L, Narayan G, Nandula SV et al. Identification of copy number gain and overexpressed genes on chromosome arm 20q by an integrative genomic approach in cervical cancer: potential role in progression. Genes Chromosomes Cancer 2008; 47: 755–765.
43. Clauss A, Lilja H, Lundwall A. A locus on human chromosome 20 contains several genes expressing protease inhibitor domains with homology to whey acidic protein. Biochem J 2002; 368: 233–242.
44. Kay NE, Eckel-Passow JE, Braggio E et al. Progressive but previously untreated CLL patients with greater array CGH complexity exhibit a less durable response to chemoimmunotherapy. Cancer Genet Cytogenet 2010; 203: 161–168.
45. Carter SL, Eklund AC, Kohane IS et al. A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers. Nat Genet 2006; 38: 1043–1048.
46. Stephens PJ, Greenman CD, Fu B et al. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. Cell 2011; 144: 27–40.

Downloaded from http://annonc.oxfordjournals.org/ by guest on January 8, 2012

Fig 2. FISH analysis using BACs as fluorescent probes (dJ1028D15-dJ781B1) on a BM sample from a CLL patient. Gain on 20q13.12 is identified as three red signals (arrows).



Supplementary Table 1. Cancer-related genes located in the minimal region of gain in 20q13. 12

Gene Symbol	Start position	End position	Size (bases)	Orientation	Gene Title	GO biological process term
HNF4A	42 984 340	43 060 030	75 691	plus strand	hepatocyte nuclear factor 4, alpha	transcription /// regulation of transcription, DNA-dependent /// negative regulation of cell proliferation /// SMAD protein signal transduction ///
SERINC3	43 124 862	43 150 750	25 889	minus strand	serine incorporator 3	induction of apoptosis
ADA	43 248 163	43 280 383	32 221	minus strand	adenosine deaminase	response to hypoxia/// germinal center B cell differentiation /// positive regulation of germinal center formation /// negative regulation of leukocyte migration /// negative regulation of mature B cell apoptosis/// positive regulation of B cell proliferation ///negative regulation of apoptosis /// positive regulation of calcium-mediated signaling /// positive regulation of T cell receptor signaling pathway ///
WISP2	43 343 485	43 357 150	13 666	plus strand	WNT1 inducible signaling pathway protein 2	regulation of cell growth /// cell adhesion /// signal transduction /// cell-cell signaling /// negative regulation of cell proliferation
YWHAB	43 514 317	43 537 173	22 857	plus strand	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, beta polypeptide	small GTPase mediated signal transduction /// Ras protein signal transduction /// activation of pro-apoptotic gene products /// negative regulation of transcription
TOMM34	43 570 771	43 589 127	18 357	minus strand	translocase of outer mitochondrial membrane 34	protein targeting to mitochondrion
PI3	43 803 517	43 805 185	1 669	plus strand	elafin	proteolysis// serine protease inhibitor
SLPI	43 880 879	43 883 206	2 328	minus strand	secretory leukocyte peptidase inhibitor	proteolysis// serine protease inhibitor
SDC4	43 953 928	43 977 064	23 137	minus strand	syndecan 4	positive regulation of protein kinase activity
PIGT	44 044 717	44 054 884	10 168	plus strand	phosphatidylinositol glycan anchor biosynthesis, class T	GPI anchor biosynthetic process /// attachment of GPI anchor to protein /// attachment of GPI anchor to protein /// neuron differentiation /// neuron apoptosis
WFDC2	44 098 346	44 110 172	11 827	plus strand	WAP four-disulfide core domain 2	proteolysis// serine protease inhibitor

Bp locations according to GRCh37, February 2009 (hg 19).

SUPPLEMENTARY METHODS

Mononuclear cells from all samples were isolated by Ficoll–Hypaque gradient centrifugation (Amersham Biosciences, Pittsburgh, PA), snap-frozen and stored at -80°C. DNA and RNA were isolated using TRIzol reagent (Invitrogen, Carlsbad, CA).

To ensure its good quality, DNA was measured using NanoDrop ND-1000 (ND-1000; NanoDrop Technologies, Wilmington, DE). Samples with OD 260nm/280 nm >1.8 were included. The integrity of the DNA was visually inspected on a 1% agarose gel.

Array Comparative Genomic Hybridization (aCGH)

BAC arrays

Tumor genomic DNA and reference DNA samples were differentially labeled by random priming with Cy3 and Cy5 fluorescent nucleotides, respectively. Labeled DNA was hybridized to 1 Mb BAC arrays produced at the Cancer Research Center (Salamanca, Spain) using a GeneTAC hybridization station (Genomic Solutions, Ann Arbor, MI) for 48 hours at 42°C according to the manufacturer's recommended protocol. These slides included 3,523 BACs spanning the genome at an average resolution of ~1 Mb. The particular BAC and P-1 derived artificial chromosome (PAC) set used to produce this array was distributed to academic institutions by the Wellcome Trust Sanger Institute (Cambridge, UK) and contains targets spaced at a density of 1 Mb over the entire genome, a set of subtelomeric sequences for each chromosome arm, and several hundred probes selected for their involvement in oncogenesis. The clone content is available online in the Cytoview windows of the Sanger Institute mapping database site (available at: http://www.ensembl.org/; accessed August 2008). Hybridized microarray slides were scanned with a GenePix 4000B microarray scanner

(Axon Instruments, Burlingame, CA, USA) and quantified using GenePix Pro4.0 microarray image analysis software (Molecular Devices, Sunnyvale, CA) [1].

Data Analysis

Cy3 and Cy5 images were scanned independently through two separate channels and the data were analyzed using the Cy5/Cy3 intensity ratio converted into $\log_2$ ratios. Array spot data showing low signal-to-background intensity ratios (signal <5 times that of the background) or low-quality replicate measurements were excluded from the analysis. Data from two-color hybridizations for both sets of DNA were normalized and the median of three spots was calculated using the DNMAD and preprocessing modules of GEPAS software [2-4]. Regions of copy number gains and losses for the BAC array-CGH data were identified by creating sample-specific thresholds using three algorithms: i. Circular Binary Segmentation, which estimates the location of change-points by calculating a likelihood ratio statistic for each probe and assesses its significance by permutation [5]; ii. Fused Lasso (cghFLasso), which smoothes the data with the fused lasso spatial smoothing technique. This algorithm is part of the cghFLasso package, which is available from http://www-stat.stanford.edu/~tibs/cghFLasso.html [6]; iii. a Gaussian Model with Adaptive Penalty (CGHseg), which estimates breakpoints by creating a cost matrix, locating all possible breakpoints from this matrix, and selecting the most likely number of breakpoints with an adaptive penalty. The algorithm has been rewritten in C, based on the MATLAB code made available by Picard et al. on their website (http://www.inapg.fr/ens_rech/maths/outil_A.html) [7]. The software enabled us to attribute a genomic status (gained, lost or normal) to each BAC clone. Chromosomal regions of copy-number change were identified as a minimum of two adjacent clones simultaneously deviating beyond the threshold values (≤-0.3 and ≥+0.3 for losses and gains, respectively) in the entire analysis in more than 5% of cases. This enabled us to characterize recurrent regions of gain or loss.

Oligonucleotide microarrays

This microarray platform provides measurements from 71,341 unique genomic loci. The hybridization, image analysis, extraction of fluorescence intensities and their $\log_2$ ratios, together with their subsequent normalization, were carried out following the manufacturer's recommendations. Gender-matched human DNA was used as reference. Each of the two reference samples, one female and one male, consisted of a pool of DNA from normal lymphocytes from several healthy people (G147A-G152A; Promega, Madison, WI). Briefly, 500 ng of tumor and reference DNA were labeled with Cy3 and Cy5, respectively. They were combined (4 µg each) and dried by vacuum centrifugation. The DNA was resuspended in 3.3 µL of Sample Tracking Control and vortexed; 8.7 µL of NimbleGen hybridization solution was added to the tube, mixed well, and heated at 95°C for 5 min in the dark. Samples were hybridized for 16–20 h at 42°C, then washed with NimbleGen wash buffers and scanned at 10-µm resolution using the GenePix 4000B dual scanner (Axon Instruments, Burlingame, CA, USA).

Data analysis

Data were extracted from scanned images using NimbleScan 2.5 extraction software (Roche NimbleGen, Madison, WI), which enables automated grid alignment, extraction, and generation of data files. Cy3 and Cy5 images were scanned independently through two separate channels and the quantified data were analyzed using the Cy5/Cy3 intensity ratio converted into $\log_2$. For mapping genomic breakpoints the segMNT v1.1 CGH segmentation analysis algorithm were run. The position of BAC clones and genes were determined according to GRCh37, February 2009 (hg 19). In the analysis of the array experiments all datasets were reviewed carefully for frequently affected chromosomal sites of physiological copy number polymorphisms. Every clone in the array was compared with the Database of Genomic Variants (available at

http://www.project.tcag.ca/cariation; accessed August 2008) and with the database of chromosomal imbalances and phenotype in humans using Ensembl Resources (DECIPHER: available at http://www.sanger.ac.uk/PostGenomics/decipher/; accessed August 2008) [8,9].

Gene expression profile analysis

The Robust Microarray Analysis (RMA) algorithm was used for background correction, intra- and inter-microarray normalization, and expression signal calculation [10]. The absolute expression signal for each gene (i.e., the signal value for each probe set) was calculated in each microarray, and Significance Analysis of Microarray (SAM) was done to calculate the significant differential expression and identify the gene probe sets characterizing the samples of each compared state. A cut-off false detection rate (FDR) of <0.05 was used in all the differential expression calculations.

Gene function analysis

The probe sets were functionally annotated and grouped according to their biological function using GeneOntology biological process descriptions. The data-sets of genes selected by statistical analysis were functionally interpreted with the aid of the DAVID bioinformatics resources [11,12].

REFERENCES

1. Chen W, Houldsworth J, Olshen AB et al. Array comparative genomic hybridization reveals genomic copy number changes associated with outcome in diffuse large B-cell lymphomas. Blood 2006; 107: 2477-2485.

2. Conde L, Montaner D, Burguet-Castell J et al. ISACGH: a web-based environment for the analysis of Array CGH and gene expression which includes functional profiling. Nucleic Acids Res 2007; 35: W81-W85.

3. Herrero J, Al-Shahrour F, Diaz-Uriarte R et al. GEPAS: A web-based resource for microarray gene expression data analysis. Nucleic Acids Res 2003; 31: 3461-3467.

4. Montaner D, Tarraga J, Huerta-Cepas J et al. Next station in microarray data analysis: GEPAS. Nucleic Acids Res 2006; 34: W486-W491.

5. Olshen AB, Venkatraman ES, Lucito R et al. Circular binary segmentation for the analysis of array-based DNA copy number data. Biostatistics 2004; 5: 557-572.

6. Tibshirani R, Wang P. Spatial smoothing and hot spot detection for CGH data using the fused lasso. Biostatistics 2008; 9: 18-29.

7. Picard F, Robin S, Lavielle M et al. A statistical approach for array CGH data analysis. BMC Bioinformatics 2005; 6: 27.

8. Sebat J, Lakshmi B, Troge J et al. Large-scale copy number polymorphism in the human genome. Science 2004; 305: 525-528.

9. Iafrate AJ, Feuk L, Rivera MN et al. Detection of large-scale variation in the human genome. Nat Genet 2004; 36: 949-951.

10. Irizarry RA, Bolstad BM, Collin F et al. Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res 2003; 31: e15.

11. Dennis G, Jr., Sherman BT, Hosack DA et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol 2003; 4: 3.

12. Huang dW, Sherman BT, Tan Q et al. DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. Nucleic Acids Res 2007; 35: W169-W175

Paper II

Hernández JA*, Rodríguez AE*, González M, Benito R, Fontanillo C, Sandoval V, Romero M, Martín-Núñez G, de Coca AG, Fisac R, Galende J, Recio I, Ortuño F, García JL, de las Rivas J, Gutiérrez NC, San Miguel JF, Hernández JM. **A high number of losses in 13q14 chromosome band is associated with a worse outcome and biological differences in patients with B-cell chronic lymphoid leukemia.** Haematologica. 2009 Mar;94(3):364-71.

*shared first authorship

A high number of losses in 13q14 chromosome is associated with a worse outcome and biological differences in patients with B-cell chronic lymphoid leukemia

José-Ángel Hernández,[1] Ana-Eugenia Rodríguez,[2] Marcos González,[3] Rocío Benito,[2] Celia Fontanillo,[2] Virgilio Sandoval,[4] Mercedes Romero,[5] Guillermo Martín-Núñez,[6] Alfonso García de Coca,[7] Rosa Fisac,[8] Josefina Galende,[9] Isabel Recio,[10] Francisco Ortuño,[11] Juan-Luis García,[2] Javier de las Rivas,[2] Norma-Carmen Gutiérrez,[2,3] Jesús-Fernando San Miguel,[2,3] and Jesús-María Hernández[2,3]

[1]Servicios de Hematología, Hospital Infanta Leonor, Madrid; [2]IBMCC, Centro de Investigación del Cáncer, Universidad de Salamanca-CSIC; [3]Hospital Clínico Universitario de Salamanca; [4]Hospital Virgen Blanca, León; [5]Hospital del Río Hortega, Valladolid; [6]Hospital Virgen del Puerto, Plasencia, Cáceres; [7]Hospital Clínico Universitario, Valladolid; [8]Hospital General de Segovia; [9]Hospital del Bierzo, León; [10]Hospital Nuestra Señora de Sonsoles, Ávila, an [11]Hospital Morales Meseguer, Murcia, Instituto de Estudios de Ciencias de la Salud de Castilla y León (IECSCYL), Unidad de Investigación, Hospital Universitario de Salamanca, Spain

ABSTRACT

Background
Among patients with B-cell chronic lymphoid leukemia, those with 13q14 deletion have a favorable outcome. However, whether the percentage of cells with 13q- influences the prognosis or the biological characteristics of this disease is unknown. We analyzed the clinico-biological characteristics and outcome of patients with B-cells chronic lymphoid leukemia with loss of 13q as the sole cytogenetic aberration.

Design and Methods
Three hundred and fifty patients with B-cell chronic lymphoid leukemia were studied. Clinical data were collected and fluorescence *in situ* hybridization and molecular studies were carried out. In addition, a gene expression profile was obtained by microarray-based analysis.

Results
In 109 out of the 350 cases (31.1%) loss of 13q was the sole cytogenetic aberration at diagnosis. In the subgroup of patients with 80% or more of cells with loss of 13q (18 cases), the overall survival was 56 months compared with not reached in the 91 cases in whom less than 80% of cells had loss of 13q ($p < 0.0001$). The variables included in the multivariate analysis for overall survival, were the percentage of losses of 13q14 ($p=0.001$) and B symptoms ($p=0.007$). The time to first therapy in the group with 80% or more *vs.* less than 80% of losses was 38 months *vs.* 87 months, respectively ($p=0.05$). In the multivariate analysis the variables selected were unmutated status of IgV_H ($p=0.001$) and a high level of $\beta_2$microglobulin ($p=0.003$). Interestingly, these differences regarding overall survival and time to first therapy were also present when other cut-offs were considered. The gene expression profile of patients with a high number of losses in 13q14 showed a high proliferation rate, downregulation of apoptosis-related genes, and dysregulation of genes related to mitochondrial functions.

Conclusions
Patients with B-cell chronic lymphoid leukemia with a high number of losses in 13q14 as the sole cytogenetic aberration at diagnosis display different clinical and biological features: short overall survival and time to first therapy as well as more proliferation and less apoptosis. A quantification of the number of cells showing a genetic abnormality should, therefore, be included in the study of the prognostic factors of B-cell chronic lymphoid leukemia.

Key words: B chronic lymphoid leukemia, 13q14 deletion, outcome, proliferation, apoptosis.

Citation: Hernández J-A, Rodríguez A-E, González M, Benito R, Fontanillo C, Sandoval V, Romero M, Martín-Núñez G, de Coca AG, Fisac R, Galende J, Recio I, Ortuño F, García J-L, de las Rivas J, Gutiérrez N-C, San Miguel J-F, and Hernández J-M. A high number of losses in 13q14 chromosome is associated with a worse outcome and biological differences in patients with B chronic lymphoid leukemia.Haematologica 2009; 94:364-371. doi:10.3324/haematol.13862

©2009 Ferrata Storti Foundation. This is an open-access paper.

Acknowledgments: we thank all the physicians from the Spanish hospitals who contributed clinical data we are also grateful to Eva Lumbreras, María Pozo, Teresa Prieto, María Ángeles Hernández, Ana Simón, Ana Díez, and Almudena Martín, from "Centro de Investigación del Cáncer, Salamanca" for their technical assistance.

Funding: partially supported by grants from "Proyectos de Investigación Biomédica del SACYL" 106/A/06; "Ayuda predoctoral FIS de formación en investigación" (AR), FI06/00126 and by Redes de Investigación RTIIC (FIS).

Manuscript received August 27, 2008. Revised version arrived October 17, 2008. Manuscript accepted November 11, 2008.

Correspondence: Jesús María Hernández, Servicio de Hematología y Departamento de Medicina Hospital Universitario de Salamanca Paseo San Vicente 58, 37007 Salamanca, Spain. E-mail: jmhr@usal.es

Introduction

B-cell chronic lymphoid leukemia (B-CLL) is a heterogeneous disorder both from genetic and prognostic points of view, with some patients displaying an indolent course while others have an aggressive disease with short survival.[1,2] In addition to classical prognostic factors,[3] new parameters such as immunophenotypic markers (ZAP70, CD38 and CD49d antigens),[4-6] molecular markers (mutational status of IGVH genes),[7,11] and cytogenetics[12-14] have been related to the prognosis in B-CLL.[15]

B-CLL patients show several cytogenetic aberrations, mainly in the regions of chromosomes 13q, 12, 11q, 17p, 14q and 6q. Some of these abnormalities can be better assessed by means of fluorescence *in situ* hybridization (FISH), which has shown that 62-80% of patients with B-CLL have cytogenetic abnormalities.[10,12] These cytogenetic changes are strongly correlated with the prognosis in terms of overall survival and time to progression (defined as the time to first therapy).[12,16-19] Patients with a deletion in 13q14 have a better outcome while patients with deletions in 11q23 or 17p13 have a shorter survival and shorter time to progression.[12] Classically, patients with B-CLL and a normal karyotype or trisomy 12 have been considered to have an intermediate prognosis.[12] It should, however, be noted that, in some series with a long-term follow-up, patients with B-CLL and a normal karyotype showed a better survival from 12 years on, as compared to patients with 13q-[9]. In addition, several studies have demonstrated that the percentage of cells displaying a particular cytogenetic abnormality (e.g. loss of *P53*)[20] or antigenic markers (e.g. CD38 or ZAP-70)[7] can be related to prognosis.

For these reasons, we decided to perform an analysis of patients diagnosed with B-CLL and deletion in 13q14, as the sole cytogenetic abnormality. The clinical features, including outcome, and the biological features of the patients displaying different degrees of infiltration by 13q- cells were assessed. Moreover, to gain further insights into the molecular mechanisms involved in 13q14 deletion B-CLL, a gene expression profile study using a microarray-based analysis was also carried out.

Design and Methods

Patients

The study population comprised 350 non-selected patients, from nine Spanish institutions, diagnosed with B-CLL. The diagnosis of B-cell was made according to World Health Organization (WHO) classification[21] and Working Group of National Cancer Institute (NCI) criteria.[22] Evidence of persistent lymphocytosis and a compatible immunophenotype were required for the diagnosis. In all cases an immunophenotypic analysis was performed by flow cytometry, including at least the following monoclonal antibodies: CD19, CD5, CD22, CD23, CD38, CD25, CD103, CD11c, FMC7, BCL2, CD10, CD20, and surface immunoglobulins κ and λ[23]. In addition, FISH studies, including specific probes for the regions 11q21, 12q13, 13q14, 14q32, and 17p13, were carried out. The study protocol was approved by local ethical committees and written informed consent was obtained from the patients.

Clinical data

Clinical data were recorded by reviewing the clinical histories of patients included in the study. In most cases (283 patients; 81%) the FISH study was performed at the time of diagnosis. In more than 95% (61 patients) of the remaining cases, (patients with a long follow-up), the FISH study was normal (25 patients) or showed alterations in 13q14 as the sole cytogenetic aberration (36 patients). Only six patients showed other cytogenetic alterations: two patients had 11q deletions, one had t(14q32) and three patients had two cytogenetic alterations (13q14 deletion plus t(14q32) in two cases and 13q14 deletion plus 11q deletion in the other one). Progression was defined according to previously reported criteria:[24] the presence of disease-related symptoms, massive or progressive organomegaly, bone marrow failure or recurrent infections.

Fluorescence in situ hybridization

Interphase FISH was performed on bone marrow samples using commercially available probes for the following regions: 13q14, 12q13, 11q22/*ATM*, 17p13/P53, and 14q32/*IGH* (Vysis/Abbott Co, Downers Grove IL, USA). The methods used for the FISH analysis have been described elsewhere.[25] 14q32 translocations, trisomy 12 and deletions were considered to be present when ≥5%, ≥3% and ≥8% interphase cells showed a split signal, three signals and one signal, respectively. Dual-color FISH using differently-labeled control probe and test probe was performed and signal screening was carried out on at least 200 cells with well-delineated signals. Hybridization was repeated on those slides with less than 80% cells showing two control-probe signals.

Mutation status of IGVH genes

Amplification and sequencing of *IGVH* genes was performed according to the ERIC recommendations on *IGHV* gene mutational status analysis in B-CLL.[26] Cases were classified as *IGVH* unmutated if there was at least 98% concordance between the tumor DNA and the respective family sequence, and *IGVH* mutated if there was less than 98% concordance.

Statistical analysis

Statistical tests were performed with SPSS 13.0 (SPSS, Chicago, IL, USA). The χ^2 test was used to assess associations between categorical variables, while continuous variables were analyzed with the Kruskal-Wallis test. The variables with statistical significance related to overall survival and time to first therapy were calculated by Kaplan-Meier method (*Log-rank*). Results were considered statistically significant for p values ≤0.05. Multivariate analysis of survival and time to first therapy was performed using the Cox regression method.

Gene expression profile analysis

Patients and samples

Bone marrow samples were obtained from 37 patients with B-CLL and deletion of 13q14 as the sole cytogenetic aberration at diagnosis. Fifteen had more than 80% of 13q- cells, while the remaining 22 cases had less than 80% of 13q- cells in the bone marrow. Mononuclear cells from all samples were isolated using Ficoll gradient, snap frozen and stored at -80°C. Both groups of patients showed more than 80% of clonal B-cell lymphocytes. RNA isolation, labeling and microarray hybridization were performed as previously reported.[27]

Normalization, signal calculation, significant differential expression, and sample/gene profile clustering

A robust microarray analysis algorithm was used for background correction, intra- and inter-microarray normalization, and expression signal calculation.[28-30] Once the absolute expression signal for each gene (i.e., the signal value for each probe set) had been calculated for each microarray, a method, called significance analysis of microarray,[31] was applied to calculate significant differential expression and find the gene probe sets that characterized the samples of each compared state. This method uses permutations to provide robust statistical inference of the most significant genes and provides *p* values adjusted to multiple testing using the false discovery rate (FDR).[32] A FDR of less than 0.05 was used for all the differential expression calculations. Finally, the resulting lists of candidate genes associated to a high degree with 13q34 band deletion were tested using another algorithm, the so-called global test,[33] which reveals the group of genes that has a global expression pattern most significantly related to the clinical feature studied. We applied all these methods using R and Bioconductor.

The function of the genes included in the expression signature of CLL with a high degree of 13q34 was assigned by applying the GeneCodis program34, that finds concurrent annotations in GO and KEGG, and thereby derives several groups of genes with functional significance. The functional analysis to identify the most relevant biological mechanisms, pathways and functional categories in the data sets of genes selected by statistical analysis was generated through the use of Ingenuity Pathways Analysis (Ingenuity Systems, Mountain View, CA, USA).

Results

Fluorescence *in situ* hybridization

Among the 350 B-CLL patients studied, 162 (46.3%) had 13q deletion. In 128 (36.6%), this aberration was the only cytogenetic abnormality. For the analysis of prognostic factors we restricted the study to the 109 out of these 128 patients in whom the 13q- was analyzed at diagnosis. Biallelic (homozygous) 13q deletion was present in 21.4% of patients, while the remaining patients had a monoallelic (heterozygous) deletion. Table 1 shows the salient cytogenetic features of the whole series of 350 patients.

Table 1. Incidence of genomic aberrations assessed by FISH in the global series of 350 patients with B-CLL.

FISH abnormality	N° of cases (%)
13q deletion	162 (46)
13q deletion as the sole alteration	128 (36)
13q deletion as the sole alteration at diagnosis	109 (31)
<80% of losses	91
≥80 of losses	18
Trisomy 12[a]	42 (12)
11q deletion[b]	30 (9)
IGH rearrangements[c]	20 (6)
17p deletion[d]	15 (4)
No FISH abnormalities	118 (33)
IGH mutated/unmutated[e]	125 (56)/ 99 (44)

[a]In 31 cases as the sole aberration. [b]In 16 cases as the sole aberration. [c]In 11 cases as the sole aberration. [d]In 7 cases as the sole aberration. [e]Performed in 224 patients.

Survival and time to progression

Overall survival and time to first therapy curves, according to cytogenetic abnormalities, in the total series of 350 patients are shown in Figure 1. There was a significant association between overall survival and the cytogenetic groups ($p<0.0001$). Thus, patients with loss of 13q as the sole anomaly and patients without abnormalities by FISH survied longer (median overall survival 159 months and not reached, respectively; median of total series, 154 months) (Figure 1A). The cytogenetic aberrations also influenced the time to first therapy, such that patients with 13q- as the sole abnormality and those with a normal karyotype had longer time to first therapy ($p<0.0001$) (Figure 1B).

Clinical and biological characteristics of patients showing 13q-

At the time of diagnosis 109 patients showed 13q deletion as the sole FISH abnormality, and the study was focused on this group of patients. The recruitment period started in October 1997 and finished in December 2006. All but three patients (with a score of 3) had a CLL immunophenotypic score of either 4 (49 cases) or 5 (58 cases).[35] The median age of this group was 65 years (range, 38-90 years) and there was a predominance of males (68%). The majority of patients had asymptomatic disease with clinical and biological characteristics of good prognosis. Thus, 81.7% of patients were in A stage according to the Binet staging system and only seven (6.4%) were in stage C. Regarding the percentage of cells with 13q-, the majority of cases (83.5%) had this abnormality in less than 80% of the analyzed cells. No relation was found between monoallelic and biallelic losses (22% in patients with <80% of 13q losses *vs.* 18% in the ≥80% group) in the 13q- patients. No major differences regarding clinical, biological, immunophenotypic or mutational status features were found in the cases with low (<80%) *vs.* high infiltration (≥80%) of 13q- cells, except for a high lymphocyte count (median of 14 *vs.* 19.7×10^9/L, respectively) ($p=0.007$) and a trend for an association with a diffuse pattern of bone marrow infil-

tration (17% *vs.* 40%; p=0.07) and splenomegaly (11% *vs.* 28%; p=0.07) in the group with high 13q- (Table 2). Of 109 cases considered, all 33 treated received, fludarabine-based therapies; in three cases allogeneic transplantation with reduced intensity conditioning was also performed.

Survival and time to progression according to degree of 13q-

A significantly longer survival was observed in the cohort of patients with losses in 13q in less than 80% of cells. Thus, in the subgroup of patients with 80% or more of cells with loss of 13q the overall survival was 56

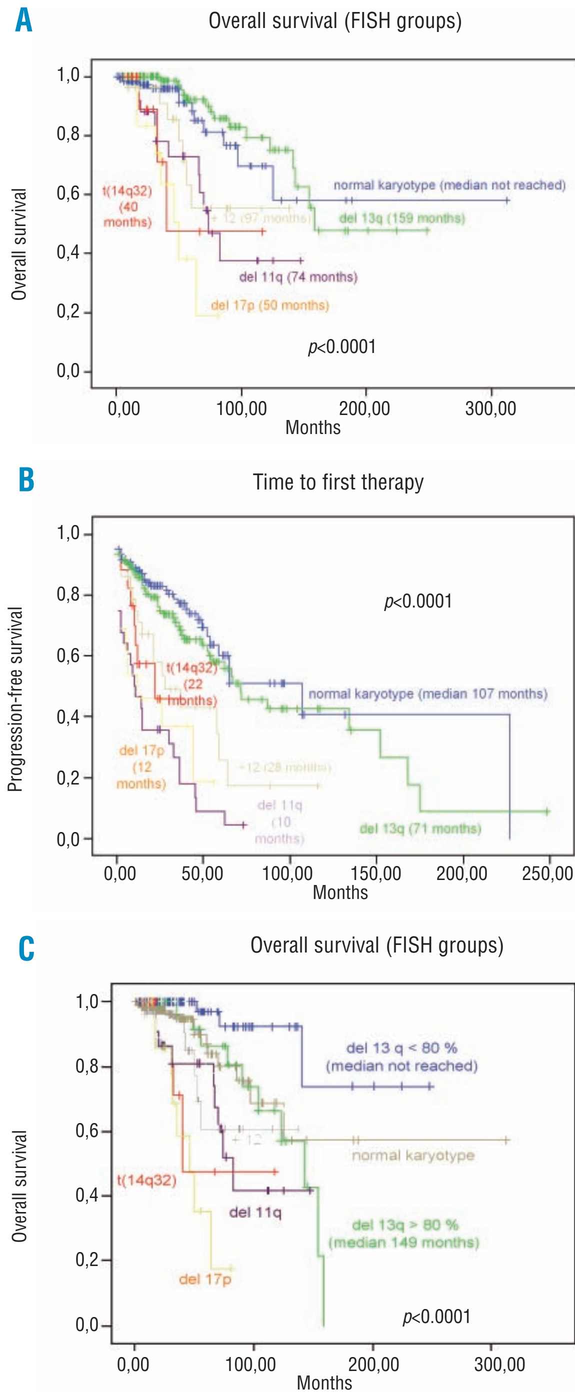


Figure 1. (A) Overall survival and (B) time to first therapy in the 350 patients with B-CLL. (C) Overall survival of patients ≥80% or <80% cells showing the 13q14 deletion against that of patients in the other cytogenetic groups. The median survival and median time to first therapy is showed in months for each cytogenetic group.

Table 2. Characteristics of 109 patients with 13q14 deletion as the sole cytogenetic aberration at diagnosis, divided according to the percentage of losses detected by FISH < 80% (n=91) or ≥80% (n=18).

Characteristic	FISH loss < 80, N= 91 (83.5%)	FISH loss ≥80, N=18 (16.5%)	*p* value
Age, years	66 (38-90)	69 (44-86)	0.58
White blood cells, n. (range), $\times 10^9$/L	19.4 (7.6-133.5)	26.9 (15.00-125.0)	0.008
Lymphocytes, n. (range), $\times 10^9$/L	14.0 (3.6-129.5)	19.7 (11.3-119.0)	0.007
Hemoglobin, n. (range), $\times 10^9$/L	14.4 (9-17.5)	14 (11-17.1)	0.91
Platelet count, n. (range), $\times 10^9$/L	183.0 (80.0-399.0)	182,000 (78.0-309.0)	0.44
IGH mutated	74%	92%	0.17
Biallelic/monoallelic 13q-*	16/69	4/13	0.43
Sex			
Male	61	13	0.49
Female	30	5	
Lactate dehydrogenase			
Normal	83	18	0.32
High	7	1	
$\beta_2$microglobulin			
Normal	56	9	0.21
High	28	8	
Binet stage			
A	76	13	0.48
B	9	4	
C	6	1	
Bone marrow pattern			
Diffuse	12	6	0.07
Other	58	9	
Lymphadenopathy			
Yes	26	5	0.95
No	65	13	
Hepatomegaly			
Yes	5	2	0.33
No	86	16	
Splenomegaly			
Yes	10	5	0.07
No	81	13	
B symptoms			
Yes	5	1	0.67
No	86	17	
CD38			
Positive	31	10	0.21
Negative	60	8	
Died during follow-up			
Yes	3	4	<0.0001
No	88	14	
Therapy during follow-up			
Yes	26	7	0.05
No	65	11	

**Number of cases.*

months (95% CI: 39-73 months), while in the group of patients in whom less than 80% of cells showed losses in 13q, the overall survival had not been reached (95% CI: 163-254 months) (p<0.0001) (Figure 2A). The proportion of deaths in both groups was 22.2% and 3.5%, respectively. Univariate analysis showed that six variables were associated with short overall survival (p<0.0001): high percentage of losses in 13q14; high level of serum lactate dehydrogenase; high level of $\beta_2$microglobulin; diffuse infiltration of the bone marrow; splenomegaly, and presence of B symptoms. In the multivariate analysis, the variables selected as independently related to overall survival were the percentage of losses of 13q14 (p=0.001) and the presence of B symptoms (p=0.007).

In addition, a significantly shorter time to first thepary was observed in the cohort of patients with 80% or more of cells showing losses in 13q (median of 38 months; 95% CI: 21-55 months) as compared to those cases with less than 80% of 13q- (median of 87 months; 95% CI: 21-153 months) (p=0.05) (Figure 2b). Thus, 38.8% of patients in the group with high infiltration required treatment *vs.* 28.9% of patients in the group with less than 80% of cells showing 13q- losses.

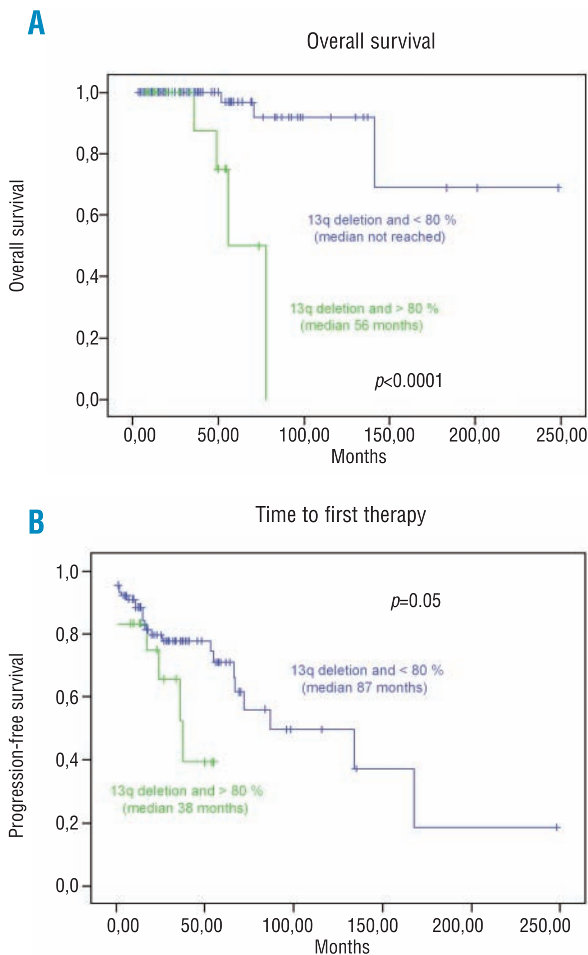


Figure 2. (A) Overall survival and (B) time to first therapy of patients with B-CLL and 13q14 deletion as the sole aberration and <80% or ≥80% FISH losses.

Univariate analysis showed that the variables associated with a short time to first therapy were: a high number of cells with deletion of 13q14 (p=0.05); a presence of biallelic losses in 13q14 (p=0.05); non-mutated pattern of *IGVH* genes; a high level of serum lactate dehydrogenase; a high level of serum $\beta_2$microglobulin; a positive Coombs' test of splenomegaly; presence of B symptoms; and a diffuse pattern of bone marrow infiltration (all with a p value <0.0001). In the multivariate analysis, the variables selected as independently related to time to first therapy were the mutational status of *IGVH* (p=0.001) and a high $\beta_2$microglobulin level (p=0.003). Interestingly, the differences were also observed when other cut-offs were analysed (i.e. ≤40%, 41-69%, ≥70%, *data not shown*, and ≤50%, 51-79% and ≥80%). (*Online Supplementary Figure S1*).

Gene expression profiles of the subsets of B-CLL patients displaying different degrees of 13q loss

For the gene expression profile analysis two groups of patients with 13q- were compared: those in whom 80% or more of cells showed 13q- (group A) and those in whom less than 80% of cells showed 13q losses (group B). The comparative analysis of the gene expression profile of both groups identified a set of 1755 differentially expressed genes: 1073 genes were upregulated in group A, and 682 were downregulated (Figure 3). The gene function analysis revealed that most of these genes are involved in apoptosis, cellular growth and proliferation, mitochondrial, endoplasmic reticulum (ER) and calcium mediated activity (Table 3). Thus, the patients with

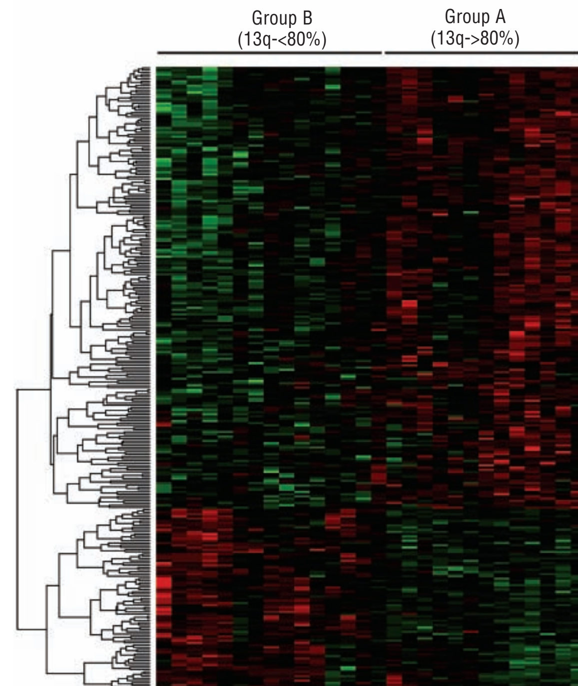


Figure 3. Unsupervised analysis of patients displaying 13q- in more than 80% of the cells (group A) and the cases with less than 80% of 13q-. Overall, 1755 genes were deregulated, most of them (n=1073) were upregulated in group A.

more than 80% of 13q- cells had overexpression of genes related to cell proliferation and MAP kinase activity such as *ATF2, FGF3, FGF10, FGFR2, RRAS2, GRB2, JUN, MAPT* and *MAP3K1*. By contrast, this group of patients had downregulation of genes related to apoptosis (*PDCD10, EGLN3, CASP6, CL*, and *DAPK1*) and cell cycle arrest, such as *CDKN2C, GAS2L1, UHMK1, GAS1, GAS2L3, ZAK*, and *GAS7*. Most of the genes involved in proteasome function were downregulated in the group with 13q losses in more than 80% of cells. The proteasome (*PSM)* genes dysregulated in our study codify proteasomal subunits, both catalytic (PSMA 1-7) and regulatory (PSMD1, PSMD7) ones. The ubiquitin process was also altered. Thus, genes for several enzymes that participate in different steps of the ubiquitin cycle were downregulated in the group with more losses in 13q: peptidases (USP27X), ubiquitin-conjugating enzymes (UBE2), thiolesterases (UCHL3) and ligases (HECTD1) (Table 3).

Table 3. Most relevant differentially expressed genes in patients with 80% or more cells showing 13q- (Up: overexpressed. Down: underexpressed).

Function	Genes	Expression
Apoptosis	*PDCD10, EGLN3, CLU, DAPK1, E2F1, UNC5B, NGFRAP1, HIPK2, APP, GLRX2, BCAP29, EIF2AK2, RAD21, BNIP2, ZBTB16, CASP6, SCARB1*	Down
Proliferation		
Cell cycle arrest	*CDKN2C, GAS2L1, UHMK1, GAS1, GAS2L3, ZAK, GAS7*	Down
MAPK signaling pathway	ATF2, FGF3, FGF10, FGFR2, RRAS2, *GRB2, JUN, MAPT, MAP3K1, GADD45B, PPP3CA, PPP3CC, PRKCA, PRKCG, RASGRF1, SOS1, TGFBR2, CACNA1A, CACNA1D, CACNA1E, CACNB1, MAP3K14, CACNA2D2, MAPK8IP1, CDC25B*	Up
Protein degradation Ubiquitin-mediated activity	*USP27X, PSMA1, PSMA2, PSMA3, PSMA4, PSMA5, PSMA6, PSMA7, UBE2D1, UBE2E1, UCHL3*	Down
	CYLD, USP22, USP40, STAMBPL1, USP28, USP32, USP6	Up

Discussion

In recent years several studies have demonstrated a relationship between genetic changes and outcome in B-CLL patients.[9,12,20] As regards cytogenetic abnormalities patients displaying a loss of *ATM* or *TP53* genes have a short survival. By contrast, both loss on 13q and the absence of cytogenetic aberrations assessed by FISH are related to long survival in B-CLL.[12] In most of the studies, patients displaying a loss of 13q were observed to have a better prognosis, but in some series with a long follow-up the outcome of patients with normal cytogenetics or 13q14 deletion was a similar[9] or even better for patients with normal karyotype. These results were reproduced in our series, with a median survival of 159 months for patients showing 13q- *vs.* a median not reached for those with a normal karyotype In order to better understand the clinical outcome of B-CLL patients with loss of 13q, we carried out a clinical and biological study.

Focusing on patients with loss of 13q as the sole cytogenetic aberration at diagnosis, we observed that the number of malignant cells carrying this genetic lesion influences the disease outcome. Thus, according to the percentage of cells with 13q- two prognostic groups could be established: the patients with a high proportion (≥80%) of 13q- cells had both a shorter overall survival than that of patients with <80% 13q- cells (56 months *vs.* not reached, $p<0.0001$) and a shorter time to first therapy (38 months *vs.* 87 months, $p=0.05$). These differences persisted when other cut-offs were considered (*Online Supplementary Figure S1*). The clinical relevance of the percentage of cells displaying a specific genetic abnormality has been recently demonstrated in CLL patients. Thus, the presence of more than 20% of cells with loss of *TP53* has been associated with an adverse prognosis, while patients with less than 20% of cells with loss of *TP53* had a prognosis similar to that of the global series.[20,36] In the present study, a high number of 13q- cells together with the presence of B-symptoms were the only independent adverse factors for a short survival in the multivariate analysis in the group of B-CLL patients with 13q-.

Our data support the concept that patients with 13q- do not constitute homogeneous group. Thus, patients with a high proportion of 13q losses had a high lymphocyte count and a trend to have more frequently a diffuse pattern of bone marrow infiltration and splenomegaly. In addition, overall survival and time to first therapy in this group were shorter than in patients with a normal karyotype or a trisomy 12, similar to that in patients with 11q deletion, although better than in the rest of cytogenetic subtypes. By contrast, CLL patients with a low number of malignant cells carrying the 13q deletion had a better prognosis than patients with a normal karyotype. It could, therefore, be useful to include quantification of the number of cells showing a genetic abnormality as part of the study of the prognostic factors in this disease.

To test our hypothesis that differences in several cellular functions such as more proliferation and less apoptosis could justify the clinical and prognostic differences found in the subsets of patients with 13q- divided according to the percentage of cells with this loss, we carried out a gene expression profiling analysis. This methodology has been applied in B-CLL showing that patients with specific genomic aberrations, gene expression phenotype or *IGVH* mutation status have distinct gene expression profiles.[37-39]

In the present study, we identified important functions differentially deregulated in the two groups of patients displaying 13q-. The main differential pathways involved are related to apoptosis, proliferation, mitochondrial and endoplastic reticulum function, as well as ubiquitin metabolism. Patients with more 13q- cells had overexpression of genes involved in proliferation. Thus, the *MAPK* signaling pathway was affected since *GRBS, RRAS, JUN* or *SOS1* were upregulated in

the group of B-CLL patients with high number of losses in 13q. In addition, some genes related to cell cycle arrest, such as *CDKN2C, GAS2L1, GAS1, GAS7, ZAK, GAS2L3* were downregulated in the group of patients with more 13q. Both the upregulation of *MAPK* and the underexpression of genes related to cell cycle arrest would lead to greater cell proliferation.[40] By contrast, genes related to apoptosis (*CASP6, CLU, DAPK1*, and *E2F1*) were found to be downregulated in the group of cases with a high proportion of 13q- cells, leading to a decrease in apoptotic activity. The accumulation of mature B cells that have escaped programmed cell death and undergone cell-cycle arrest in the G_0/G_1 phase is the hallmark of CLL.[41] Moreover, in our study we found alterations in other apoptosis-related pathways such as those mediated by mitochondria, endoplasmic reticulum and calcium metabolism. Thus, endoplasmic reticulum and protein-vesicular transport are inhibited because of an upregulation of genes involved in calcium-mediated activities, such as calmodulin binding and calcium ion binding.[42,43] Furthermore, mitochondrial oxidoreductase activity was inhibited in the group of B-CLL patients with a high number of cells with 13q-. The ATP dependency of apoptosis is well-known and both the respiratory chain and ATP synthase itself have been attributed central roles in the apoptotic process.[44,45] Therefore, patients with a high number of cells with 13q- not only had more proliferation, but also less apoptosis, which could be related to the more aggressive disease that we observed in the clinical study.

In recent years several hematologic malignancies have been treated with proteasome inhibitors.[46] Our study supports the idea that proteasome function is inhibited in the group of B-CLL patients with a high proportion of 13q-.[47] It is, therefore, logical to expect that such patients would not benefit from the use of proteasome inhibitors.

The *PI3K/NFkB* pathway seems to play a pivotal role in B-CLL cell survival and growth.[48] In our study, we found deregulation in *PI3K* beta and gamma catalytic subunits (*PI3KCB* and *PI3KCG*) as well as in other genes involved in this pathway (*RAS, c-JUN, PKC* and several integrins). This could be essential for the mitogenic and antiapoptotic functions. In addition *Ras-MAPK* and *Ras-PI3K* signaling pathways are related and both of them lead cells to greater proliferation.

In summary, our findings suggest that the number of malignant cells with 13q- can influence the outcome B-CLL patients in whom of 13q- is the single genetic abnormality. A high number of cells with 13q deletion, as quantified by FISH, is associated with short overall survival and time to progression. To the best of our knowledge, this is the first time that the percentage of losses in 13q14 has been related to survival and progression. In addition, several cellular functions such as more proliferation and less apoptosis were found in the subsets of 13q- patients.

Our results need to be confirmed by additional studies, preferably in the context of large, randomized clinical trials.

Authorship and Disclosures

J-AH designed and performed the research and statistical analyses, recorded and analyzed data and wrote the paper; A-ER: designed and performed the microarrays experiments and wrote the paper; CF, JLG, N-CG and JdR performed and interpreted the gene expression profiling studies; MG and J-MH designed and performed research and wrote the paper; VR, MR, GM-N, AGC, RF, JG, IR and FO provided patients' data and wrote the paper; J-FSM corrected and approved the final version of the manuscript.

The authors reported no potential conflicts of interest.

References

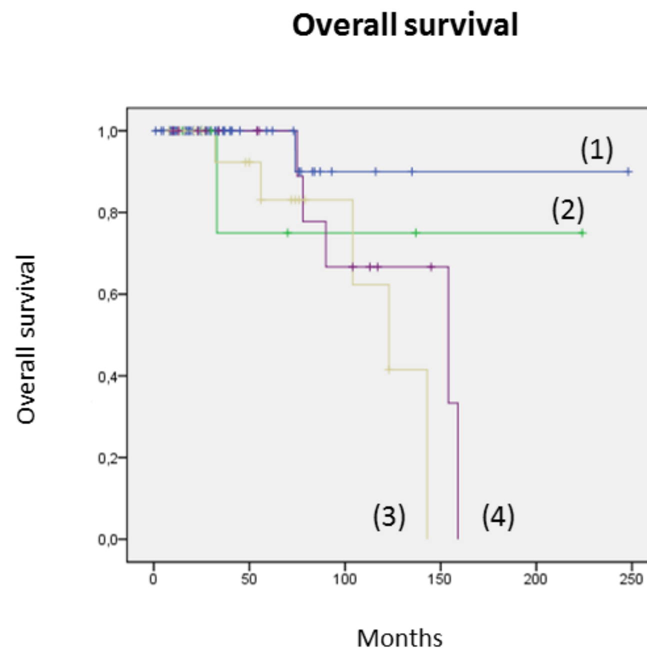
1. Chiorazzi N, Rai KR, Ferrarini M. Chronic lymphocytic leukemia. N Engl J Med 2005;352:804-15.
2. Dighiero G. Unsolved issues in CLL biology and management. Leukemia 2003;17:2385-91.
3. Montserrat E. New prognostic markers in CLL. Hematology Am Soc Hematol Educ Program 2006;279-84.
4. Crespo M, Bosch F, Villamor N, Bellosillo B, Colomer D, Rozman M, et al. ZAP-70 expression as a surrogate for immunoglobulin-variable-region mutations in chronic lymphocytic leukemia. N Engl J Med 2003; 348:1764-75.
5. Durig J, Naschar M, Schmucker U, Renzing-Kohler K, Holter T, Huttmann A, et al. CD38 expression is an important prognostic marker in chronic lymphocytic leukaemia. Leukemia 2002;16:30-5.
6. Gattei V, Bulian P, Del Principe MI, Zucchetto A, Maurillo L, Buccisano F, et al. Relevance of CD49d protein expression as overall survival and progressive disease prognosticator in chronic lymphocytic leukemia. Blood 2008;111:865-73.
7. Krober A, Seiler T, Benner A, Bullinger L, Bruckle E, Lichter P, et al. V(H) mutation status, CD38 expression level, genomic aberrations, and survival in chronic lymphocytic leukemia. Blood 2002;100:1410-6.
8. Lin K, Sherrington PD, Dennis M, Matrai Z, Cawley JC, Pettitt AR. Relationship between p53 dysfunction, CD38 expression, and IgV(H) mutation in chronic lymphocytic leukemia. Blood 2002;100:1404-9.
9. Damle RN, Wasil T, Fais F, Ghiotto F, Valetto A, Allen SL, et al. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. Blood 1999;94:1840-7.
10. Oscier DG, Gardiner AC, Mould SJ, Glide S, Davis ZA, Ibbotson RE, et al. Multivariate analysis of prognostic factors in CLL: clinical stage, IGVH gene mutational status, and loss or mutation of the p53 gene are independent prognostic factors. Blood 2002;100:1177-84.
11. Wiestner A, Rosenwald A, Barry TS, Wright G, Davis RE, Henrickson SE, et al. ZAP-70 expression identifies a chronic lymphocytic leukemia subtype with unmutated immunoglobulin genes, inferior clinical outcome, and distinct gene expression profile. Blood 2003;101:4944-51.
12. Dohner H, Stilgenbauer S, Benner A, Leupolt E, Krober A, Bullinger L, et al. Genomic aberrations and survival in chronic lymphocytic leukemia. N Engl J Med 2000;343:1910-6.
13. Hernandez JM, Mecucci C, Criel A, Meeus P, Michaux I, Van HA, et al. Cytogenetic analysis of B cell chronic lymphoid leukemias classified according to morphologic and immunophenotypic (FAB) criteria. Leukemia 1995;9:2140-6.
14. Krober A, Bloehdorn J, Hafner S, Buhler A, Seiler T, Kienle D, et al. Additional genetic high-risk features

such as 11q deletion, 17p deletion, and V3-21 usage characterize discordance of ZAP-70 and VH mutation status in chronic lymphocytic leukemia. J Clin Oncol 2006;24:969-75.
15. Hauswirth AW, Jager U. Impact of cytogenetic and molecular prognostic markers on the clinical management of chronic lymphocytic leukemia. Haematologica 2008;93: 14-9.
16. Cuneo A, Rigolin GM, Bigoni R, De AC, Veronese A, Cavazzini F, et al. Chronic lymphocytic leukemia with 6q- shows distinct hematological features and intermediate prognosis. Leukemia 2004;18:476-83.
17. Dohner H, Stilgenbauer S, James MR, Benner A, Weilguni T, Bentz M, et al. 11q deletions identify a new subset of B-cell chronic lymphocytic leukemia characterized by extensive nodal involvement and inferior prognosis. Blood 1997;89:2516-22.
18. Cavazzini F, Hernandez JA, Gozzetti A, Russo RA, De AC, Tiseo R, et al. Chromosome 14q32 translocations involving the immunoglobulin heavy chain locus in chronic lymphocytic leukaemia identify a disease subset with poor prognosis. Br J Haematol 2008; 142:529-37.
19. Cordone I, Masi S, Mauro FR, Soddu S, Morsilli O, Valentini T, et al. p53 expression in B-cell chronic lymphocytic leukemia: a marker of disease progression and poor prognosis. Blood 1998;91:4342-9.
20. Catovsky D, Richards S, Matutes E, Oscier D, Dyer MJ, Bezares RF, et al. Assessment of fludarabine plus cyclophosphamide for patients with chronic lymphocytic leukaemia (the LRF CLL4 trial): a randomised controlled trial. Lancet 2007;370:230-9.
21. Harris NL, Jaffe ES, Diebold J, Flandrin G, Muller-Hermelink HK, Vardiman J, et al. World Health Organization classification of neoplastic diseases of the hematopoietic and lymphoid tissues: report of the Clinical Advisory Committee meeting-Airlie House, Virginia, November 1997. J Clin Oncol 1999; 17:3835-49.
22. Binet JL, Caligaris-Cappio F, Catovsky D, Cheson B, Davis T, Dighiero G, et al. Perspectives on the use of new diagnostic tools in the treatment of chronic lymphocytic leukemia. Blood 2006;107:859-61.
23. Sanchez ML, Almeida J, Gonzalez D, Gonzalez M, Garcia-Marcos MA, Balanzategui A, et al. Incidence and clinicobiologic characteristics of leukemic B-cell chronic lymphoproliferative disorders with more than one B-cell clone. Blood 2003;102: 2994-3002.
24. Cheson BD, Bennett JM, Rai KR, Grever MR, Kay NE, Schiffer CA, et al. Guidelines for clinical protocols for chronic lymphocytic leukemia: recommendations of the National Cancer Institute-sponsored working group. Am J Hematol 1988;29:152-63.
25. Gonzalez MB, Hernandez JM, Garcia JL, Lumbreras E, Castellanos M, Hernandez JM, et al. The value of fluorescence in situ hybridization for the detection of 11q in multiple myeloma. Haematologica 2004;89: 1213-8.
26. Ghia P, Stamatopoulos K, Belessi C, Moreno C, Stilgenbauer S, Stevenson F, et al. ERIC recommendations on IGHV gene mutational status analysis in chronic lymphocytic leukemia. Leukemia 2007;21:1-3.
27. Gutierrez NC, Lopez-Perez R, Hernandez JM, Isidro I, Gonzalez B, Delgado M, et al. Gene expression profile reveals deregulation of genes with relevant functions in the different subclasses of acute myeloid leukemia. Leukemia 2005;19:402-9.
28. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 2003;4:249-64.
29. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res 2003;31:e15.
30. Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 2003;19:185-93.
31. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci USA 2001;98:5116-21.
32. Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I. Controlling the false discovery rate in behavior genetics research. Behav Brain Res 2001;125:279-84.
33. Goeman JJ, van de Geer SA, de KF, van Houwelingen HC. A global test for groups of genes: testing association with a clinical outcome. Bioinformatics 2004;20:93-9.
34. Carmona-Saez P, Chagoyen M, Tirado F, Carazo JM, Pascual-Montano A. GENECODIS: a web-based tool for finding significant concurrent annotations in gene lists. Genome Biol 2007;8:R3.
35. Matutes E, Owusu-Ankomah K, Morilla R, Garcia MJ, Houlihan A, Que TH, et al. The immunological profile of B-cell disorders and proposal of a scoring system for the diagnosis of CLL. Leukemia 1994;8: 1640-5.
36. Thornton PD, Gruszka-Westwood AM, Hamoudi RA, Atkinson S, Kaczmarek P, Morilla RM, et al. Characterisation of TP53 abnormalities in chronic lymphocytic leukaemia. Hematol J 2004;5:47-54.
37. Rosenwald A, Alizadeh AA, Widhopf G, Simon R, Davis RE, Yu X, et al. Relation of gene expression phenotype to immunoglobulin mutation genotype in B cell chronic lymphocytic leukemia. J Exp Med 2001;194:1639-47.
38. Klein U, Tu Y, Stolovitzky GA, Mattioli M, Cattoretti G, Husson H, et al. Gene expression profiling of B cell chronic lymphocytic leukemia reveals a homogeneous phenotype related to memory B cells. J Exp Med 2001;194:1625-38.
39. Aalto Y, El-Rifa W, Vilpo L, Ollila J, Nagy B, Vihinen M, et al. Distinct gene expression profiling in chronic lymphocytic leukemia with 11q23 deletion. Leukemia 2001;15:1721-8.
40. Smal C, Lisart S, Maerevoet M, Ferrant A, Bontemps F, Van Den NE. Pharmacological inhibition of the MAPK/ERK pathway increases sensitivity to 2-chloro-2'-deoxyadenosine (CdA) in the B-cell leukemia cell line EHEB. Biochem Pharmacol 2007;73:351-8.
41. Caligaris-Cappio F, Hamblin TJ. B-cell chronic lymphocytic leukemia: a bird of a different feather. J Clin Oncol 1999;17:399-408.
42. Lao Y, Chang DC. Study of the functional role of Bcl-2 family proteins in regulating Ca(2+) signals in apoptotic cells. Biochem Soc Trans 2007;35: 1038-9.
43. Bernardi P, Rasola A. Calcium and cell death: the mitochondrial connection. Subcell Biochem 2007;45: 481-506.
44. McClintock DS, Santore MT, Lee VY, Brunelle J, Budinger GR, Zong WX, et al. Bcl-2 family members and functional electron transport chain regulate oxygen deprivation-induced cell death. Mol Cell Biol 2002;22:94-104.
45. Shchepina LA, Pletjushkina OY, Avetisyan AV, Bakeeva LE, Fetisova EK, Izyumov DS, et al. Oligomycin, inhibitor of the F0 part of H+-ATP-synthase, suppresses the TNF-induced apoptosis. Oncogene 2002; 21:8149-57.
46. Richardson PG, Sonneveld P, Schuster MW, Irwin D, Stadtmauer EA, Facon T, et al. Bortezomib or high-dose dexamethasone for relapsed multiple myeloma. N Engl J Med 2005;352:2487-98.
47. Faderl S, Rai K, Gribben J, Byrd JC, Flinn IW, O'Brien S, et al. Phase II study of single-agent bortezomib for the treatment of patients with fludarabine-refractory B-cell chronic lymphocytic leukemia. Cancer 2006;107:916-24.
48. Cuni S, Perez-Aciego P, Perez-Chacon G, Vargas JA, Sanchez A, Martin-Saavedra FM, et al. A sustained activation of PI3K/NF-kappaB pathway is critical for the survival of chronic lymphocytic leukemia B cells. Leukemia 2004;18: 1391-400.

SUPPLEMENTARY FIGURE

Supplementary Figure 1.

Clinical outcome of del13q-only cases according to the cutoff of 80% of 13q deleted nuclei and the presence of a concomitant RB1 deletion.



Group 1. CLL patients with a low percentage of 13q- cells without a concomitant RB1 deletion.

Group 2. CLL patients with a low percentage of 13q- cells and a concomitant RB1 deletion.

Group 3. CLL patients with higher percentages of 13q- cells without a concomitant RB1 deletion.

Group 4. CLL patients with higher percentages of 13q- cells and a concomitant RB1 deletion.

CLL patients with lower percentages of 13q- cells by FISH were commonly associated with shorter deletions (not involving the RB1 locus). In addition, a correlation between the presence of either a short deletion or the presence of a low number of 13q- cells and a better outcome was present. However the patients with a short deletion and a high number of 13q- cells did not show a better prognosis than patients with high number of 13q- cells showing that the presence of high number of 13q- cells is always associated worst shorter survival independently of the deleted size on 13q-.

Paper III

Rodríguez AE, Hernández JA, Benito R, Gutiérrez NC, García JL, Hernández-Sánchez M, Risueño A, Sarasquete ME, Fermiñán E, Fisac R, García de Coca A, Martín-Núñez G, de las Heras N, Recio I, Gutiérrez O, de las Rivas J, González M, Hernández-Rivas JM. **Molecular characterization of chronic lymphocytic leukemia patients with a high number of losses in 13q14.** Submitted.

MOLECULAR CHARACTERIZATION OF CHRONIC LYMPHOCYTIC LEUKEMIA PATIENTS WITH A HIGH NUMBER OF LOSSES IN 13q14

RUNNING TITLE: Molecular characterization of 13q- CLL

Ana-Eugenia Rodríguez MS[1], José-Ángel Hernández MD,PhD[2], Rocío Benito PhD[1], Norma C Gutiérrez MD,PhD[3], Juan Luis García PhD[4], María Hernández Sánchez MS[1], Alberto Risueño MS[5], Mª Eugenia Sarasquete PhD[3], Encarna Fermiñán PhD[6], Rosa Fisac MD[7], Alfonso García de Coca MD[8], Guillermo Martín-Núñez MD[9], Natalia de las Heras MD[10], Isabel Recio MD[11], Oliver Gutiérrez MD[12], Javier de las Rivas PhD[5], Marcos González MD,PhD[1,3], Jesús-María Hernández Rivas MD,PhD[1,3]

1. IBSAL,IBMCC, Centro de Investigación del Cáncer, Universidad de Salamanca-CSIC. Salamanca, Spain.
2. Servicio de Hematología, Hospital Universitario Infanta Leonor, Madrid, Spain.
3. Servicio de Hematología, Hospital Clínico Universitario de Salamanca. Salamanca, Spain.
4. Instituto de Estudios de Ciencias de la Salud de Castilla y León, (IECSCYL)–HUSAL, Castilla y León, Spain
5. Unidad de Bioinformática y Genómica Funcional, IBMCC, Centro de Investigación del Cáncer, Universidad de Salamanca-CSIC. Salamanca, Spain.
6. Unidad de Genómica, IBMCC, Centro de Investigación del Cáncer, Universidad de Salamanca-CSIC. Salamanca, Spain.
7. Servicio de Hematología, Hospital General de Segovia, Segovia, Spain.
8. Servicio de Hematología, Hospital Clínico Universitario, Valladolid, Spain.
9. Servicio de Hematología, Hospital Virgen del Puerto, Plasencia, Spain.
10. Servicio de Hematología, Hospital Virgen Blanca, León, Spain.
11. Servicio de Hematología, Hospital Nuestra Señora de Sonsoles, Ávila, Spain.
12. Servicio de Hematología, Hospital del Río Hortega, Valladolid, Spain.

Correspondence:
Jesús María Hernández-Rivas
Servicio de Hematología y Departamento de Medicina
Hospital Universitario de Salamanca
Paseo San Vicente 58
37007 Salamanca, Spain
Fax: +34923294624
e-mail: jmhr@usal.es

Keywords: Chronic lymphocytic leukemia, FISH, 13q14 deletion, miRNA, apoptosis, proliferation

Text word count: 4537

Abstract word count: 200

Number of figures: 5

Number of tables: 3

Number of references: 46

ABSTRACT

Patients with CLL and 13q deletion as their only FISH abnormality could have a different outcome depending on the number of cells displaying this aberration. Thus, cases with a high number of 13q- cells (13q-H) had both shorter overall survival and time to first therapy. The goal of the study was to analyze the genetic profile of 13q-H CLLs. A total of 102 samples were studied, 32 of which served as a validation cohort and five were healthy donors. CLL patients with a high number of losses in 13q- showed a different level of gene expression as compared to patients with lower percentages of 13q- cells (13q-L). This deregulation affected genes involved in apoptosis and proliferation (BCR and NFkB signaling), leading to increased proliferation and decreased apoptosis in 13q-H patients. Deregulation of several microRNAs, such as miR-15a, miR-155, miR-29a and miR-223, was also observed in these CLLs. In addition, our study also suggests that the gene expression pattern of 13q-H CLL could be similar to the CLL cases with 11q- or 17p-. This study provides new evidence regarding the heterogeneity of 13q deletion in CLL patients, showing that apoptosis, proliferation as well as miRNA regulation are involved in 13q-H CLLs.

INTRODUCTION

Chronic lymphocytic leukemia (CLL) is characterized by the progressive accumulation of mature, monoclonal B lymphocytes in the blood, bone marrow (BM) and secondary lymphoid tissues.[1] The clinical course ranges from an indolent disorder with a normal lifespan to a rapidly progressive disease leading to death.[2,3] The variable clinical course of CLL is driven, at least in part, by the immunogenetic and molecular heterogeneity of the disease.[4,5] The genomic aberrations and the immunoglobulin (Ig) VH mutation status provide us with two separate genetic parameters of prognostic relevance. Thus, patients whose leukemic cells express unmutated *IgVH* regions (Ig-unmutated CLL) often have progressive disease, whereas those whose leukemic cells express mutated *IgVH* regions (Ig-mutated CLL) more often have an indolent disease.[4,6] Fluorescent in situ hybridization (FISH) can detect genomic abnormalities in more than 80% of CLL cases and the genetic subtypes of CLL show different biological and clinical features.[5] Although unfavorable aberrations (losses on 17p and 11q) are more frequent in the Ig-unmutated subgroup,[7-9] and favorable aberrations (loss on 13q as a single abnormality) are more frequent in the Ig-mutated subgroup, they have independent value in predicting outcome in CLL.[8,9]

Deletion at 13q14 (13q-) is the most common genomic aberration in CLL. It is present in more than 50% of cases, and is the sole documented cytogenetic abnormality in 36% of the patients. These latter cases are known to have a more favorable clinical course.[5,10] However, recent data from our group and others, suggest that patients with CLL and 13q deletion as the only FISH abnormality could have a different outcome depending on the number of cells displaying this aberration.[11-13] Moreover, previous studies had demonstrated that the percentage of cells displaying a particular cytogenetic abnormality (e.g. loss of *P53)*[14] or antigenic markers (e.g. CD38 or ZAP-70)[7] can be related to prognosis. We have demonstrated that cases with a high number of 13q- cells (13q-H) usually had both shorter overall survival and time to first therapy. However, to the best of

our knowledge the molecular characteristics of 13-H CLLs have not been so far analyzed in detail in order to better understand why these patients have a poor outcome.

The value of gene expression profiling (GEP) in the study of CLL is widely accepted. Such studies have identified new prognosis markers such as ZAP-70, LPL, PEG10 and CLLU1. Some of these are already well-established factors used in clinical practice, while the application of others is under study.

As a next step toward elucidation of biological differences within 13q- subgroup, the current study used the Affymetrix Human Exon arrays 1.0 ST, which offer a more fine-grained view of gene expression than the former generation of chips. Thus, the data obtained provide great insights into the biological mechanisms underlying the clinical differences observed in this CLL subgroup.[11-13]

MATERIAL AND METHODS

Patients

A total of 102 samples were selected for the study, 32 of which served as a validation cohort and five were healthy donors. CLL diagnosis was performed according to the World Health Organization (WHO) classification[15] and the Working Group of National Cancer Institute (NCI) criteria.[16] A complete immunophenotypic analysis by flow cytometry[17] and FISH studies were carried out in all cases. The median age at the time of study was 68 years (range, 35 to 90 years). Most patients were male (66%) and were in Binet clinical stage A (69%), while 26% were in stage B, and the remaining 5% were in stage C. The clinical and biological features of the CLL patients included in the study are shown in Supplementary Table S1. The study was approved by the local ethical committees. Informed consent was obtained from each patient before entering the study.

Methods

B cell isolation

Peripheral blood mononuclear cells (PBMCs) were isolated from fresh peripheral blood samples using Ficoll gradient, snap-frozen and stored at -80ºC.

For the validation cohort, CD19-positive B cells were purified by magnetically activated cell sorting (MACS) CD19 MicroBeads (Miltenyi Biotec, Bergisch Gladbach, Germany) resulting in a >98% purity, as analyzed by flow cytometry. CD19-positive normal B cells from peripheral blood of five healthy donors served as controls.

Genomic Aberrations

For the purpose of the study, only samples with one cytogenetic abnormality were included. For the gene expression profile analysis, according to our previous results,[11] two groups of patients with 13q- were compared: those in whom 80% or more of cells showed 13q- (13q-H) and those in whom fewer than 80% of cells showed 13q losses (13q-L). The distribution

of cases in the study cohort was: 13q-H (n=25; 36%), 13q-L (n=27; 39%), normal FISH (nCLL, n=8; 11%) and 17p-/11q- (n= 10; 14%).

In the validation cohort, the distribution of samples was similar: 13q-H (n=7; 22%), 13q-L (n=11; 34%) and nCLL (n=9; 28%). The remaining five cases were healthy donors.

Mutation status of IGVH genes

IGVH genes were amplified and sequenced according to the ERIC recommendations on IGHV gene mutational status analysis in CLL.[18]

Global gene expression using high density microarrays

Genome-wide expression analysis of the isolated samples was performed using Human Exon 1.0 ST microarrays (Affymetrix). RNA isolation, labeling and microarray hybridization were carried out following the manufacturer's protocols for the GeneChip platform by *Affymetrix*. Methods included synthesis of first- and second-strand cDNAs, the purification of double-stranded cDNA, synthesis of cRNA by in vitro transcription, recovery and quantization of biotin-labeled cRNA, fragmentation of this cRNA and subsequent hybridization to the microarray slide, post-hybridization washings, and detection of the hybridized cRNAs using a streptavidin-coupled fluorescent dye. Hybridized *Affymetrix* arrays were scanned with an *Affymetrix* Gene-Chip 3000 scanner. Images were generated and features extracted using Affymetrix GCOS Software.

Bioinformatic analysis: normalization, signal calculation, significant differential expression, and sample/gene profile clustering

The Robust Microarray Analysis (RMA) algorithm was used for background correction, intra- and inter-microarray normalization, and expression signal calculation.[19] The absolute expression signal for each gene was calculated for each microarray. For the expression signal calculation of the Human Exon arrays we used a new CDF package, called GeneMapper (from GATExplorer),[20] instead of the *Affymetrix* original probe-set definition.

This mapping represents an improvement thanks to the reannotation of updated Ensembl gene loci and removal of cross-hybridization noise.[20] It also allows operations to be carried out from the outset using gene identifications (Ensembl IDs) instead of probe-sets (*Affymetrix* IDs). Mapping to genome version Ensembl v53 (assembly NCBI36) was done for these analyses.

Significance Analysis of Microarray (SAM)[21] was used to calculate significant differential expression and to identify the gene probe sets that characterize the samples of each compared state. In this method, permutations provide robust statistical inference of the most significant genes and, by using a false discovery rate (FDR),[22] adjust the raw p-values to take multiple testing into account. An FDR cut-off of <0.05 was used for all the differential expression calculations.

Finally, the Global Test[23] algorithm was used to test the resulting lists of candidate genes associated with 13q-H subgroup. The Global Test allows us to identify the genes that have the global expression pattern most significantly related to the clinical feature studied.

All the bioinformatic analyses were performed with the statistical program R, using the custom packages Bioconductor[24] and GATExplorer.[20]

Principal Component Analysis

To explore and represent the differences among the different categories studied (13q-H CLLs, 13q-L CLLs, nCLLs and healthy controls), we applied Principal Component Analysis (PCA) to the expression data sets, using the normalized gene expression matrices of all samples of the validation cohort as the input. The expression matrices were filtered beforehand removing 25% of the least variable genes to avoid noise produced by non-expressed genes (i.e. the remaining 28 806 genes). For each of these genes, the median expression value across samples within each category was calculated. Next, the following formula was designed to calculate the expression values per gene and sample considering their variability within each category:

$$Y_{ij} = \frac{X_{ij(k)} - median(ik)}{sd(ik) + \beta} + median(ik)$$

where Y_{ij} is the PCA input matrix, X_{ij} is the original expression matrix, i is the gene, j the sample, k the category and β=2 is a small positive constant added to the denominator to ensure that the variance of Y_{ij} is independent of the genes.[21] This formula represents a way of calculating the dispersion of the biological replicates plus its median in each category. In this way, the clustering derived from the principal components includes a small amount of variation between individual samples, highlighting the differences between the categories.

Functional analysis and gene annotation

The functional assignment of the genes included in the expression signature of CLL cytogenetic subgroups was carried out by the Database for Annotation, Visualization and Integrated Discovery (DAVID) and the GeneCodis program,[25] which identifies concurrent annotations in GO and KEGG, and thereby constructs several groups of genes of functional significance. The most significant biological mechanisms, pathways and functional categories in the data sets of genes selected by statistical analysis were identified through the use of Ingenuity Pathways Analysis Sep2011 (Ingenuity Systems, Mountain View, CA, USA).

Gene-specific semi-quantitative PCR

Semi-quantitative SYBRgreen PCR was done in triplicate with iQ™ SYBR® Green Supermix kit (BioRad) using the IQ5 Multicolor Real-Time PCR Detection System (Bio-Rad). Expression data for selected genes were validated in a subset of CLL patients (n=40). Sense and antisense primers were designed based on the probe-sets used by Affymetrix to synthesize the GeneChip Primer sequences (Supplementary Table S2). The *ABL1* gene was used as the internal control and the expression data were analyzed by the comparative Ct method. The data were not normally distributed, so non-parametric tests

were used. Expression levels of the selected genes in both groups (13q-H and 13q-L) were analyzed using the Mann-Whitney U test with a two-tailed value of $P<0.05$ for statistical significance. All tests were performed using SPSS v19.0.

Quantification of miRNA expression levels

The expression of selected mature miRNAs was assayed using the Taqman MicroRNA Assays (Applied Biosystems) specific to hsa-mir-15a, hsa-mir-29a, hsa-mir-155 and hsa-mir-223 in 24 CLL patients displaying 13q- according to the manufacturer's recommendations. The Taqman MicroRNA Assays for U43 RNA (RNU43, Applied Biosystems) was used to normalize the relative abundance of miRNA using the $2^{-\Delta Ct}$ method. All experiments were performed in duplicate. Expression levels of the selected miRNAs in both groups (13q-H and 13q-L) were analyzed using the Mann-Whitney U test in SPSS v19.0. Values of $P<0.05$ were considered statistically significant.

Integrative analysis of miRNA and gene expression profile

A summary of the miRNA analysis performed in the study is shown in the Supplementary Figure S1. miRNAs with significantly different expression (FDR<0.05) between 13q-H and 13q-L were further analyzed to identify the networks and pathway targets. For this purpose, IPA's microRNA Target Filter, which enables prioritization of experimentally validated and predicted mRNA targets from TargetScan, TarBase, miRecords and the Ingenuity Knowledge Base was used. This tool identified the putative targets for the input miRNAs and then developed the networks among the targets and identified the known and most relevant biological functions, pathways and annotations in this enriched set of target genes. By applying the expression pairing tool, the analysis was focused on targets exhibiting altered expression in our analysis, finding miRNAs and their target genes with opposite or same expression.

RESULTS

13q-H CLLs are characterized by a specific genetic signature and miRNA expression

A total of 3 450 genes significantly distinguished 13q-H from 13q-L patients. These comprised 1 244 overexpressed genes and 2 206 underexpressed in the 13q-H group, defining the 13q-H signature. The deregulated genes of the 13q-H signature were annotated and analyzed for the presence of overrepresented "Gene Ontology categories" (Supplementary Table S3). The most significant overrepresented GO biological processes in 13q-H were related to cell cycle (P<0.0001), ribosome (P<0.0001) and regulation of transcription (P<0.0001). Moreover, 13q-H CLLs had higher levels of expression of *LEF1, BCL2, CARD11, HDAC9, NAFTC1, NFATC2, PAX5, FCRL2* and *SOS1*, while we identified several other genes downregulated in 13q-H, such as *GAS7, E2F1, RRM1, KIT, NP* and *EPOR*. Many of these genes have been reported to be deregulated in CLL, as we confirmed in our analyses that showed overexpression of *LEF1, NFATC1, NFATC2* and *PAX5* in B lymphocytes from CLL patients compared with B lymphocytes from healthy controls (data not shown). PCR results confirmed the microarray data in the analyzed genes such as *GAS7, E2F1* and *FCRL2* (Supplementary Figure S2).

Moreover, 13q-H CLL patients were also characterized by a striking overrepresentation of deregulated miRNAs. A total of 15 miRNAs were deregulated in 13q-H relative to 13q-L patients. Most of them (11) were downregulated while four were upregulated in 13q-H CLL (Table 1).

Signaling pathways and functional ontology analyses of genes differentially expressed in 13q-H patients

To determine the biological significance of the deregulated genes, a further analysis of the 3 450 deregulated genes characterizing the 13q-H CLL was carried out, revealing in this group of patients the involvement of several pathways (Table 2). These pathways are primarily related to cell proliferation, apoptosis and cell signaling. Thus the BCR pathway

was upregulated in 13q-H CLL patients. In fact, 21 genes from this pathway were overexpressed in 13q-H CLLs, some of which, such as *SYK*, *BLNK* and *PRKCB1*, were previously related to CLL pathogenesis (Supplementary Figure S3). We also observed an imbalance in proliferation and apoptosis in 13q-H patients, due to upregulation of antiapoptotic genes (*BCL2*) and decreased expression of proapoptotic genes (*RASSF5*, *BAD*, *CASP8*, *CASP10*, *FAS*) in 13q-H patients. Moreover, our analysis showed an overexpression of genes promoting proliferation, such as *LEF1*, *E2F5* and *RRAS2*. To ensure that the gene expression profiles accurately reflected the upregulation of BCR signaling pathway and the deregulation of apoptosis-related genes, representative genes that were differentially expressed in 13q-H patients were assessed by semi-quantitative SYBRgreen PCR analysis. These included *SYK*, *BLNK* and *PRKCB1* (BCR signaling pathway), *BCL2* (apoptosis) and *LEF1* and *RRAS2* (proliferation). The semi-quantitative PCR results were in close agreement with the microarray data (Figure 1) confirming the overexpression of these genes in 13q-H CLLs compared with 13q-L.

miRNA deregulation in 13q-H CLL patients

The analysis of miRNA expression in 13q-H and 13q-L CLL patients revealed that fifteen miRNAs were deregulated in 13q-H CLL patients: hsa-miR-155 was the most highly upregulated miRNA (Rfold=3.70), while hsa-miR-223 was the most significantly downregulated (Rfold=0.10). Four of the deregulated miRNAs (miR-15a, miR-29a, miR-155 and miR- 223) were further assayed by quantitative RT-PCR for validation purposes in 24 CLL samples displaying 13q-. Results confirmed the upregulation of miR-155 and the downregulation of miR-15a, miR-29a and miR- 223 in 13q-H samples relative to 13q-L (Figure 2).

The influence of these deregulated miRNAs on 13q- patients was assessed (Supplementary Figure S2). Specifically, we investigated whether observed changes in miRNAs were correlated with changes in the expression of genes. Therefore the post-transcriptional regulatory network of miRNA and genes in CLL patients with more than 80%

of 13q- cells was carried out by analyzing the miRNA-mRNA relationships. A total of 1 027 mRNA putative targets with altered expression in 13q-H CLL patients were found (Supplementary Table S4). Indeed, because miRNAs tend to downregulate the target genes, we focused our study on the subset of 11 miRNAs selected for analysis in IPA and the 432 genes predicted to be regulated by them and characterized by expression profiles stringly anticorrelated. Functional analysis revealed that transcription was the cell function most strongly affected by these miRNAs, with a total of 97 genes affected by the 11 selected miRNAs. Modification of proteins (n=41), proliferation of immune cells (n=34), and activation of protein binding sites (n=32) were other important functions affected by these miRNAs (Supplementary Table S5). Finally we performed a functional analysis of the 11 miRNAs and their 432 putative targets. The pathway analysis demonstrated that, again, B cell receptor signaling, PI3K signaling and NFkB signaling were among the most strongly affected pathways in 13q-H patients (Figure 3), highlighting the importance of miRNA regulation in CLL. MiR-155, the most overexpressed miRNA in 13q-H, was negatively correlated with the expression of 90 of the 182 expected genes (49%), demonstrating a relationship between miRNA and gene deregulation. Interestingly, most of these putative targets were assigned to the functional categories of transcription regulation (*P*=0.002). Moreover, we found several miRNAs whose targets that were experimentally observed or predicted with high confidence were strongly related to CLLs such as *BCL2* (miR-15, miR-206, miR-106b and miR-34a), *TCL1A* (miR-29a) and *LEF1* (miR-34a) (Table 3).

The GEP of 13q-H CLL patients is similar to that in CLL patients with 11q or 17p losses

We also analyzed the gene signature of CLL high risk cytogenetic subgroups in comparison with 13q- patients. Surprisingly, a significant number of deregulated genes were found to be shared between the genes that differentiate 13q subgroups and 13q-L and high risk subgroup of patients. That is, the GEP of 13q-H CLL patients resembled the gene expression pattern of patients with 17p- or 11q- abnormalities (Figure 4A). In fact, both

subgroups of CLL patients (13q-H and the 17p- and 11q- subgroup) shared 1 325 genes (46%) of the deregulated genes in the global analysis including all CLL subtypes. By contrast, the comparison between the GEP of 13q-H patients and those with losses in either 17p or 11q showed fewer differences in expression (Supplementary Figure S4).

To evaluate the biological significance of the observed similarity between the 13q-H and the 17p-/11q- signatures, we used the Ingenuity Pathway Analysis comparative tool, that facilitates the functional comparison of several panels of differentially expressed genes. Thus, we identified several commonly deregulated biological functions in both gene signatures (Figure 4B), such as cell cycle, cell death, cellular growth and proliferation. Finally, pathway analysis was performed on those genes commonly upregulated or downregulated in 13q-H, 17p- and 11q- patients in comparison with the 13q-L subgroup (Supplementary Table S6). In accordance with the comparative analysis results, several commonly deregulated pathways of relevance in CLL pathogenesis were observed. The most significant of these were the B cell receptor signaling pathway for commonly upregulated genes, and the cell cycle control of chromosomal replication pathway for commonly downregulated genes in patients showing 13q-H, 17p- or 11q- (Supplementary Table S6). The expression of the *TCL1* gene had one of the lowest q-values (0.002) with higher expression levels in patients with 13q-H, 17p- and 11q-. Of note, 13q-H, 17p- and 11q- patients also shared the deregulation of several miRNAs (Table 1).

Genome-wide expression differentiates 13q-H CLLs from 13q-L CLLs and controls

To validate the differences observed between the subgroups of 13q- CLL patients and get a visualization of these, we applied the Principal Component Analysis (PCA) in an independent series of patients. The clustering algorithm of PCA reduces complex multidimensional data to a few specified dimensions so that it can be visualized effectively. For a better characterization of the differences, we included in this cohort patients with normal FISH (nCLL) and healthy donors as two different types of controls.

Overall, the expression pattern of B lymphocytes from 13q-H and 13q-L CLL patients and nCLLs was notably different from the gene expression profile of B lymphocytes from healthy donors, as expected (Figure 5). PCA revealed a cumulative variance between groups of 48.3%, 60.9% and 68.3% corresponding to one, two and three of the initial components, respectively. Since the first three principal components explained a considerable proportion of the overall variance (68%), the 3D representation was able to show the main similarities and differences between categories. Notably, the 13q-H samples were largely separated from the others. Thus, 13q-H patients had a distinctive GEP that was different not only from healthy donors but also from all other CLLs, including 13q-L patients. By contrast, the gene expression of B lymphocytes from 13q-L CLL and nCLL was similar (Figure 5). SAM analysis revealed differences in the expression of 15 332 and 16 754 genes between CD19+ cells from 13q-L or nCLL compared with B lymphocytes from healthy donors, respectively, while both subgroups (13q-L and nCLL patients) shared the deregulation of 13 749 genes (data not shown). Moreover, the analysis failed to demonstrate differences between nCLL and 13q-L patients, while 131 genes were differentially expressed in comparison with 13q-H (data not shown).

Thus, both qualitatively (PCA) and quantitative (SAM) analysis showed that the gene expression profile of 13q- CLLs is different depending on the percentage of cells displaying this aberration.

IgVH mutational status in 13q-patients

Given that the prognostic significance of IgVH mutations is independent from that of cytogenetic abnormalities, we also analyzed the IgVH mutational status in the 13q- subgroups. There was no significant difference between both 13q- subgroups (P=0.664).

DISCUSSION

13q deletion (13q-) is the most common cytogenetic aberration in CLL and it is usually associated with the most favourable prognosis as a sole abnormality.[5] However, recent studies have shown that CLL patients carrying higher percentages of 13q- cells have more aggressive clinical courses.[11-13] By combining gene expression profile and miRNA analysis, we have shown that 13q- patients are also a biologically heterogeneous group, in which a higher number of 13q- cells (13q-H) could involve the deregulation of relevant cellular pathways. Thus, several pathways are involved in 13q-H patients (Table 2 and Supplementary Table S3), BCR signaling, NFkB signaling and antiapoptotic pathways being of special interest in CLL. Deregulation of several miRNAs (Table 1) was also observed. The influence of other factors with prognostic relevance in CLL, such as IGVH mutational status, was discarded.

The BCR is an essential signal transduction pathway for the survival and proliferation of mature B lymphocytes. In the present study, monoclonal B-cells in 13q-H CLL patients exhibit a molecular signature characterized by the overexpression of genes mainly involved in BCR signaling (Figure 1). There is now strong evidence that signaling via the B cell receptor plays a major role in the development of CLL, and it could be related to the different clinical outcomes of CLL.[26] Thus, the BCR pathway is activated in poor prognosis CLL patients (IGHV unmutated), and the overexpression of several molecules involved in this pathway has been reported in advanced stages of the disease.[27,28] In addition, SYK expression is enhanced in CLL relative to healthy B cells and also in unmutated compared with mutated CLL, possibly reflecting the increased BCR signaling in these patients.[29] In our study 13q-H CLL also overexpressed *SYK* (Figure 1), providing new evidence of the involvement of the BCR pathway in this group of CLLs. In addition, this group of patients also showed upregulation of CD79b. Chronic active BCR signaling due to point mutations in *CD79b* has recently been identified as a key pathogenic mechanism in aggressive B-cell lymphoma, and results in constitutive nuclear factor-kB (NF-kB) activation.[30] Interestingly,

CLL patients with deletions on 17p or 11q or those with losses in 13q in a high percentage of cells had an increased expression of a cluster of genes comprising several PKCs, such as *PRKCB1* and *PRKCZ.* Previous studies have shown an overexpression of *PKC* in human CLLs, which is part of a poor-prognosis gene cluster in CLL linked to the transmission of BCR signals such as calcineurin-NFAT and NF-kB, which our analysis also revealed to be deregulated (Table 2).[31,32] Furthermore, the overexpression of calcium metabolism-related genes as well as several MAPK in 13q-H patients was also observed in the present study, which would be consistent with these previous studies (Table 2).

One of the hallmarks of this clinically heterogeneous disease is defective apoptosis, which is considered to contribute not only to cell accumulation but also to disease progression and resistance to therapy.[26] In this study we report the overexpression of genes involved in promoting cell survival and antiapoptotic pathways, as well as the downregulation of several proapoptotic genes in 13q-H CLL patients (Table 2 and Supplementary Table S3). We confirm the overexpression of *LEF-1* in CLL B cells compared with B cells from healthy donors (data not shown), as previously reported,[33] but we also observed upregulation of *LEF-1* and other genes involved in the Wnt signaling pathway in 17p-, 11q- and 13q-H patients in comparison with 13q-L cases. Wnt pathway gene expression is widely known to be deregulated in CLL.[34,35] Alterations of RAS signaling are associated with potent oncogenic effects, which keep the cell in a proliferative state and block apoptosis, thereby paving the way for cancer formation. Overexpression of *RRAS* and other molecules involved in this signaling cascade, such as SOS1, RHOC and several MAP kinases, was also observed. In addition, apoptosis was also deregulated in 13q-H patients by the involvement of both mitochondrial (*BCL2* and several caspases) and extrinsic (*FAS*) pathways. Interestingly, the apoptotic signature of 13q-H patients showed a similar pattern of deregulation to that of high-risk cytogenetic groups (Figure 4B), since they both featured the alteration of several genes involved in the classic apoptotic pathway (mitochondrial). Sustained BCR signaling has also been reported to have an antiapoptotic effect.[36] Thus, in 13q-H CLL patients, our study shows an imbalance between the proliferative and apoptotic

signals, which could explain the higher level of lymphocytosis and the poor outcome previously described in these patients.[11]

An aberrant cellular miRNA expression profile in CLL cells has already been described and the changes correlate well with prognostic factors, including ZAP-70 expression status and *IgVH* mutations in CLL patients.[37] A recent study evaluating microRNAs as a signature for CLL patients with specific chromosomal abnormalities found nine miRNAs whose expression values were correlated with a specific karyotype.[38] In our study we found that several miRNAs were deregulated in 13q-H patients (Table 1), some of which had been previously reported in CLL (Table 3). Several important miRNAs, such as miR223, miR-29a and miR-181, were downregulated in 13q-H and high-risk cytogenetic subgroups, which could be related to the worse outcome in these groups of patients.[39,40] By contrast, overexpression of miR-155 was observed, which could be related to enhanced BCR-activation, as previously reported.[41] The pathogenic role of deletion 13q in CLL has been related to the lack of B-cell proliferation control allegedly determined by the deletion of the *DLEU2/MIR15A/MIR16-1* locus.[42] Interestingly, miR-15a was downregulated in 13q-H CLL patients and it has been reported to induce apoptosis through the negative regulation of *BCL2,* overexpressed in the 13q-H group of patients. It should be noted that a third of deregulated genes in 13q-H compared with 13q-L were putative targets of miRNAs also altered in this analysis, supporting the presence of a specific relationship between miRNA and gene expression in 13q-H CLL patients. Most of these genes were related to TGF or BCR signaling and confirmed these pathways to be those most commonly affected by miRNA deregulation in 13q-H patients. Among the putative target mRNAs we found many genes, such as *TCL1A, BCL2, LEF1*,[33,43,44] to be closely involved in CLLs (Table 3). These results suggest that miRNAs have a key role in the reported heterogeneity of 13q- patients. Surprisingly, our results suggest that some of the biological characteristics of 13q-H CLL patients are similar to those of high-risk cytogenetic subgroups, since they share the deregulation of several key signaling pathways (Figure 4B; Supplementary Table S6).

However, 13q-L patients had similar gene expression to that of CLL with normal FISH (Figure 5).

Therefore, this study provides new evidence regarding the heterogeneity of 13q deletion in CLL patients, showing that apoptosis, BCR and NF-kB signaling as well as miRNA regulation are the most significant affected pathways in 13q-H CLL patients. The identification of the mechanisms responsible for the clinical heterogeneity of CLL, including the mutations recently described[45,46] and the critical signaling pathways affected can lead to a better understanding of the molecular pathogenesis of the disease.

ACKNOWLEDGMENTS

Our study was partially supported by grants from the Spanish Fondo de Investigaciones Sanitarias 02/1041 and FIS 09/01543; Caja de Burgos-Banca Cívica, Proyectos de Investigación del SACYL 106/A/06 and by the Acción Transversal del Cáncer project, through an agreement between the Instituto de Salud Carlos III (ISCIII), the Spanish Ministry of Science and Innovation, the Cancer Research Foundation of Salamanca University and the Redes de Investigación RTIIC (FIS). AR is fully supported by an Ayuda Predoctoral FIS de Formación en Investigación by the Spanish Fondo de Investigaciones Sanitarias. We thank Irene Rodríguez, Sara González, Teresa Prieto, Mª Ángeles Ramos, Almudena Martín, Ana Díaz, Ana Simón, María del Pozo and Vanesa Gutiérrez of the Centro de Investigación del Cáncer, Salamanca, Spain, for their technical assistance and Jesús F. San Miguel for his critical review of the manuscript.

AUTHORSHIP CONTRIBUTION

A-ER designed the research, performed RT-qPCR assays, participated in the analysis and interpretation of the data, and wrote the manuscript; J-AH collected data, participated in discussions and critically reviewed the manuscript; RB and J-LG critically reviewed the manuscript; N-CG, EF, AR and JdR performed and interpreted the analysis of GEP experiments; MH-S performed RNA preparation and contributed to the statistical analysis; M-ES performed the miRNA expression analysis by Taqman assays; RF, AGC, GM-N, VS and IR provided patients' data; MG performed the IGHV mutational status analysis and critically reviewed the manuscript and J-MHR designed and supervised the study, performed research and wrote the manuscript.

CONFLICT- OF- INTEREST DISCLOSURE

The authors declare no competing financial interests.

Supplementary Information accompanies the paper on the Leukemia website

(http://www.nature.com/leu)

REFERENCES

1 Chiorazzi N, Rai KR, Ferrarini M. Chronic lymphocytic leukemia. *N Engl J Med* 2005; **352**: 804-815.

2 Keating MJ. Chronic lymphocytic leukemia. *Semin Oncol* 1999; **26**: 107-114.

3 Dighiero G. Unsolved issues in CLL biology and management. *Leukemia* 2003; **17**: 2385-2391.

4 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.

5 Dohner H, Stilgenbauer S, Benner A, Leupolt E, Krober A, Bullinger L et al. Genomic aberrations and survival in chronic lymphocytic leukemia. *N Engl J Med* 2000; **343**: 1910-1916.

6 Damle RN, Wasil T, Fais F, Ghiotto F, Valetto A, Allen SL et al. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. *Blood* 1999; **94**: 1840-1847.

7 Krober A, Seiler T, Benner A, Bullinger L, Bruckle E, Lichter P et al. V(H) mutation status, CD38 expression level, genomic aberrations, and survival in chronic lymphocytic leukemia. *Blood* 2002; **100**: 1410-1416.

8 Lin K, Sherrington PD, Dennis M, Matrai Z, Cawley JC, Pettitt AR. Relationship between p53 dysfunction, CD38 expression, and IgV(H) mutation in chronic lymphocytic leukemia. *Blood* 2002; **100**: 1404-1409.

9 Oscier DG, Gardiner AC, Mould SJ, Glide S, Davis ZA, Ibbotson RE et al. Multivariate analysis of prognostic factors in CLL: clinical stage, IGVH gene mutational status, and loss or mutation of the p53 gene are independent prognostic factors. *Blood* 2002; **100**: 1177-1184.

10 Mehes G. Chromosome abnormalities with prognostic impact in B-cell chronic lymphocytic leukemia. *Pathol Oncol Res* 2005; **11**: 205-210.

11 Hernandez JA, Rodriguez AE, Gonzalez M, Benito R, Fontanillo C, Sandoval V et al. A high number of losses in 13q14 chromosome band is associated with a worse outcome and biological differences in patients with B-cell chronic lymphoid leukemia. *Haematologica* 2009; **94**: 364-371.

12 Van Dyke DL, Shanafelt TD, Call TG, Zent CS, Smoley SA, Rabe KG et al. A comprehensive evaluation of the prognostic significance of 13q deletions in patients with B-chronic lymphocytic leukaemia. *Br J Haematol* 2010; **148**: 544-550.

13 Dal BM, Rossi FM, Rossi D, Deambrogi C, Bertoni F, Del G, I et al. 13q14 Deletion size and number of deleted cells both influence prognosis in chronic lymphocytic leukemia. *Genes Chromosomes Cancer* 2011; **50**: 633-643.

14 Catovsky D, Richards S, Matutes E, Oscier D, Dyer MJ, Bezares RF et al. Assessment of fludarabine plus cyclophosphamide for patients with chronic

lymphocytic leukaemia (the LRF CLL4 Trial): a randomised controlled trial. *Lancet* 2007; **370**: 230-239.

15 Harris NL, Jaffe ES, Diebold J, Flandrin G, Muller-Hermelink HK, Vardiman J et al. World Health Organization classification of neoplastic diseases of the hematopoietic and lymphoid tissues: report of the Clinical Advisory Committee meeting-Airlie House, Virginia, November 1997. *J Clin Oncol* 1999; **17**: 3835-3849.

16 Binet JL, Caligaris-Cappio F, Catovsky D, Cheson B, Davis T, Dighiero G et al. Perspectives on the use of new diagnostic tools in the treatment of chronic lymphocytic leukemia. *Blood* 2006; **107**: 859-861.

17 Sanchez ML, Almeida J, Gonzalez D, Gonzalez M, Garcia-Marcos MA, Balanzategui A et al. Incidence and clinicobiologic characteristics of leukemic B-cell chronic lymphoproliferative disorders with more than one B-cell clone. *Blood* 2003; **102**: 2994-3002.

18 Ghia P, Stamatopoulos K, Belessi C, Moreno C, Stilgenbauer S, Stevenson F et al. ERIC recommendations on IGHV gene mutational status analysis in chronic lymphocytic leukemia. *Leukemia* 2007; **21**: 1-3.

19 Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 2003; **19**: 185-193.

20 Risueno A, Fontanillo C, Dinger ME, de las RJ. GATExplorer: genomic and transcriptomic explorer; mapping expression probes to gene loci, transcripts, exons and ncRNAs. *BMC Bioinformatics* 2010; **11**: 221-

21 Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* 2001; **98**: 5116-5121.

22 Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I. Controlling the false discovery rate in behavior genetics research. *Behav Brain Res* 2001; **125**: 279-284.

23 Goeman JJ, van de Geer SA, de KF, van Houwelingen HC. A global test for groups of genes: testing association with a clinical outcome. *Bioinformatics* 2004; **20**: 93-99.

24 Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 2004; **5**: R80-

25 Carmona-Saez P, Chagoyen M, Tirado F, Carazo JM, Pascual-Montano A. GENECODIS: a web-based tool for finding significant concurrent annotations in gene lists. *Genome Biol* 2007; **8**: R3-

26 Chiorazzi N. Cell proliferation and death: forgotten features of chronic lymphocytic leukemia B cells. *Best Pract Res Clin Haematol* 2007; **20**: 399-413.

27 Guarini A, Chiaretti S, Tavolaro S, Maggio R, Peragine N, Citarella F et al. BCR ligation induced by IgM stimulation results in gene expression and functional changes only in IgV H unmutated chronic lymphocytic leukemia (CLL) cells. *Blood* 2008; **112**: 782-792.

28 Rodriguez A, Villuendas R, Yanez L, Gomez ME, Diaz R, Pollan M et al. Molecular heterogeneity in chronic lymphocytic leukemia is dependent on BCR signaling: clinical correlation. *Leukemia* 2007; **21**: 1984-1991.

29 Buchner M, Fuchs S, Prinz G, Pfeifer D, Bartholome K, Burger M et al. Spleen tyrosine kinase is overexpressed and represents a potential therapeutic target in chronic lymphocytic leukemia. *Cancer Res* 2009; **69**: 5424-5432.

30 Davis RE, Ngo VN, Lenz G, Tolar P, Young RM, Romesser PB et al. Chronic active B-cell-receptor signalling in diffuse large B-cell lymphoma. *Nature* 2010; **463**: 88-92.

31 Su TT, Guo B, Kawakami Y, Sommer K, Chae K, Humphries LA et al. PKC-beta controls I kappa B kinase lipid raft recruitment and activation in response to BCR signaling. *Nat Immunol* 2002; **3**: 780-786.

32 Bernal A, Pastore RD, Asgary Z, Keller SA, Cesarman E, Liou HC et al. Survival of leukemic B cells promoted by engagement of the antigen receptor. *Blood* 2001; **98**: 3050-3057.

33 Gutierrez A, Jr., Tschumper RC, Wu X, Shanafelt TD, Eckel-Passow J, Huddleston PM, III et al. LEF-1 is a prosurvival factor in chronic lymphocytic leukemia and is expressed in the preleukemic state of monoclonal B-cell lymphocytosis. *Blood* 2010; **116**: 2975-2983.

34 Lu D, Zhao Y, Tawatao R, Cottam HB, Sen M, Leoni LM et al. Activation of the Wnt signaling pathway in chronic lymphocytic leukemia. *Proc Natl Acad Sci U S A* 2004; **101**: 3118-3123.

35 Reya T, O'Riordan M, Okamura R, Devaney E, Willert K, Nusse R et al. Wnt signaling regulates B lymphocyte proliferation through a LEF-1 dependent mechanism. *Immunity* 2000; **13**: 15-24.

36 Longo PG, Laurenti L, Gobessi S, Sica S, Leone G, Efremov DG. The Akt/Mcl-1 pathway plays a prominent role in mediating antiapoptotic signals downstream of the B-cell receptor in chronic lymphocytic leukemia B cells. *Blood* 2008; **111**: 846-855.

37 Calin GA, Ferracin M, Cimmino A, Di LG, Shimizu M, Wojcik SE et al. A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. *N Engl J Med* 2005; **353**: 1793-1801.

38 Visone R, Rassenti LZ, Veronese A, Taccioli C, Costinean S, Aguda BD et al. Karyotype-specific microRNA signature in chronic lymphocytic leukemia. *Blood* 2009; **114**: 3872-3879.

39 Stamatopoulos B, Meuleman N, Haibe-Kains B, Saussoy P, Van Den NE, Michaux L et al. microRNA-29c and microRNA-223 down-regulation has in vivo significance in chronic lymphocytic leukemia and improves disease risk stratification. *Blood* 2009; **113**: 5237-5245.

40 Li S, Moffett HF, Lu J, Werner L, Zhang H, Ritz J et al. MicroRNA expression profiling identifies activated B cell status in chronic lymphocytic leukemia cells. *PLoS One* 2011; **6**: e16956-

41 Yin Q, Wang X, McBride J, Fewell C, Flemington E. B-cell receptor activation induces BIC/miR-155 expression through a conserved AP-1 element. *J Biol Chem* 2008; **283**: 2654-2662.

42 Cimmino A, Calin GA, Fabbri M, Iorio MV, Ferracin M, Shimizu M et al. miR-15 and miR-16 induce apoptosis by targeting BCL2. *Proc Natl Acad Sci U S A* 2005; **102**: 13944-13949.

43 Pekarsky Y, Santanam U, Cimmino A, Palamarchuk A, Efanov A, Maximov V et al. Tcl1 expression in chronic lymphocytic leukemia is regulated by miR-29 and miR-181. *Cancer Res* 2006; **66**: 11590-11593.

44 Calin GA, Cimmino A, Fabbri M, Ferracin M, Wojcik SE, Shimizu M et al. MiR-15a and miR-16-1 cluster functions in human leukemia. *Proc Natl Acad Sci U S A* 2008; **105**: 5166-5171.

45 Puente XS, Pinyol M, Quesada V, Conde L, Ordonez GR, Villamor N et al. Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. *Nature* 2011; **475**: 101-105.

46 Quesada V, Conde L, Villamor N, Ordonez GR, Jares P, Bassaganyas L et al. Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. *Nat Genet* 2011; **44**: 47-52.

FIGURES

Figure 1.

Gene expression levels of genes significantly upregulated in 13q-H CLL patients.

Box plot of the expression levels of six genes with significantly different expression between 13q-H and 13q-L patients, assessed by semi-quantitative PCR analysis. Box plots show the relative upregulation of BCR (SYK, PRKCB1 and BLNK), proliferation (LEF1 and RRAS2) and antiapoptotic (BCL2) related genes in patients with a high number of 13q- cells compared with CLL patients with lower percentages of losses in 13q. The thick line inside the box plot indicates the median expression levels and the box shows the 25th and 75th percentiles, while the whiskers show the maximum and minimum values. Outliers (extreme values falling out of the main distribution) are represented by open circles. Statistical significance was determined using the Mann-Whitney U test (*P*<0.05).

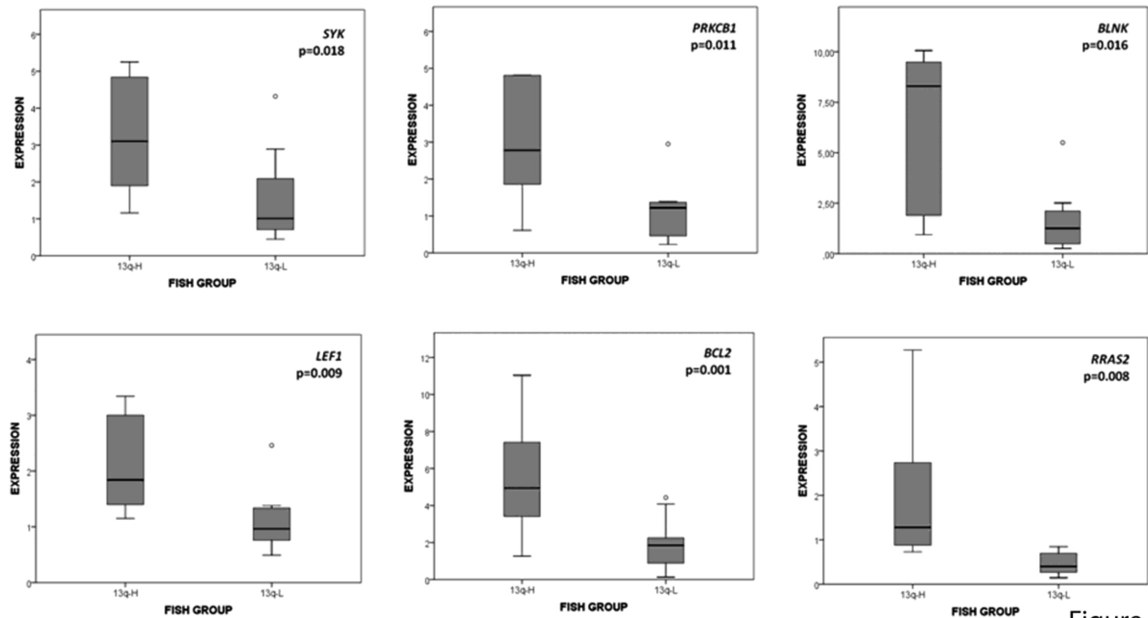


Figure 2.

Quantitative RT-PCR validation for miR-15a, miR-29a, miR-155 and miR-223 in independent CLL patients. Relative expression of miR-15a, miR-29a, miR-155 and miR-223 was evaluated by individual TaqMan miRNA assays performed in duplicate and normalized to RNU43 (2^{-dCt}). Box plots indicate the median value (horizontal line) and the 25th–75th percentile range (box) while whiskers showing the maximum and minimum values. Values outside this range are shown as outliers (open circles). *P*-values were determined by the Mann-Whitney U test. In every case, miRNAs downregulated in 13q-H CLL patients relative to 13q-L patients were also found to be downregulated by quantitative RT-PCR. Similar observations were made for miR-155, which was upregulated in 13q-H patients. All comparisons were statistically significant (P<0.05).

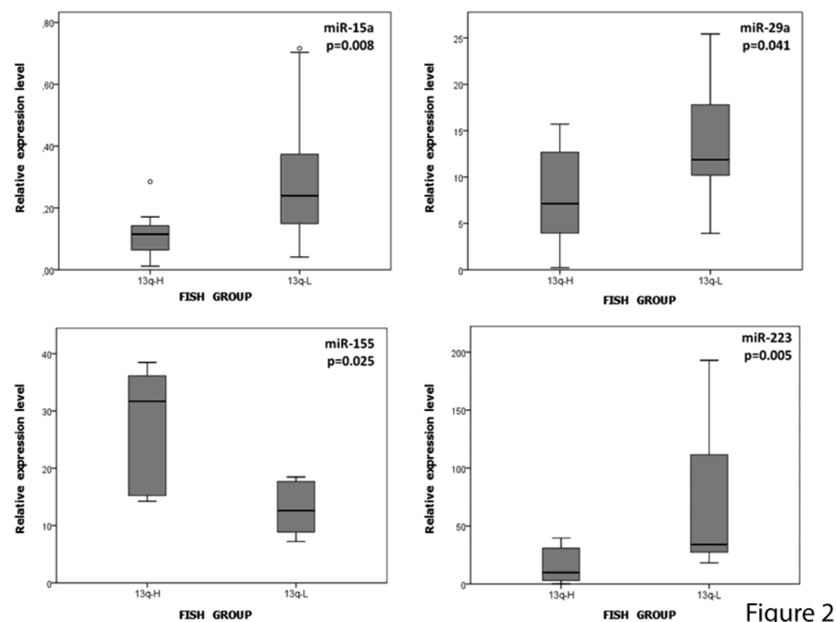


Figure 3.

Most significant cellular functions affected by the deregulation of miRNAs in 13q-H CLL patients.

432 out of the 1027 predicted mRNA target genes of the deregulated miRNAs in 13q-H CLL patients appeared also deregulated in our analysis. A functional enrichement analysis was

performed in this dataset. Category names are presented on the vertical axis. Of note, B cell receptor signaling and NF-kB signaling were among the most significant cellular functions affected. The significance of the association between the dataset and the canonical pathway was measured in two ways: (1) the ratio of the number of genes from the dataset that met the expression value cut-off that map onto the pathway divided by the total number of molecules that exist in the canonical pathway, represented by grey squares in the graph and (2) the *P*-value determining the probability of the association between the genes in the dataset and the canonical pathway, calculated by Fisher's exact test. The horizontal axis on the top indicates the −log (*P* value) and the horizontal axis at the bottom, the ratio. In both cases, the higher value indicates the higher significance.

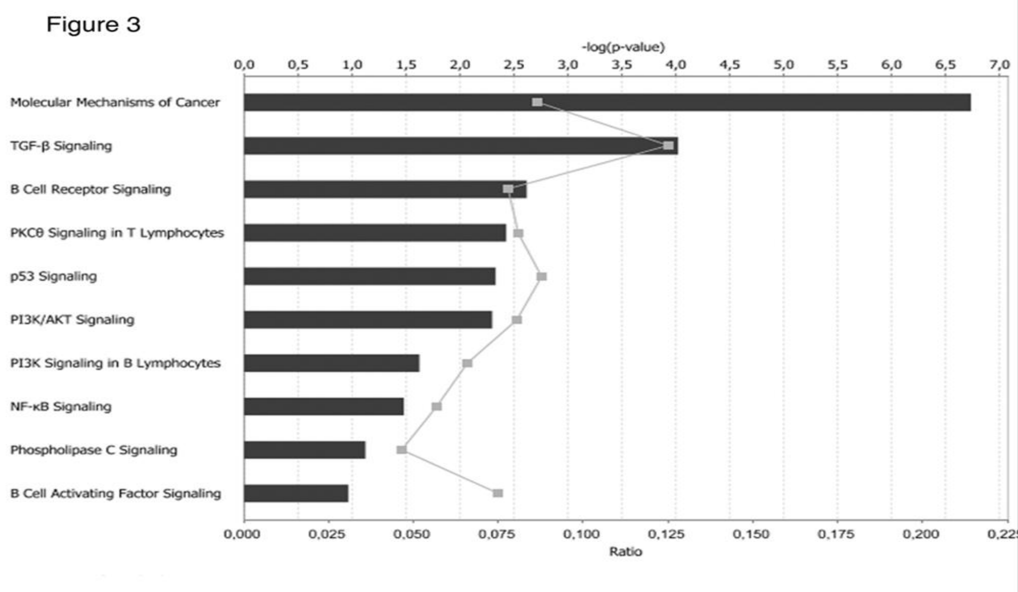
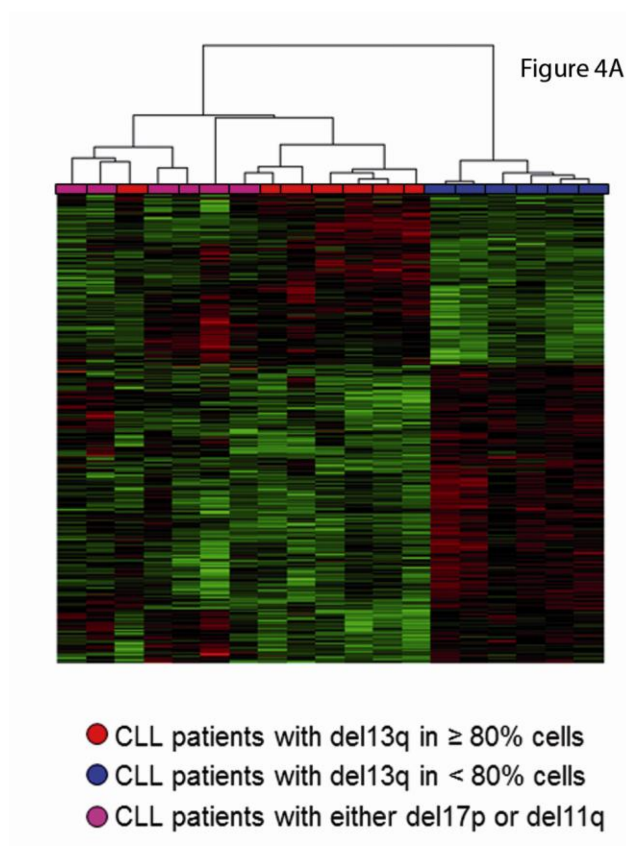


Figure 4.

Differential expression analysis followed by pathway analysis revealed commonly deregulated biological processes in CLL patients with a high load of 13q- cells, 17p- and 11q-.

A. **Heatmap of 3450 differentially expressed genes in CLL patients with a high number of losses in 13q (red), losses in 17p or 11q (magenta) and a low**

number of losses in 13q (blue). Differentially regulated genes were identified using Significance Analysis of Microarray (SAM), with a false discovery rate 5%, followed by the Global Test algorithm to test the candidate genes associated with the group of patients with a high number of losses. Individual patients are arranged in columns with the expression level for each gene across rows. Normalized gene expression values are color-coded (standard deviation from mean): red and green indicate high and low expression, respectively. All patients with 13q-L were clustered on the right side of the map in a homogeneous manner and separately from 13q-H and 17p-/11q-, which clustered together, showing that the gene expression profile (GEP) of CLL cases with higher percentages of 13q- cells is similar to that of 17p- and 11q-, while CLL patients with lower percentages of 13q- cells had a different gene profile.



B. Commonly deregulated biological functions in 13q-H and 17p-/11q- CLL patients compared with 13q-L CLL subgroup. Biological function names are presented on the vertical axis and the number of deregulated genes involved in each function, in the horizontal one . Fisher's exact test was used to examine the probability of the association between the genes in the dataset and the functional category. The color-coded bar plot (dark grey, light grey and black bars) depicts the analysis results. 13q-H patients showed marked differences in the expression of genes related to several cellular functions compared with 13q-L CLL patients (comparison 1, dark grey bars). In addition, most of these cellular functions were also deregulated in comparison with high-risk cytogenetic subgroups (17p- and 11q-) and 13q-L CLL patients (comparison 2, light grey bars). Thus, 13q-H, 17p- and 11q- patients share the deregulation of several important functions relative to 13q-L patients. Furthermore, a small number of genes related to cell cycle, cell growth and DNA repair (comparison 3, black bars) were found to be differentially expressed in the 13q-H group in a comparison of this subgroup of patients and high-risk cytogenetic subgroups.

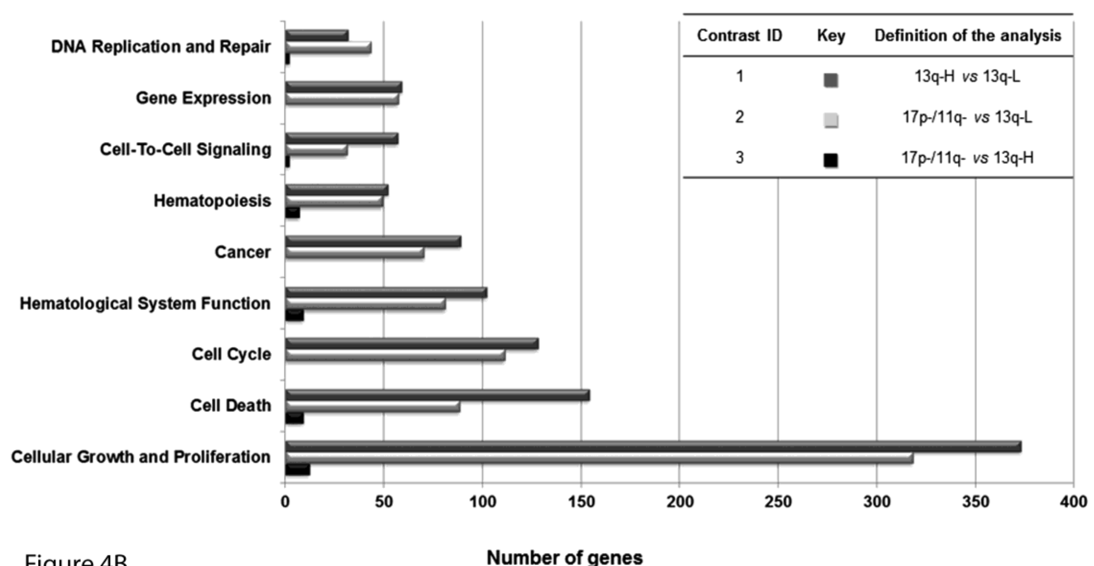
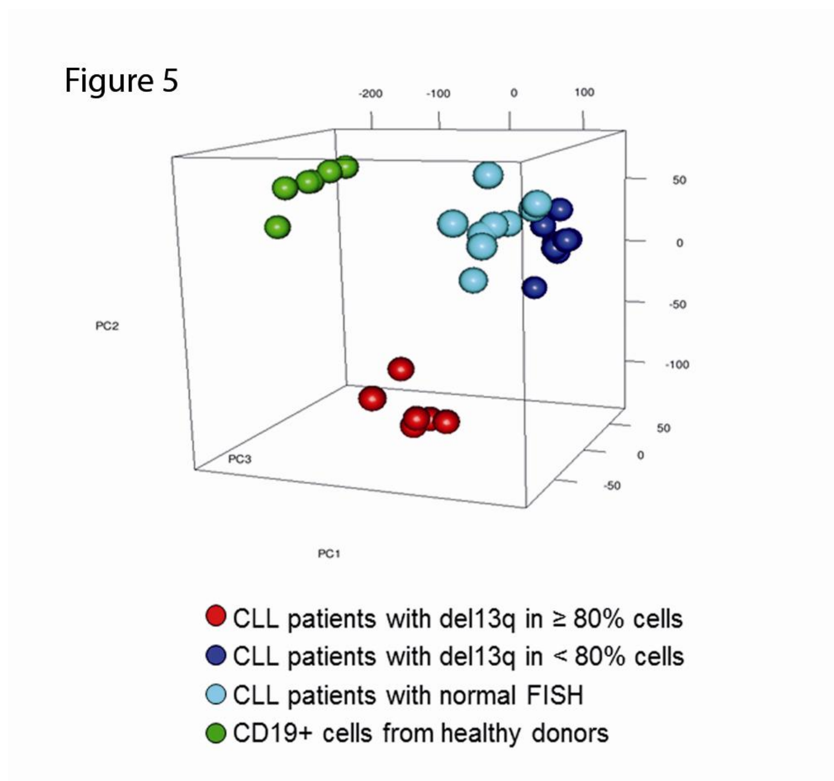


Figure 4B

Figure 5.

CLL patients with a high number of 13q- cells can be differenciated based on their expression profile.

Principal component analysis (PCA) plot of CD19+ cells from healthy controls (green), CLL with normal FISH (sky blue), 13q-H CLL (red) and 13q-L CLL (dark blue) was carried out using the 28,806 remaining genes after filtering the normalized gene expression matrices to remove the least variable genes (25%). Each sphere represents a single GEP. The result of the PCA shows a cumulative variance of 48.3%, 60.9% and 68.3% corresponding to one, two or three of the initial components, respectively. The expression pattern of CD19+ cells from CLL patients is notably different from the gene expression profile of CD19+ cells from healthy donors. Of note, the PCA analysis shows that 13q-H CLL patients have a distinctive gene expression profile. By contrast, the gene expression of B lymphocytes from 13q-L CLL and nCLL was similar.



TABLES

Table 1.

miRNAs significantly deregulated between 13q- CLL subgroups (patients with 80% or more of cells with 13q deletion and patients with less than 80% 13q cells). Upregulation or downregulation refers to 13q-H relative to13q-L CLL patients.

miRNA: microRNA

*deregulation shared with 17p/11q CLL patients.

miRNA	Map	q-value	R fold
Down-regulated			
hsa-mir-1-1*	20q13.33	0.0125	0.7027
hsa-mir-7-1	9q21.32	0.0397	0.5453
hsa-mir-15a	13q14.3	0.0329	0.4917
hsa-mir-29a	7q32.3	0.0354	0.5101
hsa-mir-34a*	1p36.23	0.0366	0.6874
hsa-mir-106b*	7q22.1	0.0280	0.5190
hsa-mir-181b	1q31.3	0.0256	0.6775
hsa-mir-204	9q21.11	0.0294	0.5693
hsa-mir-206	6p12.2	0.0476	0.7077
hsa-mir-221*	Xp11.3	0.0133	0.4622
hsa-mir-223*	Xq12	0.0017	0.1016
Up-regulated			
hsa-mir-134	14q32.31	0.0095	1.8096
hsa-mir-105-2	Xq28	0.0182	1.4040
hsa-mir-155*	21q21.3	0.0046	3.7013
hsa-mir-205	1q32.2	0.0161	1.3830

Table 2.

Enriched functional analysis of the 3450 genes differentially expressed between the two 13q- patient subgroups: 1244 genes were upregulated (i) and 2206 genes were downregulated (ii) in CLL patients with ≥80% cells displaying 13q deletion.

i.	Ingenuity Canonical Pathway	p-value	Up-regulated genes
	EIF2 Signaling	1,70E-07	RPL24,RPL27A,RPL26,RPS11,RPS27,RPS3A,SOS1,RPL35,RPL19,RPL13,RPL39L,RPL34,RPL27,RPL21,RPS19,RPL23A,RPS29,RPL36,RRAS2,RPS13,RPL26L1,RPL32,RPS25,RPS15A,RPL13A,RPS27A,RPL41,RPS14,RPSA
	B Cell Receptor Signaling	1,95E-05	MAP2K6,BLNK,MAP3K14,MAP3K9,CD19,CD79B,BAD,POU2F2,IKBKE,NFATC1,FCGR2B,PTEN,MAP3K12,RRAS2,CAMK2D,SYK,SOS1,CD22,NFATC2,PIK3AP1,PPP3CA,PRKCB
	PI3K Signaling in B Lymphocytes	6,92E-05	BLNK,CD19,CD79B,IKBKE,NFATC1,FCGR2B,PRKCZ,PTEN,BLK,CAMK2D,RRAS2,CD180,SYK,IRS1,SH2B2,NFATC2,PIK3AP1,PPP3CA,PRKCB
	CD27 Signaling in Lymphocytes	2,75E-03	MAP2K6,MAP3K12,MAP3K9,MAP3K14,CD70,IKBKE,TRAF5,CD27,MAP2K5
	mTOR Signaling	5,37E-03	VEGFB,RHOC,RPS19,RPS11,PRKCZ,RPS29,RPS27,RPS3A,RRAS2,RPS13,IRS1,GPLD1,RPS15A,RPS25,GNB1L,RPS27A,RPS14,RPSA,PRKCB
	Role of JAK1 and JAK3 in γc Cytokine Signaling	6,76E-03	BLNK,IL2RG,RRAS2,IRS1,SYK,SH2B2,JAK2,STAT1,IL7
	Nucleotide Excision Repair Pathway	1,20E-02	ERCC4,ERCC1,GTF2H1,ERCC2,MNAT1,XPA
	Regulation of eIF4 and p70S6K Signaling	1,66E-02	RPS19,RPS11,PRKCZ,RPS29,RPS27,RRAS2,RPS3A,RPS13,IRS1,SOS1,RPS25,RPS15A,RPS27A,RPS14,RPSA
	Phospholipase C Signaling	1,86E-02	BLNK,PEBP1,CD79B,RHOC,MEF2A,HDAC9,NFATC1,FCGR2B,MYL6B,PRKCZ,RRAS2,SYK,SOS1,GPLD1,NFATC2,MEF2C,GNB1L,ARHGEF9,PPP3CA,PRKCB
	PKCθ Signaling in T Lymphocytes	1,86E-02	MAP3K12,MAP3K9,MAP3K14,POU2F1,RRAS2,CAMK2D,SOS1,NFATC2,IKBKE,NFATC1,CARD11,PPP3CA
	April Mediated Signaling	2,34E-02	MAP3K14,NFATC2,IKBKE,NFATC1,TRAF5,TNFRSF17
	Interferon Signaling	2,63E-02	OAS1,IFI35,JAK2,STAT1,BCL2
	IL-4 Signaling	2,69E-02	IL2RG,RRAS2,IRS1,SOS1,NFATC2,NFATC1,JAK2,FCER2
	B Cell Activating Factor Signaling	2,95E-02	MAP3K14,NFATC2,IKBKE,NFATC1,TRAF5,TNFRSF17
	NF-κB Signaling	6,46E-02	MAP2K6,MAP3K14,FLT1,BMPR2,PRKCZ,TNFRSF17,TLR10,RRAS2,BMPR1A,TLR6,TLR7,TRAF5,CARD11,PRKCB

ii.	Ingenuity Canonical Pathway	p-value	Down-regulated genes
	Mitotic Roles of Polo-Like Kinase	1,35E-05	KIF23,CDC25C,ESPL1,CDC20,PPP2CA,PRC1,CDC7 (includes EG:12545),CCNB2,CDC23,PLK1,PPP2R5A,CDK1,CCNB1,SLK,HSP90B1,PLK4,PKMYT1,PPP2R1B,KIF11,CDC27,CDC25A
	Cell Cycle Control of Chromosomal Replication	2,75E-05	MCM6,CDC45,CDT1,CDC6,CDC7 (includes EG:12545),CDK6,ORC6,MCM4,MCM3,MCM2,CDK2,MCM7,ORC1
	Caveolar-mediated Endocytosis Signaling	6,17E-04	FYN,ITGA2B,ITSN1,RAB5A,ACTB,COPA,ITGA6,ITGA5,COPB1,ACTG1,COPG,COPB2,DYRK3,ITGB2,ITGAE,ITGAM,ITGA9,ITGAV,HLA-C,ITGA4
	Glycolysis/Gluconeogenesis	7,94E-04	PGK1,ALDH4A1,PGM2,PKLR,GAPDH,PGM1,BPGM,PDHA1,HK1,ALDH2,GPI,HK2,ALDH1A1,DHRS9,ENO1,DLAT,DLD,FBP1,ALDH3B1,LDHA,ACSL1
	Integrin Signaling	1,12E-04	RAP2B,RAF1,FYN,ITGA2B,ARHGAP26,TSPAN7,PIK3R1,PIK3R5,PPP1CB,NCK1,SHC1,ITGAE,PARVB,ARF6,WASL,RHOG,ITGA9,ARF4,PIK3CG,RHOU,ITGAV,VCL,MAP2K1,ACTN1,ITGA4,PXN,NRAS,ASAP1,CRKL,ACTB,ITGA6,TSPAN2,ITGA5,ACTG1,ITGB2,ARF1,ITGAM,TLN2,ZYX,PIK3CB,ACTN4,CTTN
	Cyclins and Cell Cycle Regulation	1,17E-03	RAF1,E2F4,CCNE2,TFDP1,HDAC2,PPP2CA,SUV39H1,CDK6,CDKN2C,CCNB2,E2F3,PPP2R5A,CDK1,CCNB1,CCNA2,CCNE1,E2F1,PPP2R1B,E2F2,CDK2,CDC25A
	Role of CHK Proteins in Cell Cycle Checkpoint Control	1,32E-03	CDC25C,E2F4,E2F1,RFC2,E2F3,BRCA1,CDK1,E2F2,CDK2,CDC25A,CHEK1,RFC3
	Nicotinate and Nicotinamide Metabolism	1,66E-03	DAPK1,PRKCQ,SGK1,MAPK6,CSNK1A1,CDK6,CSNK1D,PLK1,TTK,CDK1,SACM1L,VNN1,NEK2,ARAF,GRK6,PRKAA1,PNP,CD38,HIPK1,MAP2K1,NMNAT3,CDK2,BST1,DUSP16
	Inositol Phosphate Metabolism	1,78E-03	MINPP1,SGK1,PIK3R1,PIK3R5,CSNK1A1,TTK,OCRL,NEK2,PIK3CG,PRKAA1,PLCB1,IMPA2,PI4K2B,HIPK1,MAP2K1,PMPCA,MTMR3,DAPK1,IMPA1,PRKCQ,MTMR14,MAPK6,CDK6,CSNK1D,PLK1,CDK1,ARAF,SYNJ1,GRK6,PIK3CB,CDK2
	Cell Cycle Regulation by BTG Family Proteins	4,57E-03	CCNE2,E2F4,CCNE1,PPP2CA,E2F1,E2F3,PPP2R1B,CCRN4L,E2F2,CDK2,PPP2R5A
	Clathrin-mediated Endocytosis Signaling	6,61E-03	AP2A1,STON2,PIK3R1,PIK3R5,PDGFC,VEGFA,ARF6,ARRB1,WASL,SNX9,PIK3CG,DAB2,CSNK2B,AAK1,AP2M1,RAB5A,ACTB,CHP,CLTC,RAB7A,ITGA5,ACTG1,TSG101,ITGB2,ARRB2,LDLR,SYNJ1,TFRC,PIK3CB,DNM1L,CTTN
	Sphingolipid Metabolism	6,76E-03	LASS6,GLA,GALC,SGMS2,ASAH1,SACM1L,LASS2,VNN1,LPIN1,GBA,SMPD4,GLB1,PPAP2B,SPHK1,ARSB,FUT4,KDSR,DUSP16
	Role of BRCA1 in DNA Damage Response	7,08E-03	E2F4,BARD1,RBBP8,PLK1,E2F3,CHEK1,RAD51,GADD45A,E2F1,RFC2,BRIP1,BRCA1,HLTF,E2F2,RFC3
	Protein Ubiquitination Pathway	7,94E-03	USP24,USP14,USP12,UBE2H,PSMD7,CDC20,USP20,DNAJC3,CDC23,HSPA5,USP39,SMURF1,USP3,HSP90B1,USP42,USP47,NEDD4L,BRCA1,PSMC2,HLA-C,DNAJB12,USP15,MED20,USP36,USP38,HSPA9,USP19,PSMD6,PSMD5,HSPD1,PSMD3,USP1,UBE2D1,NEDD4,TRAF6,PSMD11,DNAJC5,USP4,PSMD2,DNAJB11,PSMD12

Table 3.

Most significant target genes affected by deregulation in miRNA in 13q-H CLL patients.

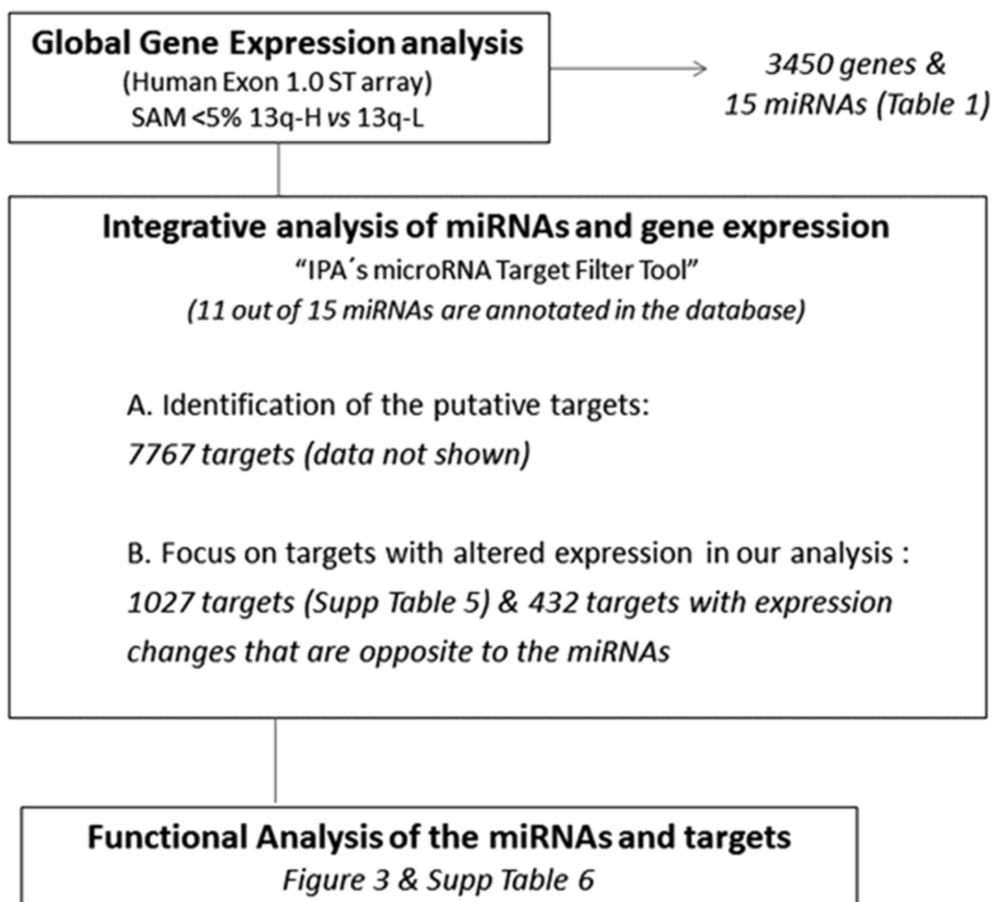
Target			miRNA	
Symbol	**Fold Change**	**B-cells related pathways**	**ID**	**Fold Change**
BCL2	2.132	Apoptosis	hsa-mir-206	0.708
			hsa-mir-15a	0.492
			hsa-mir-106b	0.519
			hsa-mir-204	0.569
			hsa-mir-34a	0.687
E2F5	2.624	DNA Damage Response	hsa-mir-206	0.708
			hsa-mir-106b	0.519
			hsa-mir-34a	0.687
FOS	0.447	B Cell Activating Factor,CD27	hsa-mir-155	3.701
LEF1	2.835	ILK, Wnt	hsa-mir-34a	0.687
MAP2K6	3.558	BCR,CD27	hsa-mir-29a	0.510
MAP3K12	1.254	BCR,CD27	hsa-mir-106b	0.519
MAP3K14	1.348	Apoptosis,B Cell Activating Factor,BCR,CD27	hsa-mir-106b	0.519
MAP3K9	1.400	BCR,CD27	hsa-mir-106b	0.519
MYD88	0.752	NF-κB,Toll-like Receptor	hsa-mir-155	3.701
PLCB1	0.773	PI3K	hsa-mir-205	1.383
RRAS2	1.931	Apoptosis, BCR	hsa-mir-223	0.102
SOS1	2.352	BCR	hsa-mir-15a	0.492
			hsa-mir-106b	0.519
			hsa-mir-204	0.569
TCL1A	7.848	Akt	hsa-mir-29a	0.510

SUPPLEMENTARY FIGURES

Supplementary Figure S1.

Summary of the miRNA analysis performed in the study.

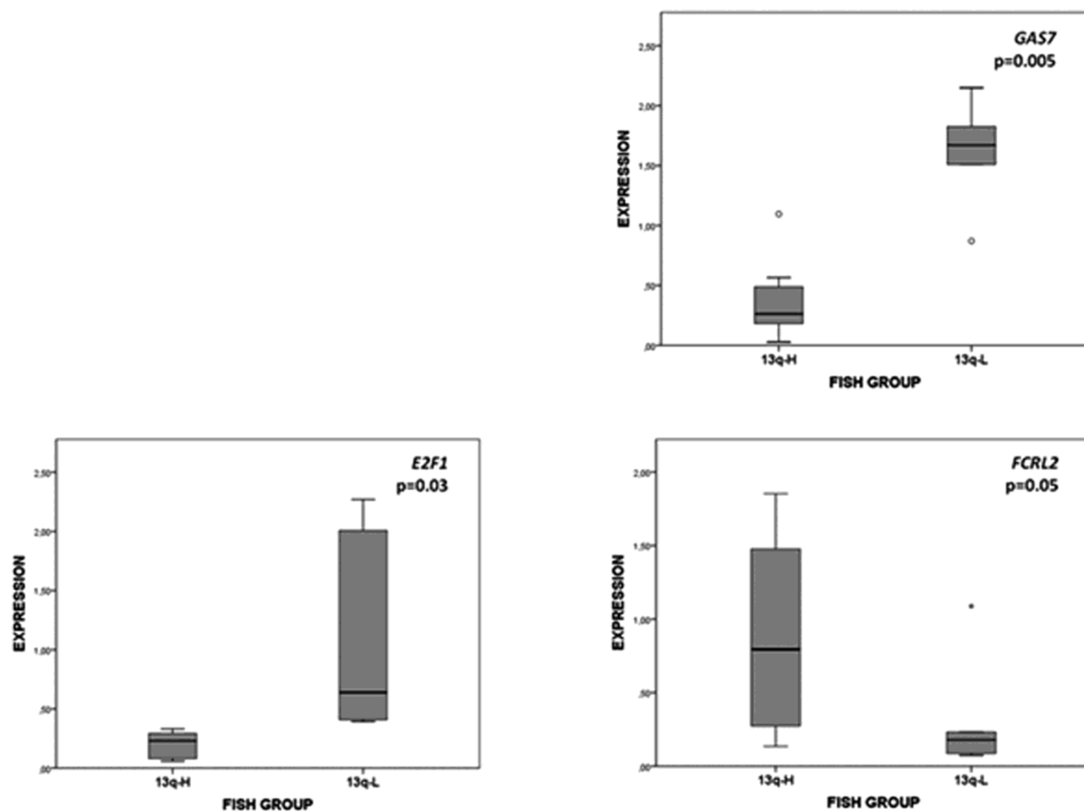
The chart explains the steps involved in the identification and validation of the miRNAs and their deregulated targets in 13q- CLL patients.



Supplementary Figure S2.

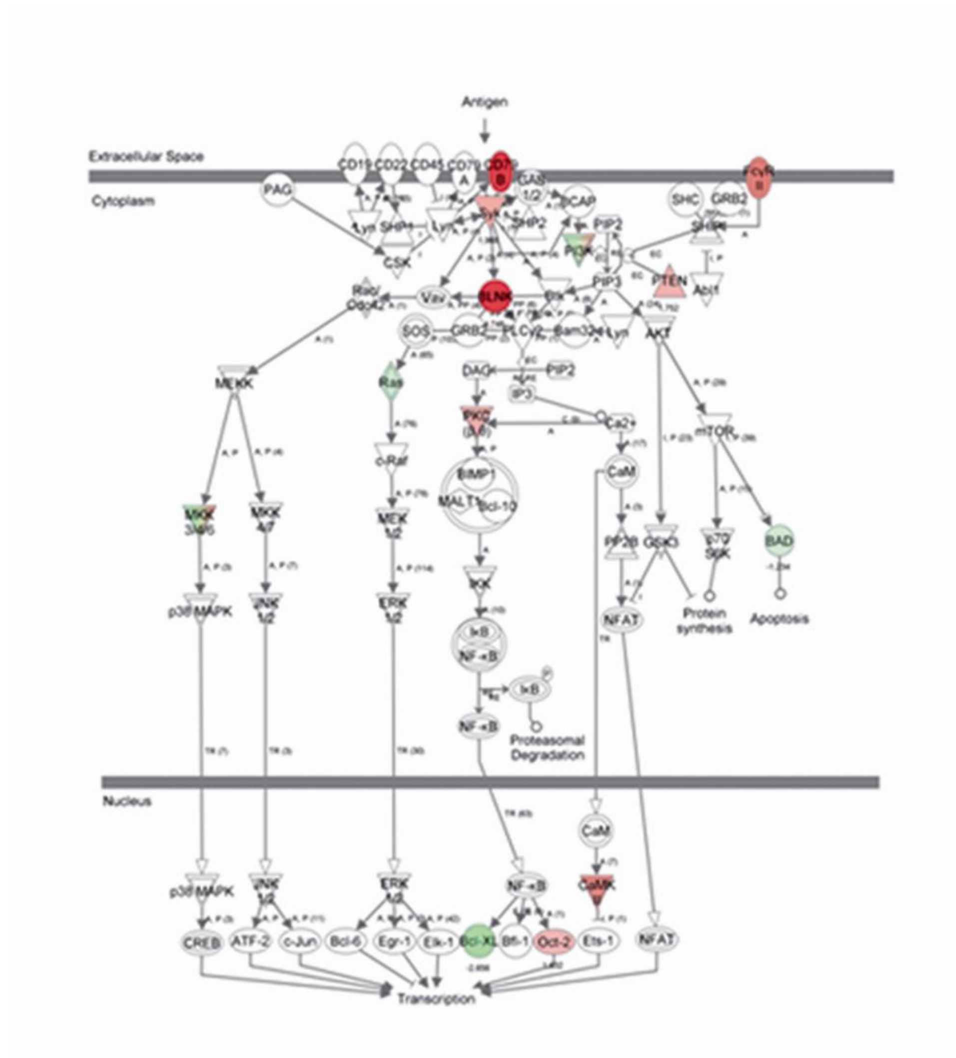
Box plot of the expression levels of three genes with significant differences between 13q-H and 13q-L patients, assessed by semi-quantitative PCR.

Box plots show the values for *GAS7*, *E2F1* and *FCRL2* expression, showing a significant difference in the level of expression between 13q-H and 13q-L CLL patients. The thick line inside the box plot indicates median expression levels, the limits of the box represent the 25th and 75th percentiles, and the whiskers show the maximum and minimum values. Outliers (extreme values falling outside the main distribution) are represented by open circles. Statistical significance was determined using the Mann-Whitney U test (*P*<0.05).



Supplementary Figure S3.

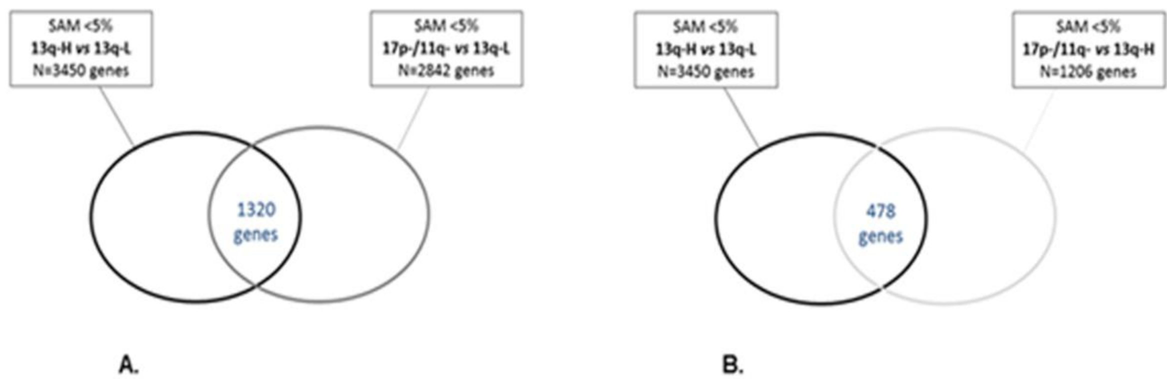
BCR signaling pathway identified as the top canonical pathway altered in CLL patients with higher percentages of 13q- losses according to the Ingenuity Pathway Analysis knowledge base. Genes significantly differentially expressed between CLL with 80% or more of cells with loss of 13q (13q-H) and CLL with losses in 13q in fewer than 80% of cells (13q-L) were mapped to the pathway and colored in red if the expression levels were higher, or in green if they were lower in 13q-H than in 13q-L cases. Significant positions of the pathway are occupied by genes deregulated in our analysis, indicating that this pathway is affected in 13q-H patients. CLL patients with 17p and 11q deletions showed similar deregulation in this pathway.



Supplementary Figure S4.

Overlap of differentially expressed genes as analyzed by SAM.

Venn diagram illustrating the number of significantly affected genes in common and distinct for the contrasts (1) and (2). 13q-H and 17p-/11q- shared the deregulation of 46% of genes (n=1325) relative to 13q-L.



Paper IV

Rodríguez AE, Quwaider D, Benito R, Misiewicz-Krzeminska I, Hernández-Sánchez M, Rojas-Ricardo E, Sarasquete ME, Gutiérrez NC, Hernández-Rivas JM. **A common polymorphism at miRNA-223 binding site in 3´untranslated region deregulates HSP90B1 expression in chronic lymphocytic leukemia.** Manuscript in preparation.

A COMMON POLYMORPHISM AT miRNA-223 BINDING SITE IN 3´UNTRANSLATED REGION DEREGULATES HSP90B1 EXPRESSION IN CHRONIC LYMPHOCYTIC LEUKEMIA

Running title: HSP90B1 overexpression in CLL

Ana-Eugenia Rodríguez[1], Dalia Quwaider[1], Rocío Benito[1], Irena Misiewicz-Krzeminska[1], María Hernández[1], Mª Eugenia Sarasquete[2], José Ángel Hernández[3], Norma C Gutiérrez[2], Jesús-María Hernández-Rivas[1,2]

[1]IBSAL, IBMCC, Centro de Investigación del Cáncer, Universidad de Salamanca-CSIC; [2]Hospital Clínico Universitario de Salamanca; [3]Hospital Infanta Leonor, Madrid.

Correspondence:

Jesús María Hernández-Rivas

Servicio de Hematología y Departamento de Medicina

Hospital Universitario de Salamanca

Paseo San Vicente 58

37007 Salamanca, Spain

Fax: +34923294624

e-mail: jmhr@usal.es

Keywords: Chronic lymphocytic leukemia, Next-generation sequencing, miRNA, polymorphisms.

Introduction

Chronic lymphocytic leukemia (CLL) has been established as a heterogeneous disease of remarkable diversity. *IGVH* mutational status and genomic aberrations subdivide CLL into distinct clinical subgroups,[1-3] although no single genetic variation or abnormality responsible for CLL development has been identified. A number of single nucleotide polymorphisms (SNP) have been implicated to impact upon the genetic susceptibility and the disease course in CLL.[4-9]

MicroRNAs (miRNAs) have recently emerged as a major class of gene expression regulators. miRNAs are single-stranded RNA molecules approximately 19 to 24 nucleotides in length and post-transcriptionally regulate gene expression by forming base pairing with sequences in the 3´-untranslated region (3´UTR) of target mRNAs. There is an increasing evidence suggesting that SNPs in the 3´UTR targeted by miRNAs (known as miRSNPs) are associated with diseases by affecting gene expression.[10-13] Moreover, deregulation of the normal microRNA expression profile could play a critical role in human disease, particularly in solid tumors,[14] but also in hematologic malignancies,[15] and aberrant expression of microRNAs has been recently associated with chronic lymphocytic leukemia outcome.[16,17] Moreover, the down-regulation of miR-223 has been demonstrated to be associated with disease aggressiveness and poor prognostic factors in CLL patients.[18] Thus, it may become a new reliable prognostic predictor. However, there is no evidence of the pathogenetic mechanism of this microRNA in CLL patients, and no target has been proposed or validated for miR-223 in CLL until date.

Over the last decade, several studies have implicated Heat shock proteins (HSPs) as major contributors to cancer progression and the development of chemo-resistance. In tumor cells, including CLL cells, HSPs are upregulated and may contribute to prolonged tumor cell survival via several mechanisms that remain to be fully revealed.[19-21] Hsp90 is an important mediator of VEGF-induced cell survival, a mechanism that involves up-regulation and binding of Bcl-2 and APAF-1 to Hsp90.[22]

The protective effect of HSP90 is also related to its ability to interfere with apoptotic pathways.[23,24] Preclinical studies in CLL have shown that HSP90 inhibition causes the degradation of ZAP-70 and other proteins associated with poor survival and this may ultimately lead to apoptosis.[23,25-28] Furthermore, inhibitors of Hsp90 have been proposed as a novel therapeutic option for CLL, particularly as their mechanism of action appears to be independent of mutations of *ATM* or *TP53*. This could represent a therapeutic option to drug resistance in CLL associated with lesions in the ATM/TP53 pathway.[29-31] Thus, targeting Hsp90 is an attractive strategy in CLL.

The aim of the present study was to identify the role of Hsp90 in CLL. By applying next generation sequencing (NGS) techniques we have detected a common 4 bp deletion SNP (rs2307842) in 25% of CLL patients, which disrupts the binding site for miR-223 in HSP90B1 3´UTR. We demonstrated that this SNP alters the regulation of HSP90B1 expression in CLL, leading to its overexpression only in B lymphocytes.

Material and Methods

Patients and controls

A total of four patients with CLL and four patients with other hematological malignancies (used as controls) were selected for a Targeted sequence capture and DNA Sequencing assay. CLL diagnosis was performed according to World Health Organization (WHO) classification[32] and Working Group of National Cancer Institute (NCI) criteria. The CLL patients were three males and one female with age at diagnosis ranging from 41 to 66 years. Two patients had *IGHV*-unmutated gene, and two *IGHV*-mutated gene (<98% homology). Two patients had 13q deletion, one had 11q deletion and one did not show any cytogenetic aberration by FISH. Tumor samples (CD19+ fraction) were used for sequencing and were obtained before administration of any treatment. To determine the clinical impact of HSP90B1 3´UTR polymorphism, we expanded the study to 109 additional patients with CLL and 32 healthy controls. The clinical and biological characteristics of the CLL patients are summarized in Supplementary Table 1. FISH studies and *IGVH* mutational status were assessed in all patients. The median age was XX (range, 35 to 90 years). Most patients were male (66%) and were in Binet clinical stage A (69%), while 26% were in stage B and the remaining 5% in stage C. The study was approved by the local ethical committees. Informed consent was obtained from each patient before entering the study.

Cells and culture conditions

The human multiple myeloma cell lines NCI-H929 and MM1S were acquired from ATCC (American Type Culture Collection). Both cell lines were cultured in RPMI 1640 medium supplemented with 20% of fetal bovine serum and antibiotics (Gibco). Cells were routinely checked for the presence of mycoplasma with MycoAlert kit (Lonza GmBH) and only mycoplasma free cells were used in the experiments. Phenotypic and

cytogenetic identity of the cell lines were verified by flow cytometry and FISH before the experiments.

Collection and preparation of samples

Peripheral blood mononuclear cells (PBMCs) from all CLL patients were isolated by Ficoll–Hypaque gradient centrifugation (Amersham Biosciences, Pittsburgh, PA), snap-frozen and stored at -80°C. CLL B-lymphocytes were purified using magnetically activated cell sorting (MACS) CD19 MicroBeads (Miltenyi Biotec, Bergisch Gladbach, Germany). CD19 selection resulted in >98% purity, as analyzed by flow cytometry.

DNA was extracted from fresh-frozen samples by using a Qiagen kit. To ensure a good quality, DNA was measured using NanoDrop ND-1000 (ND-1000; NanoDrop Technologies, Wilmington, DE). Onli samples with OD 260nm/280 nm >1.8 were included. The integrity of the DNA was visually inspected on a 1% agarose gel. RNA was extracted using Trizol reagent (Invitrogen) according to the standard protocol. The RNA integrity was assessed using Agilent 2100 Bioanalyzer (Agilent Technologies).

Targeted Sequence Capture and DNA Sequencing assay

We applied array-based sequence capture (Roche NimbleGen) followed by next-generation sequencing (Roche GS FLX Titanium sequencing platform) to analyze 93 genes that we considered relevant in CLL (Supplementary Table S3) and two chromosomal regions: 13q14.3 (50043128-50382849 bp) and 17p13.1 (7500000-7535000). Details regarding the design of the array, 454 sequencing, coverage statistics and data analysis are also provided in the Supplementary Methods and Supplementary Tables S3 and S4.

Pyrosequencing assays

The following oligonucleotide primers were used for amplifying a 170-bp genomic fragment spanning the sequence for 3UTR region of the HSP90B1 gene: HSP90B1-For_Bio: 5-CTGCACTGTAAAATGTGGGATTAT-3 and HSP90B1-Rev: 5-

AGACACTGAGTATTTGGGATCTTT-3. The 5-ends of the forward primer were conjugated with biotin (Bio). PCR was performed using a DNA thermal cycler (Applied Biosystem, ABI 9700). The template was denatured initially for 5 minutes at 96°C followed by 40 amplification cycles containing: initial denaturation at 95°C for 30 seconds, followed by annealing for 30 seconds at 60°C and extension at 72°C for 30 seconds. Final extension was carried at 72°C for 5 minutes. The amplified PCR products were checked by electrophoresis on 1% agarose gels and stored at -20ºC.
Pyrosequencing was carried out on a PyroMark Q24 system (Qiagen) according to the manufacturer's protocol. The pyrosequencing primer HSP90B1-Py1-5-TGACAAGATTTTACATCA-3 was used with the nucleotide dispensation order: CAGAGTAGTCA. A total of 10 μl PCR product, 2 μl Streptavidin Sepharose High Performance beads (GE Healthcare Bio-Sciences AB, Sweden Uppsala), 28 μl water and 40 μl binding buffer (Qiagen) were mixed and agitated constantly for 10 min at 1,400 rpm. The PCR products attached to the beads were washed in 70% ethanol, followed by denaturation in 0.2 N NaOH and washing buffer (Qiagen). Purified DNA samples were annealed to the sequencing primer (0.3 μM) in 25 μl annealing buffer (Qiagen) and denaturated for 2 min at 80ºC, followed by cooling down to room temperature for 5 min. The samples were then processed in the PyroMark Q24 Instrument, with a running time of 15 min for 24 samples. Setup of assay and sequence-run as well as analysis were performed by the PyroMark Q24 Software.

Luciferase reporter assay

The double-stranded oligonucleotides corresponding to the wild-type (WT-3'UTR) or mutant (MUT-3'UTR) miR-223 binding site in the 3´UTR of HSP90B1 (NM_003299) were synthesized (Sigma-Aldrich) and ligated between the PmeI and XbaI restriction sites of the pmirGLO vector (Promega). The mutant (MUT-3'UTR) miR-223 binding site was generated based on 3´UTR HSP90B1 sequence in which 4 nucleotides were

deleted at miR-223 seeding region, corresponding to the SNP found in the Targeted Sequence Capture and DNA Sequencing assay (rs2307842). The oligonucleotides sequences are presented in Supplemental Table 4. For luciferase assays, HEK293 cells were transfected with 500 ng of the above constructs and cotransfected with 25 nM miRNA precursor molecule by nucleofection using HEK293 cell line program in the Amaxa II nucleofector system. At 24 hours after transfection cells were collected and Firefly and Renilla luciferase activities were measures using Dual-Glo® Luciferase Assay System (Promega) according to the manufacturer's protocol. Measurements were performed on Tekan Infinite® F500 microplate reader. Firefly luciferase activity was normalized to Renilla luciferase activity.

Transfection with synthetic miRNAs

H929 and MM1S cell lines were transfected with Pre-miR™ miRNA precursors pre-miR-223 or pre-miR™ miRNA negative, non-targeting control#1 (Ambion) at 50 nM concentration, using the nucleofector II system with C-16 program and Q-023 program, respectively (Amaxa). Transfection efficiency was assessed with Block-iT™ Fluorescent Oligo (Invitrogen) by flow cytometry.

Gene-specific semi-quantitative PCR

To detect the mRNA expression of both *HSP90B1* and *BCL2*, total RNA (1 µg) was reverse transcribed to cDNA using SuperScript™ III First-Strand Synthesis SuperMIx (Invitrogen). Semi-quantitative SYBRgreen PCR was done in triplicate with iQ™ SYBR® Green Supermix kit (BioRad) using the IQ5 Multicolor Real-Time PCR Detection System (Bio-Rad) with gene-specific primers (Supplementary Table S5). The *GAPDH* gene was used as the internal control and the expression data were analyzed by the comparative Ct method.

Immunoblotting

Whole cell lysates were collected using RIPA buffer (Sigma-Aldrich) containing protease inhibitors (Complete Protease Inhibitor Cocktail Tablets, Roche) and phosphatase inhibitors (Phosphatase Inhibitor Cocktail A and B, Santa Cruz Biotechnology). Protein concentration was measured using the Bradford assay (BioRad). Protein samples (40 μg/lane) were subjected to sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to polyvinylidene fluoride (PVDF) membrane 0.45 μm (BioRad). The primary antibodies used for immunoblotting were anti-Hsp90b1 (Cell Signalling) and anti-β-actin (Sigma-Aldrich) as an internal control for protein loading. The membranes were then washed and subsequently incubated with the secondary horseadish peroxidase-linked anti-mouse IgG and anti-rabbit IgG antibodies (PierceNet) (1:10000). Chemiluminescence was detected using Amersham ECL Plus™ Western Blotting Detection Reagent (GE Healthcare).

Clinical/Laboratory prognostic parameters

To assess the role of rs2307842 polymorphism in CLL patients, comparison was made between various clinical/laboratory parameters including expression of HSP90B1 mRNA when available, IGVH gene mutational status, Binet staging, age at presentation, lymphocyte doubling time (> or < 12 months), time to first treatment, progression-free survival and overall survival.

Statistical analysis

The two-sided Student's *t* test was used to analyse differences in experiments. Data are reported as mean values ± SD of at least triplicate determinations. To analyze the results from the semiquantitative PCR experiments with CLL patients, the Mann-Whitney *U* test was used to identify statistically significant differences. *P*

values < 0.05 were considered statistically significant. All statistical analyses were conducted using SPSS 19.0 statistical package (SPSS).

Results

A targeted genome capture and next-generation sequencing strategy allows the identification of a common polymorphism in 3´UTR *HSP90B1*

Using a custom NimbleGen array we captured and sequenced 93 genes and two entire chromosomal regions of eight individuals: four CLLs and four haematological malignancies not CLL. The enrichment assay followed by NGS allowed the detection of over 1 600 variations/sample (median 1 721, range 1 618-1 823). All putative variants were first compared with published single nucleotide polymorphism (SNP) data (dbSNP build 130; http://www.ncbi.nlm.nih.gov/projects/SNP). Most of the variants detected were identified as known SNPs and 226 variants were present in all the patients. Thus, they were discarded. Overall, 10% of variants detected in each sample were mutations not described previously. 73 missense variations affecting 33 genes were detected. Most of the genes had one (70%) or two (12%) variations. Results are summarized in Supplementary Table S6.

By applying a custom-made data analysis pipeline, we have annotated the detected variants, including known single-nucleotide polymorphisms (SNPs), amino acid consequences, genomic location and miRNA binding sites. Thus we have identified a polymorphism in 3´UTR HSP90B1 in one CLL patient (25%). This polymorphism is filled as rs2307842 (102865778-102865781) in the NCBI SNP database and the frequency is 21% in the rs2307842 results in the deletion of four nucleotides in 3´UTR sequence, three of them are part of the predicted binding site for miR-223 (Figure 1A). We hypothesized that this 'GACT' deletion disrupts the binding site for miR-223 thereby

increasing the translation of HSP90B1. Given the importance of miRNA regulation of gene expression in cancer and also in CLL, we decided to gain insight in this field.

HSP90B1 is a direct target gene of miR-223

We have confirmed that miR-223 regulates HSP90B1 expression by 3´UTR reporter assays. First, the double-stranded oligonucleotides, corresponding to the wild-type (WT-3'UTR) or mutant (MUT-3'UTR) miR-223 binding site in the 3`UTR of HSP90B1 (NM_003299), were synthesized and pmirGLO Vectors made up of an SV40 promoter, the Renilla luciferase gene, and the 3´UTR of HSP90B1 were transfected into HEK293 cells along with miR-223 or negative control (NC) mimics and relative luciferase activity was measured at 24h. The relative luciferase activity of the construct with wild-type 3´UTR was significantly repressed by 31% following miR-223 transfection ($p<0.05$)(Figure 1B). However, the presence of rs2307842 polymorphism in 3´UTR HSP90B1 abolished this suppression ($p<0.05$)(Figure 1B), suggesting that miR-223 directly binds to this site.

We also validated HSP90B1 as a target gene of miR-223 by transfecting MM1S and H929 cell lines with miR-223/NC mimics and then measuring HSP90B1 expression by semi-quantitative PCR and Western blot. Sanger sequencing showed that rs2307842 polymorphism was present in 3´UTR HSP90B1 of MM1S cell line. All the experiments were done in triplicate. Exogenous expression of miR-223 downregulated the expression levels of HSP90B1 in H929 cell line (3´UTR-WT) in both mRNA and protein levels (Figures 1C and 1D). By contrast, HSP90B1 expression was not modified in MM1S cell line (3´UTR-MUT) (Figures 1C and 1D).

Taken together, all these results demonstrate that HSP90B1 is a *bona fide* target gene of miR-223 and the rs2307842 polymorphism abolishes the miR-223 regulation on HSP90B1 expression.

rs2307842 is a common polymorphism in CLL patients and is not associated with any clinical feature of the disease

To determine the clinical impact of HSP90B1 3´UTR polymorphism in CLL, we screened 109 additional patients with CLL and 32 healthy controls for this SNP by pyrosequencing. A total of eighteenpaired DNA samples (CD19+ and CD19- cells) immunomagnetically purified from CLL patients showed complete concordance in their 3´UTR sequence, confirming that rs2307842 was the result of a single nucleotide polymorphism and not an acquired mutation. The SNP was found in 27/109 (25%) CLL patients and 8/32 (25%) healthy controls, which is consistent with the data obtained from dbSNP (http://www.ncbi.nlm.nih.gov/projects/SNP). Overall, we found no association between rs2307842 and any clinical characteristic of CLL patients.

rs2307842 polymorphism determines HSP90B1 overexpression in B-lymphocytes of CLL patients

To test the hypothesis that rs2307842 determines overexpression of HSP90B1 mRNA by abolishing the miR-223 regulation in CLL patients, we have performed semi-quantitative PCR on CD19+ cells from a subset of the patients previously characterized for the presence of the SNP. To gain insight into the influence of the SNP on gene expression, we measured HSP90B1 mRNA levels in paired samples (tumoral and normal) from CLL patients with rs2307842 (VAR, n=6) and wild-type (WT, n=12). PCR results showed that B-lymphocytes from VAR-CLLs have a higher expression of HSP90B1 than B-lymphocytes from WT-CLLs (P=0.002) and also from their normal counterpart (P=0.011) (Figure 2). Despite the presence of the SNP in the normal counterpart of the VAR-CLLs, no changes in mRNA expression were observed in comparison with the expression levels of WT samples (Figure 2)// when comparing with

WT samples. Interestingly, rs2307842 determined HSP90B1 overexpression only in the tumoral fraction of the CLL patients with the SNP (Figure 2).

In addition, taking into account that the down-regulation of miR-223 is associated with disease aggressiveness and poor prognostic factors in CLL, such as unmutated status of *IGVH,* we have investigated HSP90B1 expression in these patients (*IGVH*-CLLs). As shown in Supplementary Figure 1, B-lymphocytes from *IGVH*-CLL patients also showed overexpression of HSP90B1 in comparison with WT patients (P=0.003) and their normal counterpart (P=0.006).

CLL patients with HSP90B1 overexpression also have overexpression of BCL-2.

Considering the key role of the antiapoptotic Bcl-2 in CLL pathogenesis and the fact that Hsp-90 can modulate Bcl-2 expression, we also have measured the mRNA levels of BCL-2 in our series of patients (VAR-CLL, IGVH-UM-CLL and WT-CLL patients). As shown in Figure 3, the expression levels of HSP90B1 mRNA were positively correlated with the mRNA expression levels of BCL-2. Direct correlation was obtained using Spearman´s correlation, r= 0.517, P= 0.006.

Discussion

MicroRNAs are known to inhibit gene expression by binding to the 3´UTR of the target transcript. In this study, we present evidence that rs2307842, a common SNP in HSP90B1 3´UTR, modulates HSP90B1 expression by interfering with miR-223 function, resulting in HSP90B1 overexpression. Furthermore, we propose that this could represent a pathogenic mechanism for miR-223, as HSP90B1 overexpression has been involved in several types of cancer.

Functional polymorphisms in 3´UTRs of several genes (also known as miRSNPs or miR-polymorphisms) have been reported to be associated with diseases affecting gene expression. Loss of microRNA function due to a defective miRNA-mRNA binding results in overexpression of the target mRNA, which can be involved in key biological processes, oncogenic mechanisms or drug resistance.[10,11,13,33] To the best of our knowledge, only one SNP in the 3´UTR region of IRF4 had been previously associated with an increased risk of CLL,[5] but, as far as we know, no functional studies of miRSNP in CLL have been previously reported.. Firstly we have confirmed that miR-223 regulates HSP90B1 expression by 3´UTR reporter assays and forced overexpression of miR-223/transfection with synthetic miR-223. HSP90B1 was validated as miR-223 direct target, as miR-223 reduced mRNA and protein levels of HSP90B1 in 31%. Sincethe level of target gene expression is reduced but not abolished by the correspondent miRNA, it is considered a fine “tuning” miRNA.[34]

Moreover, our results showed that the presence of rs2307842 SNP alter the interaction between the target site in HSP90B1 and miR-223 in CLL patients. We have performed semi-quantitative PCR using CD19+ peripheral blood lymphocytes from CLL patients, both with the SNP and wild-type. As expected, B lymphocytes from CLL patients with the SNP showed higher levels of HSP90B1 in comparison with B lymphocytes from wild-type CLL patients. Surprisingly, non-clonal cells from CLL patients with the SNP

showed levels of mRNA HSP90B1 similar to that of wild-type CLL patients (both tumoral and normal counterparts). These findings suggest that a regulatory mechanism of HSP90B1 expression could be present in cells with rs2307842 polymorphism. This mechanism may be responsible for maintaining the optimal levels of mRNA HSP90B1 and could be damaged in CLL B lymphocytes. However,other possible hypothesis, such asis that CLL B lymphocytes accumulate so many genetic aberrations that HSP90B1 expression is also affected by several mechanisms, could not be excluded. Thus, to explain gene alteration in cancer by a microRNA-dependent mechanism, besides looking for mutations inside or surrounding microRNA genetic loci, it could bepossible also to search for mutations altering the 3´UTR-site targeted by the miRNA.

There is still little known about miR-223 function in CLL. MicroRNA-223 expression levels decreased significantly with the progression of the disease thus associating miR-223 down-regulation with higher tumor burden, disease aggressiveness, and poor prognostic factors.[16,35] Interestingly, miR-223 and miR-29c have been used to create a quantitative PCR-based score able to improve CLL patients stratification in terms of treatment free survival and overall survival, when combined with two other prognostic factors.[18] Despite the proven implication of miR-223 expression in CLL prognosis, little is known on the molecular mechanisms which could be responsible for the poor outcome of CLL patients showing miR-223 downregulation. Unlike other miRNAs with prognostic value in CLL such as miR-181b and miR-29c, the target of miR-223 in CLL is still unknown.[36,37] Thus, we also performed semi-quantitative PCR on purified cells from unmutated-IGVH (UM) CLL patients. HSP90B1 mRNA levels were upregulated in UM B cells in comparison with mutated and wild-type B cells. Moreover, the overexpression was also significant in comparison with the non clonal cells of UM CLL patients. Thus we hypothesized that the deregulation on HSP90B1 expression in B cells with the SNP could be similar to that of B cells showing downexpression of miR-223. Our work provides novel information about how the downregulation of miR-223

can be determining the poor outcome of CLL patients, maybe through upregulation of HSP90B1 expression.

Hsp90 is a molecular chaperone that catalyzes the conformational maturation of a number of oncogenic signalling proteins with the hydrolysis of ATP. Inhibitors such as 17-AAG and 17-DMAG prevent the binding of ATP to Hsp90, resulting in the release and degradation of signalling proteins required for the growth of cancer. Identification of a novel Hsp90 client that plays an important role in the course of a malignancy is of great interest for clinical development of Hsp90 inhibitors. One such client is the tyrosine kinase ζ-associated protein of 70 kD (ZAP-70). ZAP-70 is physically associated with Hsp90 in B-CLL cells and the protein expression is ablated by treatment with 17-AAG.[26] Recent evidence suggests that Hsp90 inhibitors could be a therapeutic option in CLL. These drugs have shown preclinical efficacy in the treatment of CLL independently of p53 function, indicating its value to a broad set of patients with limited therapeutic options.[29-31] Hsp90 inhibitors have a novel mechanism of action targeting multiple pathways that has not been fully elucidated in CLL and finally leads to apoptosis. Of note, a correlation between HSP90B1 and BCL2 overexpression in CLL patients was observed in the present study. The role of Hsp90 in Bcl-2 regulation has not been yet elucidated, except in mast cells.[38] In this sense, therapies such as 17-DMAG, which target both of these proteins, are of great clinical interest.[25]

In summary our study highlights the proven importance of miRNAs as critical players in the pathogenesis of CLL and shows for the first time that a single-nucleotide polymorphism in the miR-223 binding site modulates HSP90B1 expression in B lymphocytes of CLL. These results could provide a plausible explanation as to why CLL patients harboring miR-223 downregulation are associated with a poor outcome. This work also point out HSP90B1 as a new pathogenic mechanism in CLL, given its role in apoptosis resistance and cell survival promotion in several types of cancer and also in

CLL. Future work is needed to understand the relevance and functional consequence of this common polymorphism in CLL patients and future efforts are warranted to explore whether miR-223 and HSP90B1 could be potentially useful for CLL prognosis and treatment.

.

Acknowledgements

The study was partially supported by grants from the Spanish Fondo de Investigaciones Sanitarias 02/1041 and FIS 09/01543; Caja de Burgos-Banca Cívica, Proyectos de Investigación del SACYL 106/A/06 and by the Acción Transversal del Cáncer project, through an agreement between the Instituto de Salud Carlos III (ISCIII), the Spanish Ministry of Science and Innovation, the Cancer Research Foundation of Salamanca University and the Redes de Investigación RTIIC (FIS). AR was fully supported by an Ayuda Predoctoral FIS de Formación en Investigación by the Spanish Fondo de Investigaciones Sanitarias. We thank Irene Rodríguez, Sara González, Teresa Prieto, Mª Ángeles Ramos, Almudena Martín, Ana Díaz, Ana Simón, María del Pozo and Vanesa Gutiérrez of the Centro de Investigación del Cáncer, Salamanca, Spain, for their technical assistance and Jesús F. San Miguel for his critical review of the manuscript.

Conflict- of- interest disclosure

The authors declare no competing financial interests.

References

1. 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.

2. Damle RN, Wasil T, Fais F et al. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. Blood 1999;94:1840-1847.

3. Dohner H, Stilgenbauer S, Benner A et al. Genomic aberrations and survival in chronic lymphocytic leukemia. N Engl J Med 2000;343:1910-1916.

4. Sellick GS, Goldin LR, Wild RW et al. A high-density SNP genome-wide linkage search of 206 families identifies susceptibility loci for chronic lymphocytic leukemia. Blood 2007;110:3326-3333.

5. Di Bernardo MC, Crowther-Swanepoel D, Broderick P et al. A genome-wide association study identifies six susceptibility loci for chronic lymphocytic leukemia. Nat Genet 2008;40:1204-1210.

6. Sellick GS, Wade R, Richards S et al. Scan of 977 nonsynonymous SNPs in CLL4 trial patients for the identification of genetic variants influencing prognosis. Blood 2008;111:1625-1633.

7. Zenz T, Benner A, Duhrsen U et al. BCL2-938C>A polymorphism and disease progression in chronic lymphocytic leukemia. Leuk Lymphoma 2009;50:1837-1842.

8. Starczynski J, Pepper C, Pratt G et al. Common polymorphism G(-248)A in the promoter region of the bax gene results in significantly shorter survival in patients with chronic lymphocytic Leukemia once treatment is initiated. J Clin Oncol 2005;23:1514-1521.

9. Gryshchenko I, Hofbauer S, Stoecher M et al. MDM2 SNP309 is associated with poor outcome in B-cell chronic lymphocytic leukemia. J Clin Oncol 2008;26:2252-2257.

10. Mishra PJ, Humeniuk R, Mishra PJ et al. A miR-24 microRNA binding-site polymorphism in dihydrofolate reductase gene leads to methotrexate resistance. Proc Natl Acad Sci U S A 2007;104:13513-13518.

11. Bertino JR, Banerjee D, Mishra PJ. Pharmacogenomics of microRNA: a miRSNP towards individualized therapy. Pharmacogenomics 2007;8:1625-1627.

12. Gehring NH, Frede U, Neu-Yilik G et al. Increased efficiency of mRNA 3' end formation: a new genetic mechanism contributing to hereditary thrombophilia. Nat Genet 2001;28:389-392.

13. Mayr C, Hemann MT, Bartel DP. Disrupting the pairing between let-7 and Hmga2 enhances oncogenic transformation. Science 2007;315:1576-1579.

14. Volinia S, Calin GA, Liu CG et al. A microRNA expression signature of human solid tumors defines cancer gene targets. Proc Natl Acad Sci U S A 2006;103:2257-2261.

15. Fabbri M, Garzon R, Andreeff M et al. MicroRNAs and noncoding RNAs in hematological malignancies: molecular, clinical and therapeutic implications. Leukemia 2008;22:1095-1105.

16. Calin GA, Ferracin M, Cimmino A et al. A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. N Engl J Med 2005;353:1793-1801.

17. Fulci V, Chiaretti S, Goldoni M et al. Quantitative technologies establish a novel microRNA profile of chronic lymphocytic leukemia. Blood 2007;109:4944-4951.

18. Stamatopoulos B, Meuleman N, Haibe-Kains B et al. microRNA-29c and microRNA-223 down-regulation has in vivo significance in chronic lymphocytic leukemia and improves disease risk stratification. Blood 2009;113:5237-5245.

19. Nylandsted J, Brand K, Jaattela M. Heat shock protein 70 is required for the survival of cancer cells. Ann N Y Acad Sci 2000;926:122-125.

20. Broemer M, Krappmann D, Scheidereit C. Requirement of Hsp90 activity for IkappaB kinase (IKK) biosynthesis and for constitutive and inducible IKK and NF-kappaB activation. Oncogene 2004;23:5378-5386.

21. Sato S, Fujita N, Tsuruo T. Modulation of Akt kinase activity by binding to Hsp90. Proc Natl Acad Sci U S A 2000;97:10832-10837.

22. Dias S, Shmelkov SV, Lam G, Rafii S. VEGF(165) promotes survival of leukemic cells by Hsp90-mediated induction of Bcl-2 expression and apoptosis inhibition. Blood 2002;99:2532-2540.

23. Trentin L, Frasson M, Donella-Deana A et al. Geldanamycin-induced Lyn dissociation from aberrant Hsp90-stabilized cytosolic complex is an early event in apoptotic mechanisms in B-chronic lymphocytic leukemia. Blood 2008;112:4665-4674.

24. Rodina A, Vilenchik M, Moulick K et al. Selective compounds define Hsp90 as a major inhibitor of apoptosis in small-cell lung cancer. Nat Chem Biol 2007;3:498-507.

25. Hertlein E, Wagner AJ, Jones J et al. 17-DMAG targets the nuclear factor-kappaB family of proteins to induce apoptosis in chronic lymphocytic leukemia: clinical implications of HSP90 inhibition. Blood 2010;116:45-53.

26. Castro JE, Prada CE, Loria O et al. ZAP-70 is a novel conditional heat shock protein 90 (Hsp90) client: inhibition of Hsp90 leads to ZAP-70 degradation, apoptosis, and impaired signaling in chronic lymphocytic leukemia. Blood 2005;106:2506-2512.

27. Johnson AJ, Wagner AJ, Cheney CM et al. Rituximab and 17-allylamino-17-demethoxygeldanamycin induce synergistic apoptosis in B-cell chronic lymphocytic leukaemia. Br J Haematol 2007;139:837-844.

28. Jones DT, Addison E, North JM et al. Geldanamycin and herbimycin A induce apoptotic killing of B chronic lymphocytic leukemia cells and augment the cells' sensitivity to cytotoxic drugs. Blood 2004;103:1855-1861.

29. Best OG, Che Y, Singh N et al. The Hsp90 inhibitor SNX-7081 synergizes with and restores sensitivity to fludarabine in chronic lymphocytic leukemia cells with lesions in the TP53 pathway: a potential treatment strategy for fludarabine refractory disease. Leuk Lymphoma 2012

30. Best OG, Singh N, Forsyth C, Mulligan SP. The novel Hsp-90 inhibitor SNX7081 is significantly more potent than 17-AAG against primary CLL cells and a range of haematological cell lines, irrespective of lesions in the TP53 pathway. Br J Haematol 2010;151:185-188.

31. Lin K, Rockliffe N, Johnson GG, Sherrington PD, Pettitt AR. Hsp90 inhibition has opposing effects on wild-type and mutant p53 and induces p21 expression and cytotoxicity

irrespective of p53/ATM status in chronic lymphocytic leukaemia cells. Oncogene 2008;27:2445-2455.

32. Harris NL, Jaffe ES, Diebold J et al. World Health Organization classification of neoplastic diseases of the hematopoietic and lymphoid tissues: report of the Clinical Advisory Committee meeting-Airlie House, Virginia, November 1997. J Clin Oncol 1999;17:3835-3849.

33. Zhou X, Chen X, Hu L et al. Polymorphisms involved in the miR-218-LAMB3 pathway and susceptibility of cervical cancer, a case-control study in Chinese women. Gynecol Oncol 2010;117:287-290.

34. Bartel DP. MicroRNAs: target recognition and regulatory functions. Cell 2009;136:215-233.

35. Visone R, Rassenti LZ, Veronese A et al. Karyotype-specific microRNA signature in chronic lymphocytic leukemia. Blood 2009;114:3872-3879.

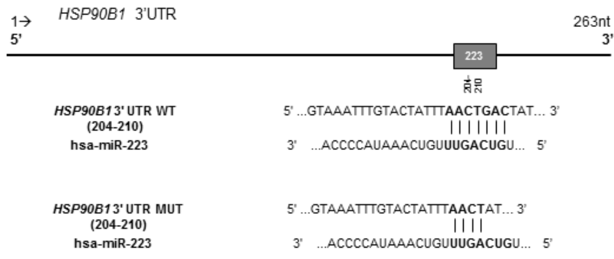
36. Pekarsky Y, Santanam U, Cimmino A et al. Tcl1 expression in chronic lymphocytic leukemia is regulated by miR-29 and miR-181. Cancer Res 2006;66:11590-11593.

37. Mott JL, Kobayashi S, Bronk SF, Gores GJ. mir-29 regulates Mcl-1 protein expression and apoptosis. Oncogene 2007;26:6133-6140.

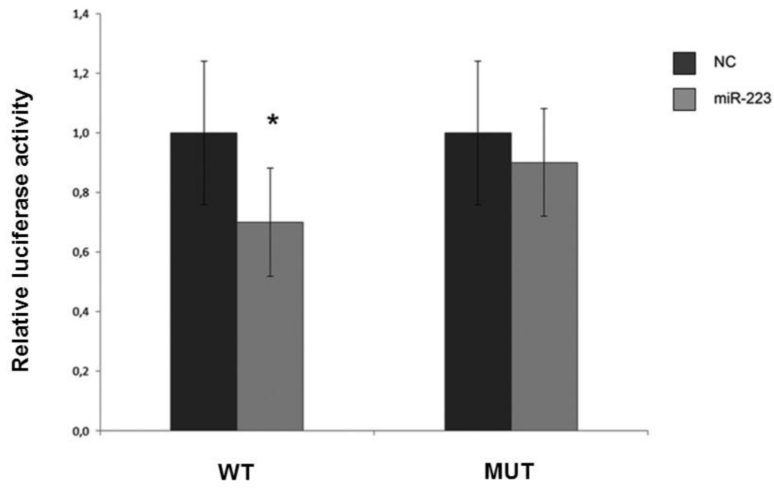
38. Cohen-Saidon C, Carmi I, Keren A, Razin E. Antiapoptotic function of Bcl-2 in mast cells is dependent on its association with heat shock protein 90beta. Blood 2006;107:1413-1420.

FIGURE LEGENDS

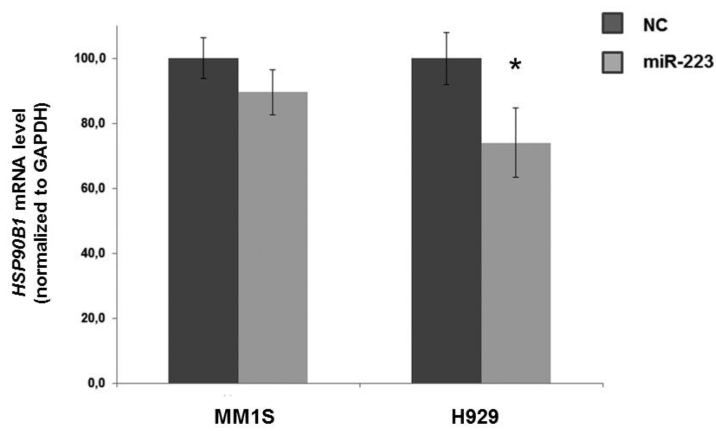
A



B



C



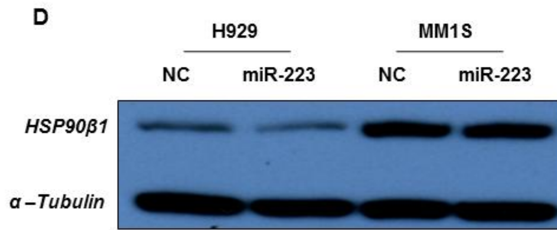


Figure 1. HSP90B1 is a target of miR-223

(A) 3´UTR HSP90B1 region (263 nt length) with a predicted binding site for miR-223 at 204-210 nt (grey box). The figure shows the mature miR-223 sequence (hsa-miR-223) aligned with both HSP90B1 3´UTR wild type (WT, up) and with the polymorphism (MUT, below). The seed region is shown in bold. The rs2307842 polymorphism disrupts the putative binding site for miR-223 by deleting the last three nucleotides of the seed region.

(B) Luciferase reporter assays to confirm targeting of HSP90B1 3´UTR by miR-223. Ectopic miR-223 expression inhibits wild-type but not mutant HSP90B1 3´UTR reporter activity in HEK293 cells. Cells were co-transfected with miR-223 precursor/negative control (NC) miRNA and with either WT or MUT HSP90B1 3´UTR reporter construct. Luciferase activity assay was performed at 24h after transfection. The columns represent normalized relative luciferase activity (RLU) by means with 95% confidence intervals from 4 independent experiments (Mann-Whitney test, *P<0.05).

(C) and (D) Forced overexpression of miR-223 downregulated HSP90B1 mRNA (C) and protein (D) expression in H929 cell line (wt) but not in MM1S (mutant). Cells were transfected with miR-223 precursors and negative controls. Twenty-four hours later, cells were analyzed for HSP90B1 expression by semi-quantitative PCR (C) and western-blot (D). The data shown are representative for 3 independent experiments (Mann-Whitney test, *P<0.05).

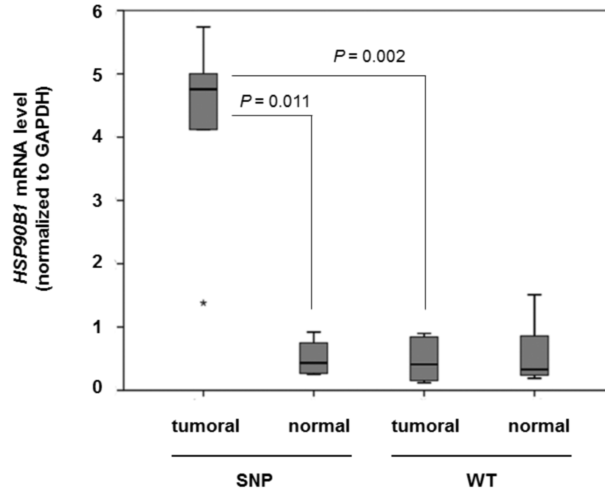


Figure 2. Gene expression levels of HSP90B1 in CLL patients with rs2307842 assessed by semi-quantitative PCR analysis.

Box plots show the relative upregulation of HSP90B1 in CLL patients with rs rs2307842 (MUT) compared with wt-CLL patients (WT). The thick line inside the box plot indicates the median expression levels and the box shows the 25th and 75th percentiles, while the whiskers show the maximum and minimum values. Outliers (extreme values falling out of the main distribution) are represented by open circles. Statistical significance was determined using the Mann-Whitney U test (*P*<0.05).

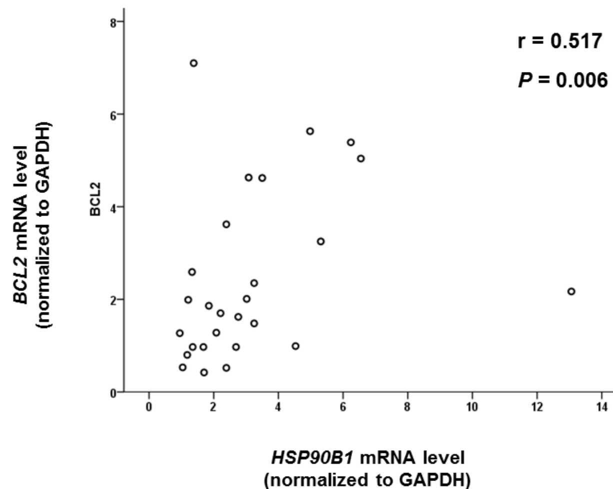
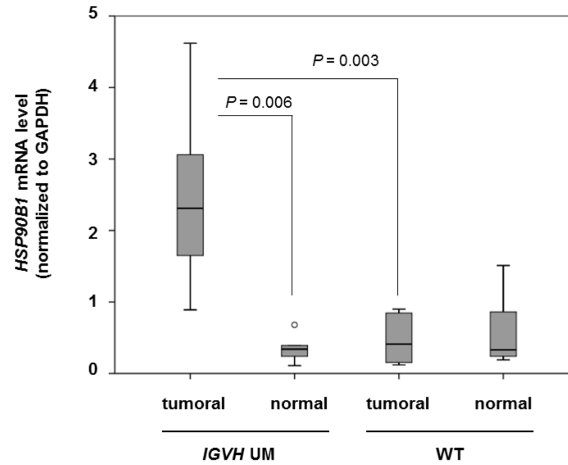


Figure 3. Expression levels of HSP90B1 are positively correlated with BCL-2 expression in CLL patients. Pearson´s correlation analysis of the relative expression levels of both genes determined using semi-quantitative PCR in 27 paired samples.



Supplementary Figure 1. Gene expression levels of HSP90B1 in *IGVH*-UM CLL patients assessed by semi-quantitative PCR analysis.

Box plots show the relative upregulation of HSP90B1 in CLL patients with rs2307842 (MUT) compared with wt-CLL patients (WT). The thick line inside the box plot indicates the median expression levels and the box shows the 25th and 75th percentiles, while the whiskers show the maximum and minimum values. Outliers (extreme values falling out of the main distribution) are represented by open circles. Statistical significance was determined using the Mann-Whitney U test (P<0.05).

SUPPLEMENTARY METHODS

1. NimbleGen Target-Region Capture

A custom Sequence Capture 385K Human Array was designed and manufactured by Roche NimbleGen. A total of 385,000 unique, overlapping probes 60–90 nucleotides in length were designed including all known exons and untranslated regions (UTRs). In total, 1564 exons from 93 distinct target genes and 2 target regions 13q14.3: 50043128-50382849 bp and 17p13.1: 7500000-7535000; NCBI build 36.1, hg18) were selected, targeting 845212 bases. The genes had been selected according to our previous gene expression data and their relevance in CLL, and included, for example, HSP90B1, TP53, ATM, PHLPP1, E2F1, RAPGEF2 or PI3K. Approximately 5 µg of genomic DNA from 4 CLL patients was fragmented to a size range of 300–500 base pairs (bp) with the use of a GS Nebulizer Kit (Roche Applied Science) to generate blunt-ended fragments. The fragmented DNA was purified (DNA Clean & Concentrator-25, Zymo Research) and analyzed on an Agilent Bioanalyzer 2100 DNA Chip 7500 according to the manufacturer's instructions. The fragmented DNA was then processed according to the recommended NimbleGen protocol (Roche Applied Science, User Guide 3.1; July 2008). In brief, linkers were ligated to the polished fragments in the library to provide a priming site for post-enrichment amplification of the eluted fragment pool. The linker-terminated fragments were then denatured to produce single-stranded products. The resulting library was hybridized to a custom 385K array for 72 h at 42ºC, with the use of the NimbleGen Sequence Capture Hybridization System 4. The hybridized DNA from the target regions was washed and eluted with the use of a NimbleGen Wash and Elution Kit according to the manufacturer's instructions. The eluted sample was amplified by ligation-mediated PCR with the use of primers complementary to the sequence of the adaptors

2. 454 Sequencing

We applied NGS technology using 454 FLX Titanium chemistry according to the manufacturer's protocols (Roche Applied Science).[1] Sequencing-compatible linkers were ligated to the eluted samples from the capture microarrays. The libraries were subsequently diluted, amplified on beads using emulsion PCR and sequenced using the 454 FLX sequencing instrument.

3. Sequencing data analysis

Basic raw data analysis was carried out using the GS Run Browser and GS Reference Mapper software version 2.0.01 (Roche Applied Science). Following in silico removal of the linker sequence, each sequence read was compared with the entire appropriate version of the human genome. Captured sequences mapped uniquely back to regions within the target regions were considered sequencing hits. These were then used to calculate the percentage of reads that did hit target regions, and the fold sequencing coverage for the entire target region. All putative variances were first compared with published single nucleotide polymorphism (SNP) data (dbSNP build 130; http://www.ncbi.nlm.nih.gov/projects/SNP). We used a custom-made data analysis pipeline to annotate detected variants with various kinds of information, including known single-nucleotide polymorphisms (SNPs), amino acid consequences, genomic location and miRNA binding sites.

4. Coverage statistics

According to NCBI build 36.1, hg18 reference genome, the final target bases covering the 93 target genes and target regions 13q14.3 and 17p13.1 were defined to be 845

212 bp; of those 750 594, target bases (99.39%) were covered by capture oligonucleotides as defined by NimbleGens default settings for probe selection. 5 134 bp (0.6%) of the initial target region were omitted due to reasons of specificity and uniqueness (Supplementary Table S4). This was sufficient to reach an average target coverage of 21.7- fold per individual (Supplementary Table S5).

Reference List

1. Margulies M, Egholm M, Altman WE et al. Genome sequencing in microfabricated high-density picolitre reactors. Nature 2005;437:376-380.

Discusión general

La leucemia linfática crónica es una hemopatía maligna heterogénea, no solo en su curso clínico sino también en su biología. Esta diversidad se debe, en gran medida, a las alteraciones genéticas que presentan sus células. Las alteraciones citogenéticas y el estado mutacional de IGVH constituyen los factores genéticos de mayor relevancia pronóstica en la LLC,[8,124] pero hasta el momento no se ha identificado una única mutación o anomalía genética responsable del origen y desarrollo de esta enfermedad.[125,126] Teniendo en cuenta la gran heterogeneidad molecular de la LLC, los estudios genómicos masivos y el análisis global de las posibles rutas celulares afectadas pueden ser de gran utilidad en el análisis genético de estos enfermos, mucho más eficaces que el estudio enfocado hacia dianas únicas.

El análisis combinado del genoma y del perfil de expresión de los enfermos con LLC ha permitido la identificación de una nueva ganancia recurrente en 20q13.

En la LLC, las alteraciones citogenéticas determinan subgrupos con diferente valor pronóstico. Al ser una enfermedad tan heterogénea, los microarrays constituyen una herramienta muy útil para su estudio global. Por esta razón nos planteamos realizar un análisis combinado del genoma y transcriptoma de 67 enfermos con LLC mediante arrays genómicos y de expresión.

Los arrays genómicos permiten el estudio global del genoma mediante la identificación de las ganancias y de las pérdidas del material tumoral. No requieren células en división como ocurre con la citogenética convencional, lo que supone una gran ventaja dado el bajo índice mitótico de la LLC. Su principal ventaja es que permiten el análisis de todo el genoma en un solo experimento, lo cual se puede comparar con la realización de miles de estudios de HISF independientes.[127] Además tienen una resolución alta, que varía según el tipo de plataforma utilizada (200 pb-10 Kb), pero en cualquier caso muy superior a la de la FISH y la citogenética convencional. Permiten la detección de alteraciones crípticas pero no de traslocaciones. De esta manera, la aplicación de los arrays genómicos al estudio de la LLC ha permitido la detección de nuevas alteraciones citogenéticas, como las ganancias en 2p, o las deleciones de 8p, 18q y 22q.[66,128-131] Sin embargo, ningún estudio ha investigado la presencia de nuevas alteraciones citogenéticas en relación con los subgrupos citogenéticos definidos por FISH. Por eso el objetivo de nuestro trabajo fue estudiar la presencia de nuevas alteraciones recurrentes en LLC y su prevalencia en los diferentes subgrupos citogenéticos definidos por FISH.

En nuestro estudio detectamos alteraciones genéticas en el 75% de los pacientes analizados, un porcentaje muy similar a la clonalidad detectada mediante los estudios de FISH. Predominaban ligeramente las pérdidas (53%) sobre las ganancias, como corresponde a los procesos linfoides y como se ha descrito en LLC. Los arrays genómicos nos permitieron detectar nuevas alteraciones recurrentes además de las que se pueden identificar mediante los estudios de FISH. De esta manera, describimos una nueva ganancia en el cromosoma 20, en el 19% de los pacientes con LLC. Esta frecuencia es similar a la de otras alteraciones citogenéticas "clásicas" (+12, 11q- y 17p-).[19] El análisis genético identificó la región mínimamente ganada en 20q13.12, con un tamaño de ~2.31 Mb, próxima al amplicón ganado en el cáncer de mama, pero no comprendida en él.[132,133] Las ganancias en 20q13 no aparecen como única alteración citogenética en la LLC, sino que todos los pacientes con ganancias en esta región presentaban alteraciones genéticas adicionales, aunque su presencia no se asociaba significativamente con ninguna en concreto. En realidad, las ganancias en 20q estaban asociadas a una mayor complejidad genética, definida como un mayor número de cambios detectados por CGH arrays. La elevada complejidad genética se ha relacionado recientemente con una mayor progresión y menor respuesta al tratamiento en la LLC.[134-136] La presencia de un mayor número de cambios en los pacientes de LLC con ganancias en 20q sugiere que esta nueva entidad podría estar asociada a una enfermedad más avanzada, tal y como se ha sugerido en los linfomas no Hodgkin.[137] En los pacientes con ganancias en 20q se observó una tendencia a presentar un mayor recuento linfocitario y síntomas B, lo que sugiere que esta alteración podría constituir un marcador de evolución. Sin embargo en nuestro estudio no encontramos una correlación entre 20q+ y los datos de progresión de la enfermedad

Las alteraciones en el cromosoma 20 son muy frecuentes en cáncer y la presencia de ganancias en 20q se asocia con un comportamiento más agresivo y peor pronóstico. Pensamos que la ganancia de 20q13 en LLC puede tener importancia en la patogénesis o en la evolución de la enfermedad debido al papel que pueden desempeñar los genes que se localizan en ella, todos sobreexpresados en el estudio del perfil de expresión. Se han identificado 11 genes que codifican para proteínas relacionadas con cáncer, como PIGT, PI3, SLPI y WFDC2, cuya sobreexpresión se relaciona con progresión en varios tipos de tumores.[138-140] Por ello, un estudio más detalladode los genes localizados en esta región sería de gran interés en LLC.

El análisis de la expresión génica mediante microarrays de ARN ha permitido en los pacientes con leucemia mejorar el proceso diagnóstico y la clasificación de estas neoplasias.[106-110,141,142] En el caso de la LLC, estos análisis han permitido incluso definir nuevos factores pronósticos,

como ZAP70, gen que se encontró más expresado al analizar las diferencias en el perfil de expresión de los pacientes con y sin mutaciones de IGVH.[62,143] El estudio del perfil de expresión ha permitido profundizar en los mecanismos moleculares implicados en la patogénesis de la LLC mediante el análisis de la influencia de las alteraciones genéticas en la expresión de los genes localizados en los loci correspondientes.[35,49,144,145] Nuestros datos confirman que existe una asociación entre los niveles de expresión y las regiones alteradas en los diferentes subgrupos genéticos de LLC. Este efecto de dosis génica podría tener un efecto patogénico en la LLC, incluso en las nuevas alteraciones descritas, como es el caso de las ganancias en 20q.

Los arrays genómicos, además de describir nuevas alteraciones citogenéticas, han permitido profundizar en las características de los subgrupos definidos mediante FISH, como las deleciones en 13q14,[40-42,146] que son objeto de estudio en la segunda parte de esta tesis doctoral.

Los enfermos con LLC y deleción de 13q presentan diferentes características clínicas y biológicas.

Aunque las deleciones en 13q son a menudo citogenéticamente crípticas, representan la alteración citogenética más frecuente en los enfermos con LLC. Clásicamente se ha considerado que los pacientes con LLC que tienen un mejor pronóstico son los que presentan deleción en 13q14 como única alteración citogenética, incluso ligeramente mejor que los que tienen citogenética normal. Sin embargo esta afirmación está siendo cuestionada actualmente, ya que los enfermos de LLC con 13q- constituyen un grupo con un pronóstico heterogéneo. En el análisis de algunas de las series publicadas[124] y en la serie de 350 pacientes objeto de estudio del artículo II, se ha podido observar un entrecruzamiento de las curvas de supervivencia y de tiempo hasta la progresión en los pacientes con citogenética normal y pérdida del cromosoma 13q14. En relación con la supervivencia, este hecho tiene lugar a los 12-13 años. A partir de este momento, la supervivencia de los enfermos con FISH normal es mejor. En este trabajo de tesis doctoral se demuestra que el porcentaje de células que presentan deleción de 13q influye en la evolución de los enfermos de LLC. Además varios estudios han demostrado que el tamaño de la pérdida se puede asociar con características clínicas diferentes: la del(13q) de tipo II es mayor, implica la pérdida del gen *RB1* y se relaciona con progresión de la enfermedad, mientras que las deleciones de tipo I son de menor tamaño,

no suponen la pérdida de *RB1* y solo se relacionan con progresión si aparecen asociadas a otras alteraciones. [40,147]

En el artículo I los estudios de FISH y CGH arrays demostraron que los enfermos 13q- constituyen un grupo muy heterogéneo tanto en el tamaño de la deleción como en el porcentaje de células con esta alteración. A pesar de utilizar para nuestro estudio arrays genómicos, no pudimos analizar el tamaño de la pérdida en estos pacientes. La correlación obtenida entre los resultados de FISH y de CGH arrays fue excelente para todos los subgrupos citogenéticos excepto para los pacientes con 13q-. Pensamos que esto puede ser debido a (1) la ausencia de clones en nuestro BAC-array que mapeen adecuadamente 13q14 y (2) que las alteraciones genéticas, especialmente las pérdidas, son más difíciles de detectar mediante CGHa cuando están presentes en porcentajes bajos (determinados por FISH), y éste era el caso de, aproximadamente, la mitad de los pacientes 13q- de nuestra serie . En realidad, los casos en los que no se pudo determinar la presencia de deleción en 13q mediante el BAC-array presentaban un porcentaje de pérdida, por FISH, menor del 30%. Esto podría justificar la ausencia de correlación entre los resultados obtenidos con ambas técnicas en este subgrupo de enfermos. Por lo tanto, en nuestra serie inicial evaluamos el tamaño de la deleción de 13q mediante sondas de FISH (LSI-D13S319, para determinar si 13q está delecionado, y LSI-RB, que identifica las pérdidas de mayor tamaño) y analizamos la supervivencia. Esta aproximación es semejante a la realizada por Dal Bo et al.[47] La hipótesis de este grupo es que los pacientes con un mayor porcentaje de células 13q- y peor pronóstico presentan también deleciones de mayor tamaño. Nuestro análisis mostró que los pacientes con LLC y menores porcentajes de 13q- por FISH solían presentar deleciones de menor tamaño (que no afectaban a RB1). Además un menor porcentaje de células 13q- o una pérdida de menor tamaño se correlacionaban con una mayor supervivencia. Sin embargo, encontramos pacientes con porcentajes de pérdidas de 13q altos y tamaños de deleción pequeños, y estos casos no tenían una mejor supervivencia que los pacientes con porcentajes altos y deleciones de mayor tamaño, demostrando que la presencia de un porcentaje alto de pérdidas se asocia con una peor supervivencia independientemente del tamaño de la deleción.

La relevancia clínica del número de células que tienen una alteración citogenética se ha demostrado recientemente en un estudio del grupo inglés de LLC. Así, la presencia de más de un 20% de células con pérdidas en *TP53* se asocia con un pronóstico peor en los pacientes con LLC, mientras que los pacientes con pérdidas en *TP53* en menos del 20% de sus células tienen un pronóstico similar a los enfermos de la serie global.[148] Nuestros resultados reflejan que los pacientes que tienen una infiltración alta por células 13q- (≥ 80%) no muestran características

clínicas diferentes del resto de enfermos con LLC y 13q-, excepto por el hecho de tener un recuento linfocitario mayor así como una tendencia a presentar patrón de infiltración difuso de la médula ósea y más esplenomegalia. Sin embargo, estos pacientes tienen un curso clínico significativamente peor.

Aunque el seguimiento fue relativamente corto para una enfermedad crónica (mediana de 42 meses), el subgrupo de pacientes con un número alto de células con 13q- presentó tanto una menor supervivencia como un menor tiempo hasta la progresión. Estos resultados fueron similares al utilizar diferentes puntos de corte en el número de células 13q- (≤ 50; 51-79; ≥ 80; 10-40, 41-70, ≥ 70). En el grupo de pacientes con pérdida de 13q los únicos factores adversos para la supervivencia global en el análisis multivariante fueron la presencia de un número mayor de células 13q- y la sintomatología B. En nuestra experiencia, el mayor número de pérdidas en 13q14 no está relacionado con un mayor porcentaje de pérdidas bialélicas de 13q- en los pacientes. Este subgrupo de enfermos con homocigosis de 13q14 no presenta una supervivencia peor, aunque parecen precisar tratamiento antes (53 meses vs 134 meses; p = 0,05).

En nuestro conocimiento, es la primera vez que el porcentaje de células con pérdidas en el cromosoma 13q14 se asocia con la supervivencia y la progresión en la LLC. Nuestros resultados han sido posteriormente corroborados por dos grupos independientes, que confirman que los enfermos de LLC y un mayor porcentaje de células con deleción de 13q como única alteración citogenética tienen un peor pronóstico.[46,47] Sin embargo, hasta el momento no se han analizado en profundidad las características moleculares de los pacientes con LLC y 13q-. Este análisis de las diferencias moleculares que existen entre los enfermos con deleción de 13q podría explicar por qué estos pacientes presentan esta variabilidad clínica. Por eso decidimos realizar un estudio del perfil de expresión génico de los enfermos con LLC y 13q- con dos plataformas de arrays: HU 133 Plus 2.0 (artículo II) y Human Exon 1.0 ST (artículo III).

El análisis funcional preliminar realizado en el artículo II identificó importantes funciones celulares alteradas en los pacientes con 13q- en función del porcentaje de pérdidas. La mayoría de los genes cuya expresión estaba alterada estaban implicados en procesos de crecimiento y proliferación celular, apoptosis y metabolismo del calcio y del retículo. En el artículo III el análisis del perfil de expresión se realizó con el microarray Human Exon 1.0 ST que ofrece un análisis más detallado de la expresión génica que la anterior generación de chips. De esta manera los resultados obtenidos en el artículo III caracterizan las diferencias biológicas que subyacen en las diferencias clínicas observadas en los pacientes de LLC y

deleción de 13q con mayor detalle. En general, la correlación entre los resultados obtenidos con las dos plataformas fue excelente.

Para validar las diferencias observadas entre los subgrupos de enfermos con 13q- y visualizarlas, realizamos un análisis de componentes principales (PCA) sobre linfocitos B purificados en una serie independiente de pacientes. Para una mejor caracterización de las diferencias y semejanzas en esta serie de pacientes incluimos, además de los enfermos 13q-, pacientes con FISH normal y controles sanos. De nuevo los resultados mostraron que los enfermos de LLC con un elevado porcentaje de células 13q- tienen un perfil de expresión característico, que los diferencia de los linfocitos B sanos y también de los enfermos con menores porcentajes de células 13q-, cuyo perfil de expresión está muy próximo al de los enfermos con FISH normal.

El perfil de expresión génica de los enfermos con un mayor porcentaje de células 13q- es diferente al de los enfermos con porcentajes menores. Estas diferencias involucraban a más de 3000 genes, que están implicados en funciones celulares de gran importancia como la señalización por *BCR*, la apoptosis y la supervivencia celular. Los enfermos con un elevado porcentaje de células 13q- se caracterizan por la sobreexpresión de genes implicados en la señalización por BCR como *SYK, BLNK* y *PRKCB1*. La señalización por *BCR* es esencial para la supervivencia de los linfocitos B y se ha demostrado que desempeña un papel fundamental en el desarrollo de la LLC.[72] Además esta ruta está activada en los pacientes con LLC de peor pronóstico (IGVH no mutado) y la sobreexpresión de muchas de las moléculas implicadas en ella, como *SYK* y *PKC* se ha descrito en los casos no mutados y en estadios avanzados de la enfermedad.[67,149-151] Los enfermos con pérdidas de 13q >80% presentan una desregulación del balance entre proliferación y apoptosis debido a la alteración de la expresión de genes implicados en estos procesos, como sobreexpresión de *LEF1, RRAS, SOS1* y *BCL2* así como infraexpresión de *BAD, FAS* y varias caspasas. Recientemente se ha descrito que una activación sostenida de *BCR* podría tener un efecto antiapoptótico. Esta desregulación, que favorece que exista una mayor proliferación celular y menor apoptosis en los enfermos con mayores porcentajes de 13q-, podría justificar la mayor linfocitosis observada en este subgrupo de enfermos.

Nuestros resultados también demuestran que los enfermos con 13q- presentan un perfil específico de expresión de microRNAs. En la LLC se han definido varias firmas de microRNAS que caracterizan a los linfocitos B de los enfermos con LLC o se correlacionan con parámetros biológicos de relevancia, fundamentalmente con el estado mutacional de IGVH.[80,85,152,153] En

nuestro estudio describimos un perfil de microRNAs característico de los enfermos con porcentajes altos de pérdidas de 13q, en el que destaca la infraexpresión de miR-34a (relacionado con resistencia al tratamiento) y sobreexpresión de miR-155 así como una infraexpresión de miR-223 y miR-29, relacionados con progresión en LLC. Además comprobamos que la desregulación de los microRNAs influye en el perfil de expresión de los enfermos 13q-, ya que más de 400 de los genes de éste son dianas de los microRNAs alterados. Destaca, por su interés en LLC, la sobreexpresión de *BCL2, TCL1* y *LEF1*, dianas de miR-15a, miR-29a y miR-34a, respectivamente.

Además nuestros resultados mostraron que los enfermos con un mayor porcentaje de células 13q- tienen algunas características biológicas semejantes a las de los enfermos de los grupos citogenéticos de peor pronóstico como son 17p- y 11q-, al compartir la desregulación de genes implicados en varias funciones celulares clave como ciclo celular, muerte celular y proliferación. También compartían la desregulación de algunos miRNAs como la infraexpresión de miR-34a, miR-223 y la sobreexpresión de miR-155. Este hecho respalda nuestra hipótesis de que los enfermos de LLC con un elevado porcentaje de células con deleción de 13q tienen un comportamiento clínico más agresivo.

Las diferencias en el pronóstico clínico de estas dos variantes permanecen aún sin una explicación clara, tanto si los linfocitos clonales con 13q- del grupo de pacientes con un alto porcentaje de células con 13q- representan una forma más activada de la enfermedad, lo que favorecería el potencial proliferante del clon maligno, como si se trata de pacientes que pasan de tener un porcentaje de células 13q- bajo a un porcentaje elevado, por lo que la deleción de 13q constituiría un marcador no sólo de peor pronóstico sino también de evolución. Se necesitan por tanto, mayores evidencias experimentales para justificar estas hipótesis, y tanto el análisis de los perfiles de expresión como el análisis del estado mutacional de los linfocitos 13q- en estos subgrupos pueden resultar de gran utilidad para definir qué genes estarían interviniendo en esta posible evolución. En este sentido, la posibilidad de estudiar de manera evolutiva a enfermos con pérdidas de 13q en los que el porcentaje de pérdidas no se modifica sustancialmente en el tiempo y compararlos con las características de los enfermos con 13q- en los que se observa un aumento de pérdidas en la evolución de la enfermedad podría aportar datos relevantes y será objeto de estudios posteriores a esta tesis doctoral.

Por todo lo expuesto, nuestros hallazgos confirman que los pacientes con LLC y pérdida de 13q14 no pueden considerarse globalmente. Un gran número de células con pérdidas en 13q

cuantificadas por FISH se asocia con una supervivencia global y tiempo hasta la progresión más cortos y con un perfil de expresión génico y de microRNAs característico.

Análisis genético de la LLC por secuenciación masiva mediante captura de secuencias.

En la última parte de esta tesis doctoral realizamos un estudio mutacional de la LLC.

Hasta ahora los estudios de mutaciones en los pacientes con LLC estaban restringidos a un número limitado de genes: mutaciones somáticas de IGVH y mutaciones de P53 y ATM, asociadas o no a la deleción del gen. La importancia de las primeras en el valor pronóstico de la LLC es indiscutible y las segundas van cobrando cada vez más importancia.[8,48,50,51,124,154-158]

La aplicación de las nuevas tecnologías de secuenciación masiva permite detectar mutaciones presentes con una frecuencia muy baja, y analizar varios genes en diferentes pacientes de manera simultánea.En los trabajos publicados recientemente por el Consorcio Español para la secuenciación de la LLC,[125,126] del que nuestro grupo forma parte, se han identificado mutaciones recurrentes en genes que no se habían relacionado previamente con LLC, como NOTCH1, XPO1, MYD88 y SF3B1. Estos datos refuerzan el concepto de LLC como enfermedad heterogénea, ya que además se ha comprobado que algunas de estas mutaciones tienen importancia clínica, por lo que podrían ser en un futuro marcadores pronósticos como ahora lo son las alteraciones citogenéticas o el estado mutacional de IGVH.

Sin duda uno de los aspectos que más llama la atención en estos estudios es que, a pesar del gran número de pacientes analizados (105-310 pacientes) en la población española son muy pocos (ocho) los genes con alteraciones recurrentes y muy bajas las frecuencias de aparición (2.4%-12.2%). Ésta es una de las mayores dificultades de los estudios mutacionales en una enfermedad tan heterogénea como es la LLC.

Por eso decidimos seguir una estrategia de captura de secuencias seguida de secuenciación masiva con el sistema 454-FLX de Roche. De esta manera, en lugar de secuenciar todo el genoma, nos centramos en los genes (93) y las regiones cromosómicas (13q14.3 y 17p13.1) que consideramos de mayor interés en LLC. La selección se llevó a cabo teniendo en cuenta los datos publicados y nuestros resultados previos de perfil de expresión. Se secuenció el ADN procedente de linfocitos B purificados de 4 pacientes con LLC. En otros estudios, incluido los del Consorcio Español, se secuencian en paralelo el ADN tumoral y el normal del mismo paciente, para descartar *a priori* los polimorfismos (variaciones que estarían presentes en ambas fracciones). Nosotros decidimos secuenciar los linfocitos B clonales de 4 pacientes y

descartar posteriormente los polimorfismos mediante secuenciación convencional de la correspondiente fracción no tumoral. Además analizamos 4 pacientes con neoplasias hematológicas no LLC, como controles.

De esta manera identificamos más de 1600 variaciones por muestra (media: 1721, rango: 1618-1823 variaciones). De éstas, 226 estaban presentes en las 8 muestras, por lo que pensamos que podrían tratarse de polimorfismos y los descartamos. Además el análisis que diseñamos nos permitió caracterizar las variantes detectadas: polimorfismos descritos (dbSNP build 130), consecuencias de la variación en la secuencia de aminoácidos, localización genómica e interacción con microRNAs. Después de aplicar varios de estos filtros, obtuvimos un total de 73 variaciones que implicaban cambio de aminoácido en 33 genes. Algunos resultados eran esperables, como las variaciones en *ATM* y *P53*. Las variaciones detectadas en otros genes como *E2F1*, *RAPGEF2* o *PI3K* se han comprobado mediante secuenciación convencional, aunque no hemos obtenido resultados satisfactorios, bien porque se trataba de polimorfismos no descritos, presentes también en la fracción no tumoral del paciente (*RAPGEF2*) o bien porque no pudimos demostrar su presencia en más pacientes con LLC . La investigación del resto de genes en los que se han identificado mutaciones en este estudio se encuentra actualmente en desarrollo.

Estos resultados reflejan la dificultad de analizar la LLC desde el punto de vista mutacional, ya que a pesar de diseñar una estrategia dirigida hacia genes de interés, la baja frecuencia de aparición de las mutaciones en estos pacientes puede resultar decepcionante y son necesarias series muy amplias para obtener resultados significativos.

El polimorfismo rs2307842 en la región 3´UTR de *HSP90B1* provoca su sobreexpresión en los linfocitos B de los enfermos con LLC .

El análisis que diseñamos nos permitió, entre otras cosas, anotar las variantes detectadas en la base de datos de polimorfismos (dbSNP build 130; http://www.ncbi.nlm.nih.gov/projects/SNP) y determinar si alguna variante estaba relacionada con algún microRNA.
Los microRNAs (miRNAs) son una familia de RNAs pequeños, de unos 21-25 nucleótidos, que regulan la expresión de otros genes uniéndose a secuencias específicas de regiones 3´UTR del RNA mensajero diana y reprimiendo su traducción en función de la complementariedad entre el miRNA y el RNA mensajero. Algunos miRNAs desempeñan un papel directo en la oncogénesis, pudiendo actuar como oncogenes o como genes supresores de tumores, y su

expresión aberrante se ha asociado con muchos tipos de cánceres incluyendo tumores sólidos y hematológicos.[83,159-162] En el caso de la LLC, se han descrito varios miRNAs cuya expresión aberrante caracteriza a los linfocitos B de estos enfermos.[80,152,153]

En nuestro estudio mutacional identificamos un polimorfismo en la región 3´UTR de HSP90B1. Se trata de un polimorfismo ya descrito y anotado en las bases de datos (rs2307842), que supone la deleción de 4 nucleótidos, tres de ellos pertenecientes al sitio de unión de miR-223, predicho en las bases de datos (TargetScan, TarBase, miRecords e Ingenuity Knowledge Base). Teniendo en cuenta el importante papel que desempeñan los microRNAs en la patogénesis de la LLC, decidimos profundizar en este hallazgo.

Recientemente se ha descrito un nuevo tipo de polimorfismos, conocido como miR-SNPs. Son polimorfismos localizados en los sitios de unión de los microRNAS o próximos a ellos, que pueden afectar a la expresión de los genes regulados por estos miRNAs al interferir en la unión miRNA-mRNA y, por lo tanto, en la regulación de la diana.[163-166]

En primer lugar validamos, por primera vez, HSP90B1 como diana de miR-223 mediante experimentos de luciferasa. Además identificamos la deleción en la línea celular MM1S y comprobamos que en estas células la transfección con miR-223 no causaba la disminución de los niveles de HSP90B1, al contrario de lo observado en la línea celular H929, que era salvaje para rs2307842. De esta manera confirmamos que miR-223 regula la expresión de HSP90B1 y que la presencia de rs2307842 impide la correcta regulación por parte del miRNA, al alterar el sitio de unión en la región 3´UTR.

El siguiente paso en nuestro estudio fue analizar la presencia de este polimorfismo en los enfermos de LLC. Identificamos el SNP en el 25% de nuestros pacientes, lo que está de acuerdo con la incidencia descrita en las bases de datos pero no se observó asociación con ninguna variable clínica. Sin embargo, el estudio de la expresión de HSP90B1 en los linfocitos B clonales de enfermos con LLC demostró que la presencia del polimorfismo condiciona un incremento en la expresión de HSP90B1 en estas células con respecto a los linfocitos B sin el polimorfismo y, de forma sorprendente, también con respecto a las células no clonales de los pacientes con el polimorfismo. Esto podría sugerir que en las células con el polimorfismo existe un mecanismo regulador de la expresión de HSP90B1, que estaría alterado en los linfocitos B de la LLC. O bien que la regulación fisiológica de la expresión de HSP90B1 esté dañada en las células clonales de la LLC debido a la gran cantidad de alteraciones genéticas que pueden acumular.

Los miRNAs también son importantes en LLC porque regulan la expresión de muchos genes claves en la patogénesis de la enfermedad. Así se ha descrito un papel fundamental para miR-

29 y miR-181 al regular el oncogen *TCL1*, que previamente se había asociado a un fenotipo agresivo.[89] miR15a y miR-16-1 regulan no sólo la expresión de *BCL2* sino también de varios factores con actividad pro-proliferativa, como *CCND2, CCND3, CD4* y *CDK6*.[39,94] Recientemente se ha descrito la relación que existe entre la señalización por *BCR* en los linfocitos B de la LLC y miR-155,[95] y cada vez son más numerosos los estudios que intentan descifrar la compleja relación que existe entre la deleción de 17p y la familia miR-34, dianas de TP53.[167] Se ha demostrado que la infraexpresión de miR-223 se relaciona con un peor pronóstico en los pacientes con LLC.[152,153] Además es uno de los factores incluidos en el score que se ha propuesto recientemente para clasificar a los pacientes con LLC en términos de supervivencia sin enfermedad y supervivencia global.[168] Dada la importancia de miR-223 en LLC y al no haberse comprobado la existencia de ninguna diana conocida en esta enfermedad, nos planteamos si la sobreexpresión de HSP90B1 podría ser el mecanismo, o uno de los mecanismos, responsables del peor comportamiento clínico de los enfermos que infraexpresan miR-223. Por eso medimos la expresión de HSP90B1 en los enfermos de LLC con IGVH no mutado. Así comprobamos que los linfocitos B de estos enfermos también se caracterizaban por la sobreexpresión de HSP90B1. Por lo tanto, proponemos que la desregulación de la expresión de HSP90B1 en los linfocitos B de LLC con el polimorfismo podría ser similar a la de los linfocitos B sin mutación somática de IGVH y con infraexpresión de miR-223.

HSP90 es una proteína de choque térmico involucrada en el plegamiento, la activación y el ensamblaje de muchas proteínas. Se sobreexpresa en muchos tipos de cáncer, favoreciendo la supervivencia celular.[169-171] Su inhibición es una estrategia terapéutica de interés creciente en cáncer y también en LLC.[172] Los inhibidores de HSP90 han demostrado ser eficaces en estudios preclínicos incluso en pacientes con LLC refractarios a los tratamientos convencionales, por lo que se presentan como una alternativa muy interesante.[173-175] En LLC se ha demostrado que la inhibición de HSP90 tiene como consecuencia la degradación de un gran número de moléculas necesarias para la supervivencia celular y finalmente, la apoptosis.[176-179] En nuestro estudio hemos observado que en los pacientes con LLC la sobreexpresión de *HSP90B1* se asocia con sobreexpresión de *BCL2*, molécula que desempeña un papel fundamental en la apoptosis, aunque no está claro aún el mecanismo por el cual *HSP90* podría actuar sobre *BCL2*.[180]

Nuestro trabajo resalta la importancia de la regulación de la expresión génica por parte de los miRNAS en la LLC, así como de la presencia de polimorfismos. Además estos resultados abren nuevas vías en la comprensión de los mecanismos etiopatogénicos de la LLC y pueden resultar útiles para la investigación de nuevas terapias.

El trabajo de investigación llevado a cabo durante el desarrollo de esta tesis doctoral ha permitido identificar, gracias a un análisis global (genoma-exoma-transcriptoma), nuevas alteraciones genéticas que podrían tener relevancia clínica en la LLC.

En primer lugar, la integración de los datos procedentes de los BAC arrays y de los arrays de expresión ha permitido describir una nueva alteración citogenética recurrente en el brazo largo del cromosoma 20 en los enfermos con LLC, que no se asocia con ninguna alteración citogenética establecida pero que podría representar un marcador de evolución en estos pacientes. El análisis del perfil de expresión ha proporcionado evidencias biológicas sobre las diferencias clínicas observadas en los enfermos con LLC y deleción de 13q, al demostrar que presentan un perfil de expresión génica y de microRNAs característico según el porcentaje de células con esta alteración citogenética. Finalmente, el estudio de la LLC desde el punto de vista mutacional ha permitido identificar un polimorfismo común en la población y en los pacientes con LLC que podría tener importancia en la patogénesis y en el tratamiento.

En conjunto estos resultados confirman que la LLC es una enfermedad con un marcado componente genético y muy heterogénea. Además resaltan la utilidad de las técnicas de análisis masivo (microarrays y secuenciación masiva) en el estudio del cáncer en general y de la LLC en particular, al tratarse de una enfermedad con tantas variantes y sin una causa genética establecida.

Conclusiones

1. La aplicación de los microarrays genómicos al estudio de la LLC permite la detección de alteraciones cromosómicas en el 75% de los pacientes. Además de las alteraciones determinadas por FISH, los microarrays genómicos identifican nuevas alteraciones recurrentes, como pérdidas en 11q13.3, 5q13.3–q14.1, 5q31.1 y 7q22, así como ganancias en 1q21.3–q22, 11q13.3, 16q23.2–q24.2, 6p21.31–p21.1, 20q13.12 y 10q22.3. Estos datos corroboran la heterogeneidad de la LLC.

2. El análisis del perfil de expresión génica de los enfermos de LLC demuestra que la mayoría de los genes desregulados se localizan en las regiones de ganancia o pérdida de material genético, tanto en los subgrupos citogenéticos “clásicos”, determinados mediante FISH, como en las nuevas alteraciones recurrentes identificadas. Esta observación sugiere que los cambios en la expresión génica están relacionados con las alteraciones en el número de copias, lo que confirma el efecto de dosis génica previamente descrito en las hemopatías malignas.

3. El análisis combinado del genoma y del transcriptoma de los enfermos de LLC revela la presencia de una nueva ganancia recurrente en 20q, en el 19% de los pacientes. El análisis genético sitúa la región mínimamente ganada en 20q13.12, con un tamaño de ~2.31 Mb. Esta ganancia se caracteriza por la sobreexpresión de los genes localizados ella y mayor complejidad genética, por lo que podría representar un marcador de evolución en estos pacientes.

4. Los enfermos con LLC y pérdida de 13q como única alteración citogenética constituyen un grupo muy heterogéneo, en el que el porcentaje de células que presentan esta alteración influye en el pronóstico de los pacientes. Los enfermos con un mayor porcentaje de células con pérdidas en 13q (≥80%) se caracterizan por tener un tiempo hasta el tratamiento y una supervivencia más

corta que los enfermos con un menor porcentaje de células con pérdida en 13q (<80%).

5. Los linfocitos B clonales de los enfermos con muchas pérdidas en 13q tienen más proliferación y menos apoptosis que los linfocitos B de los enfermos con pocas pérdidas en 13q, debido a la desregulación de genes de estas rutas, así como de algunos microRNAs tales como miR-15a, miR-29a, miR-34a, miR-155 y miR-181b. Además su perfil de expresión es similar al de las LLC de los grupos citogenéticos de alto riesgo (17p- y 11q-). Estos resultados justificarían las diferencias clínicas observadas entre los dos subgrupos de LLC con pérdida de 13q

6. El polimorfismo rs2307842, localizado en la región 3´UTR de *HSP90B1*, impide la correcta regulación de miR-223 sobre este gen al alterar su sitio de unión. Esto provoca la sobreexpresión de *HSP90B1* en los linfocitos B de los enfermos con LLC, pero no en las células normales de enfermos con el polimorfismo.

7. Hsa-miR-223 regula la expresión de *HSP90B1*, inhibiéndola. Éste podría ser uno de los mecanismos patogénicos que explicaría por qué los enfermos que infraexpresan miR-223 presentan un peor pronóstico. Además en los enfermos con LLC existe una correlación directa entre los niveles de expresión de *HSP90B1* y *BCL2*, lo que podría condicionar una mayor resistencia a la apoptosis en los enfermos con sobreexpresión de *HSP90B1*.

.

Referencias

1. Rozman C, Montserrat E. Chronic lymphocytic leukemia. N Engl J Med 1995;333:1052-1057.

2. Kay NE, Hamblin TJ, Jelinek DF et al. Chronic lymphocytic leukemia. Hematology Am Soc Hematol Educ Program 2002193-213.

3. Orfao A, Gonzalez M, San Miguel JF et al. B-cell chronic lymphocytic leukaemia: prognostic value of the immunophenotype and the clinico-haematological features. Am J Hematol 1989;31:26-31.

4. Chiorazzi N, Rai KR, Ferrarini M. Chronic lymphocytic leukemia. N Engl J Med 2005;352:804-815.

5. Carney DA, Wierda WG. Genetics and molecular biology of chronic lymphocytic leukemia. Curr Treat Options Oncol 2005;6:215-225.

6. Guipaud O, Deriano L, Salin H et al. B-cell chronic lymphocytic leukaemia: a polymorphic family unified by genomic features. Lancet Oncol 2003;4:505-514.

7. Ghia P, ten BE, Sanz E et al. Ordering of human bone marrow B lymphocyte precursors by single-cell polymerase chain reaction analyses of the rearrangement status of the immunoglobulin H and L chain gene loci. J Exp Med 1996;184:2217-2229.

8. 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.

9. Maloum K, Davi F, Merle-Beral H et al. Expression of unmutated VH genes is a detrimental prognostic factor in chronic lymphocytic leukemia. Blood 2000;96:377-379.

10. Ghia P, Stamatopoulos K, Belessi C et al. Geographic patterns and pathogenetic implications of IGHV gene usage in chronic lymphocytic leukemia: the lesson of the IGHV3-21 gene. Blood 2005;105:1678-1685.

11. Zenz T, Dohner H, Stilgenbauer S. Genetics and risk-stratified approach to therapy in chronic lymphocytic leukemia. Best Pract Res Clin Haematol 2007;20:439-453.

12. Rozman C, Bosch F, Montserrat E. Chronic lymphocytic leukemia: a changing natural history? Leukemia 1997;11:775-778.

13. Stilgenbauer S, Lichter P, Dohner H. Genetic features of B-cell chronic lymphocytic leukemia. Rev Clin Exp Hematol 2000;4:48-72.

14. Juliusson G, Oscier DG, Fitchett M et al. Prognostic subgroups in B-cell chronic lymphocytic leukemia defined by specific chromosomal abnormalities. N Engl J Med 1990;323:720-724.

15. Dicker F, Schnittger S, Haferlach T, Kern W, Schoch C. Immunostimulatory oligonucleotide-induced metaphase cytogenetics detect chromosomal aberrations in 80% of CLL patients: A study of 132 CLL cases with correlation to FISH, IgVH status, and CD38 expression. Blood 2006;108:3152-3160.

16. Rigolin GM, Cibien F, Martinelli S et al. Chromosome aberrations detected by conventional karyotyping using novel mitogens in chronic lymphocytic leukemia with "normal" FISH: correlations with clinicobiologic parameters. Blood 2012;119:2310-2313.

17. Heerema NA, Byrd JC, Dal Cin PS et al. Stimulation of chronic lymphocytic leukemia cells with CpG oligodeoxynucleotide gives consistent karyotypic results among laboratories: a CLL Research Consortium (CRC) Study. Cancer Genet Cytogenet 2010;203:134-140.

18. Muthusamy N, Breidenbach H, Andritsos L et al. Enhanced detection of chromosomal abnormalities in chronic lymphocytic leukemia by conventional cytogenetics using CpG oligonucleotide in combination with pokeweed mitogen and phorbol myristate acetate. Cancer Genet 2011;204:77-83.

19. Dohner H, Stilgenbauer S, Benner A et al. Genomic aberrations and survival in chronic lymphocytic leukemia. N Engl J Med 2000;343:1910-1916.

20. Stilgenbauer S, Bullinger L, Benner A et al. Incidence and clinical significance of 6q deletions in B cell chronic lymphocytic leukemia. Leukemia 1999;13:1331-1334.

21. Cuneo A, Rigolin GM, Bigoni R et al. Chronic lymphocytic leukemia with 6q- shows distinct hematological features and intermediate prognosis. Leukemia 2004;18:476-483.

22. Dohner H, Stilgenbauer S, James MR et al. 11q deletions identify a new subset of B-cell chronic lymphocytic leukemia characterized by extensive nodal involvement and inferior prognosis. Blood 1997;89:2516-2522.

23. Dickinson JD, Smith LM, Sanger WG et al. Unique gene expression and clinical characteristics are associated with the 11q23 deletion in chronic lymphocytic leukaemia. Br J Haematol 2005;128:460-471.

24. Dohner H, Fischer K, Bentz M et al. p53 gene deletion predicts for poor survival and non-response to therapy with purine analogs in chronic B-cell leukemias. Blood 1995;85:1580-1589.

25. Pittman S, Catovsky D. Prognostic significance of chromosome abnormalities in chronic lymphocytic leukaemia. Br J Haematol 1984;58:649-660.

26. Cordone I, Masi S, Mauro FR et al. p53 expression in B-cell chronic lymphocytic leukemia: a marker of disease progression and poor prognosis. Blood 1998;91:4342-4349.

27. Hernandez JM, Mecucci C, Criel A et al. Cytogenetic analysis of B cell chronic lymphoid leukemias classified according to morphologic and immunophenotypic (FAB) criteria. Leukemia 1995;9:2140-2146.

28. Reddy KS. Chronic lymphocytic leukaemia profiled for prognosis using a fluorescence in situ hybridisation panel. Br J Haematol 2006;132:705-722.

29. Amiel A, Leopold L, Gronich N et al. The influence of different chromosomal aberrations on molecular cytogenetic parameters in chronic lymphocytic leukemia. Cancer Genet Cytogenet 2006;167:145-149.

30. Athanasiadou A, Stamatopoulos K, Tsompanakou A et al. Clinical, immunophenotypic, and molecular profiling of trisomy 12 in chronic lymphocytic leukemia and comparison with other karyotypic subgroups defined by cytogenetic analysis. Cancer Genet Cytogenet 2006;168:109-119.

31. Del Principe MI, Del PG, Venditti A et al. Clinical significance of soluble p53 protein in B-cell chronic lymphocytic leukemia. Haematologica 2004;89:1468-1475.

32. Dickinson JD, Gilmore J, Iqbal J et al. 11q22.3 deletion in B-chronic lymphocytic leukemia is specifically associated with bulky lymphadenopathy and ZAP-70 expression but not reduced expression of adhesion/cell surface receptor molecules. Leuk Lymphoma 2006;47:231-244.

33. Grever MR, Lucas DM, Dewald GW et al. Comprehensive assessment of genetic and molecular features predicting outcome in patients with chronic lymphocytic leukemia: results from the US Intergroup Phase III Trial E2997. J Clin Oncol 2007;25:799-804.

34. Sindelarova L, Michalova K, Zemanova Z et al. Incidence of chromosomal anomalies detected with FISH and their clinical correlations in B-chronic lymphocytic leukemia. Cancer Genet Cytogenet 2005;160:27-34.

35. Hernandez JA, Rodriguez AE, Gonzalez M et al. A high number of losses in 13q14 chromosome band is associated with a worse outcome and biological differences in patients with B-cell chronic lymphoid leukemia. Haematologica 2009;94:364-371.

36. Liu Y, Corcoran M, Rasool O et al. Cloning of two candidate tumor suppressor genes within a 10 kb region on chromosome 13q14, frequently deleted in chronic lymphocytic leukemia. Oncogene 1997;15:2463-2473.

37. Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T. Identification of novel genes coding for small expressed RNAs. Science 2001;294:853-858.

38. Migliazza A, Bosch F, Komatsu H et al. Nucleotide sequence, transcription map, and mutation analysis of the 13q14 chromosomal region deleted in B-cell chronic lymphocytic leukemia. Blood 2001;97:2098-2104.

39. Calin GA, Dumitru CD, Shimizu M et al. Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proc Natl Acad Sci U S A 2002;99:15524-15529.

40. Parker H, Rose-Zerilli MJ, Parker A et al. 13q deletion anatomy and disease progression in patients with chronic lymphocytic leukemia. Leukemia 2011;25:489-497.

41. Mosca L, Fabris S, Lionetti M et al. Integrative genomics analyses reveal molecularly distinct subgroups of B-cell chronic lymphocytic leukemia patients with 13q14 deletion. Clin Cancer Res 2010;16:5641-5653.

42. Ouillette P, Erba H, Kujawski L et al. Integrated genomic profiling of chronic lymphocytic leukemia identifies subtypes of deletion 13q14 1. Cancer Res 2008;68:1012-1021.

43. Bullrich F, Fujii H, Calin G et al. Characterization of the 13q14 tumor suppressor locus in CLL: identification of ALT1, an alternative splice variant of the LEU2 gene. Cancer Res 2001;61:6640-6648.

44. Rondeau G, Moreau I, Bezieau S et al. Comprehensive analysis of a large genomic sequence at the putative B-cell chronic lymphocytic leukaemia (B-CLL) tumour suppresser gene locus. Mutat Res 2001;458:55-70.

45. Palamarchuk A, Efanov A, Nazaryan N et al. 13q14 deletions in CLL involve cooperating tumor suppressors. Blood 2010;115:3916-3922.

46. Van Dyke DL, Shanafelt TD, Call TG et al. A comprehensive evaluation of the prognostic significance of 13q deletions in patients with B-chronic lymphocytic leukaemia. Br J Haematol 2010;148:544-550.

47. Dal BM, Rossi FM, Rossi D et al. 13q14 Deletion size and number of deleted cells both influence prognosis in chronic lymphocytic leukemia. Genes Chromosomes Cancer 2011;50:633-643.

48. Rossi D, Gaidano G. ATM and chronic lymphocytic leukemia: mutations, and not only deletions, matter. Haematologica 2012;97:5-8.

49. Haslinger C, Schweifer N, Stilgenbauer S et al. Microarray gene expression profiling of B-cell chronic lymphocytic leukemia subgroups defined by genomic aberrations and VH mutation status. J Clin Oncol 2004;22:3937-3949.

50. Guarini A, Marinelli M, Tavolaro S et al. ATM gene alterations in chronic lymphocytic leukemia patients induce a distinct gene expression profile and predict disease progression. Haematologica 2012;97:47-55.

51. Zenz T, Eichhorst B, Busch R et al. TP53 mutation and survival in chronic lymphocytic leukemia. J Clin Oncol 2010;28:4473-4479.

52. Callet-Bauchu E, Salles G, Gazzo S et al. Translocations involving the short arm of chromosome 17 in chronic B-lymphoid disorders: frequent occurrence of dicentric rearrangements and possible association with adverse outcome. Leukemia 1999;13:460-468.

53. Dohner H, Stilgenbauer S, Fischer K, Bentz M, Lichter P. Cytogenetic and molecular cytogenetic analysis of B cell chronic lymphocytic leukemia: specific chromosome aberrations identify prognostic subgroups of patients and point to loci of candidate genes. Leukemia 1997;11 Suppl 2:S19-S24.

54. Haferlach C, Dicker F, Schnittger S, Kern W, Haferlach T. Comprehensive genetic characterization of CLL: a study on 506 cases analysed with chromosome banding

analysis, interphase FISH, IgV(H) status and immunophenotyping. Leukemia 2007;21:2442-2451.

55. Ross FM, Stockdill G. Clonal chromosome abnormalities in chronic lymphocytic leukemia patients revealed by TPA stimulation of whole blood cultures. Cancer Genet Cytogenet 1987;25:109-121.

56. Mayr C, Speicher MR, Kofler DM et al. Chromosomal translocations are associated with poor prognosis in chronic lymphocytic leukemia. Blood 2006;107:742-751.

57. Michaux L, Wlodarska I, Rack K et al. Translocation t(1;6)(p35.3;p25.2): a new recurrent aberration in "unmutated" B-CLL. Leukemia 2005;19:77-82.

58. Cavazzini F, Cuneo A, De AC et al. Abnormalities of chromosomes 1p34-36, 4p16, 4q35, 9q11-32 and +7 represent novel recurrent cytogenetic rearrangements in chronic lymphocytic leukemia. Leuk Lymphoma 2004;45:1197-1203.

59. Klein U, Tu Y, Stolovitzky GA et al. Transcriptional analysis of the B cell germinal center reaction. Proc Natl Acad Sci U S A 2003;100:2639-2644.

60. Gutierrez NC, Ocio EM, de las RJ et al. Gene expression profiling of B lymphocytes and plasma cells from Waldenstrom's macroglobulinemia: comparison with expression patterns of the same cell counterparts from chronic lymphocytic leukemia, multiple myeloma and normal individuals. Leukemia 2007;21:541-549.

61. Lanasa MC, Allgood SD, Slager SL et al. Immunophenotypic and gene expression analysis of monoclonal B-cell lymphocytosis shows biologic characteristics associated with good prognosis CLL. Leukemia 2011;25:1459-1466.

62. Rosenwald A, Alizadeh AA, Widhopf G et al. Relation of gene expression phenotype to immunoglobulin mutation genotype in B cell chronic lymphocytic leukemia. J Exp Med 2001;194:1639-1647.

63. Vasconcelos Y, De VJ, Vallat L et al. Gene expression profiling of chronic lymphocytic leukemia can discriminate cases with stable disease and mutated Ig genes from those with progressive disease and unmutated Ig genes. Leukemia 2005;19:2002-2005.

64. Crespo M, Bosch F, Villamor N et al. ZAP-70 expression as a surrogate for immunoglobulin-variable-region mutations in chronic lymphocytic leukemia. N Engl J Med 2003;348:1764-1775.

65. Rassenti LZ, Huynh L, Toy TL et al. ZAP-70 compared with immunoglobulin heavy-chain gene mutation status as a predictor of disease progression in chronic lymphocytic leukemia. N Engl J Med 2004;351:893-901.

66. Schwaenen C, Nessling M, Wessendorf S et al. Automated array-based genomic profiling in chronic lymphocytic leukemia: development of a clinical tool and discovery of recurrent genomic alterations. Proc Natl Acad Sci U S A 2004;101:1039-1044.

67. Rodriguez A, Villuendas R, Yanez L et al. Molecular heterogeneity in chronic lymphocytic leukemia is dependent on BCR signaling: clinical correlation. Leukemia 2007;21:1984-1991.

68. Chiaretti S, Tavolaro S, Marinelli M et al. Evaluation of TP53 mutations with the AmpliChip p53 research test in chronic lymphocytic leukemia: Correlation with clinical outcome and gene expression profiling. Genes Chromosomes Cancer 2011

69. Dameshek W. Chronic lymphocytic leukemia--an accumulative disease of immunolgically incompetent lymphocytes. Blood 1967;29:Suppl-84.

70. Montserrat E, Sanchez-Bisono J, Vinolas N, Rozman C. Lymphocyte doubling time in chronic lymphocytic leukaemia: analysis of its prognostic significance. Br J Haematol 1986;62:567-575.

71. Messmer BT, Messmer D, Allen SL et al. In vivo measurements document the dynamic cellular kinetics of chronic lymphocytic leukemia B cells. J Clin Invest 2005;115:755-764.

72. Chiorazzi N. Cell proliferation and death: forgotten features of chronic lymphocytic leukemia B cells. Best Pract Res Clin Haematol 2007;20:399-413.

73. Buske C, Gogowski G, Schreiber K et al. Stimulation of B-chronic lymphocytic leukemia cells by murine fibroblasts, IL-4, anti-CD40 antibodies, and the soluble CD40 ligand. Exp Hematol 1997;25:329-337.

74. Granziero L, Ghia P, Circosta P et al. Survivin is expressed on CD40 stimulation and interfaces proliferation and apoptosis in B-cell chronic lymphocytic leukemia. Blood 2001;97:2777-2783.

75. Messmer BT, Albesiano E, Efremov DG et al. Multiple distinct sets of stereotyped antigen receptors indicate a role for antigen in promoting chronic lymphocytic leukemia. J Exp Med 2004;200:519-525.

76. 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;109:259-270.

77. Caligaris-Cappio F. Role of the microenvironment in chronic lymphocytic leukaemia. Br J Haematol 2003;123:380-388.

78. Chen L, Widhopf G, Huynh L et al. Expression of ZAP-70 is associated with increased B-cell receptor signaling in chronic lymphocytic leukemia. Blood 2002;100:4609-4614.

79. 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.

80. Calin GA, Liu CG, Sevignani C et al. MicroRNA profiling reveals distinct signatures in B cell chronic lymphocytic leukemias. Proc Natl Acad Sci U S A 2004;101:11755-11760.

81. Calin GA, Sevignani C, Dumitru CD et al. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. Proc Natl Acad Sci U S A 2004;101:2999-3004.

82. Calin GA, Croce CM. MicroRNA-cancer connection: the beginning of a new tale. Cancer Res 2006;66:7390-7394.

83. Calin GA, Croce CM. MicroRNA signatures in human cancers. Nat Rev Cancer 2006;6:857-866.

84. Cimmino A, Calin GA, Fabbri M et al. miR-15 and miR-16 induce apoptosis by targeting BCL2. Proc Natl Acad Sci U S A 2005;102:13944-13949.

85. Li S, Moffett HF, Lu J et al. MicroRNA expression profiling identifies activated B cell status in chronic lymphocytic leukemia cells. PLoS One 2011;6:e16956.

86. Fulci V, Chiaretti S, Goldoni M et al. Quantitative technologies establish a novel microRNA profile of chronic lymphocytic leukemia. Blood 2007;109:4944-4951.

87. Zanette DL, Rivadavia F, Molfetta GA et al. miRNA expression profiles in chronic lymphocytic and acute lymphocytic leukemia. Braz J Med Biol Res 2007;40:1435-1440.

88. Zhang J, Jima DD, Jacobs C et al. Patterns of microRNA expression characterize stages of human B-cell differentiation. Blood 2009;113:4586-4594.

89. Pekarsky Y, Santanam U, Cimmino A et al. Tcl1 expression in chronic lymphocytic leukemia is regulated by miR-29 and miR-181. Cancer Res 2006;66:11590-11593.

90. Bichi R, Shinton SA, Martin ES et al. Human chronic lymphocytic leukemia modeled in mouse by targeted TCL1 expression. Proc Natl Acad Sci U S A 2002;99:6955-6960.

91. Mott JL, Kobayashi S, Bronk SF, Gores GJ. mir-29 regulates Mcl-1 protein expression and apoptosis. Oncogene 2007;26:6133-6140.

92. Longo PG, Laurenti L, Gobessi S et al. The Akt/Mcl-1 pathway plays a prominent role in mediating antiapoptotic signals downstream of the B-cell receptor in chronic lymphocytic leukemia B cells. Blood 2008;111:846-855.

93. Pepper C, Lin TT, Pratt G et al. Mcl-1 expression has in vitro and in vivo significance in chronic lymphocytic leukemia and is associated with other poor prognostic markers. Blood 2008;112:3807-3817.

94. Liu Q, Fu H, Sun F et al. miR-16 family induces cell cycle arrest by regulating multiple cell cycle genes. Nucleic Acids Res 2008;36:5391-5404.

95. Yin Q, Wang X, McBride J, Fewell C, Flemington E. B-cell receptor activation induces BIC/miR-155 expression through a conserved AP-1 element. J Biol Chem 2008;283:2654-2662.

96. Nana-Sinkam SP, Croce CM. MicroRNA in chronic lymphocytic leukemia: transitioning from laboratory-based investigation to clinical application. Cancer Genet Cytogenet 2010;203:127-133.

97. Tjio JH. The chromosome number of man. Am J Obstet Gynecol 1978;130:723-724.

98. Watson JD, Crick FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature 1953;171:737-738.

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

100. Armstrong SA, Kung AL, Mabon ME et al. Inhibition of FLT3 in MLL. Validation of a therapeutic target identified by gene expression based classification. Cancer Cell 2003;3:173-183.

101. Yeoh EJ, Ross ME, Shurtleff SA et al. Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. Cancer Cell 2002;1:133-143.

102. Shaffer AL, Rosenwald A, Hurt EM et al. Signatures of the immune response. Immunity 2001;15:375-385.

103. Ferrando AA, Neuberg DS, Staunton J et al. Gene expression signatures define novel oncogenic pathways in T cell acute lymphoblastic leukemia. Cancer Cell 2002;1:75-87.

104. Schoch C, Kohlmann A, Schnittger S et al. Acute myeloid leukemias with reciprocal rearrangements can be distinguished by specific gene expression profiles. Proc Natl Acad Sci U S A 2002;99:10008-10013.

105. 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;101:3109-3117.

106. Bullinger L, Dohner K, Bair E et al. Use of gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia. N Engl J Med 2004;350:1605-1616.

107. Alvarez P, Saenz P, Arteta D et al. Transcriptional profiling of hematologic malignancies with a low-density DNA microarray. Clin Chem 2007;53:259-267.

108. Chng WJ, Schop RF, Price-Troska T et al. Gene-expression profiling of Waldenstrom macroglobulinemia reveals a phenotype more similar to chronic lymphocytic leukemia than multiple myeloma. Blood 2006;108:2755-2763.

109. Dunphy CH. Gene expression profiling data in lymphoma and leukemia: review of the literature and extrapolation of pertinent clinical applications. Arch Pathol Lab Med 2006;130:483-520.

110. Margalit O, Somech R, Amariglio N, Rechavi G. Microarray-based gene expression profiling of hematologic malignancies: basic concepts and clinical applications. Blood Rev 2005;19:223-234.

111. Solinas-Toldo S, Lampel S, Stilgenbauer S et al. Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. Genes Chromosomes Cancer 1997;20:399-407.

112. Pinkel D, Segraves R, Sudar D et al. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. Nat Genet 1998;20:207-211.

113. Pollack JR, Perou CM, Alizadeh AA et al. Genome-wide analysis of DNA copy-number changes using cDNA microarrays. Nat Genet 1999;23:41-46.

114. Snijders AM, Nowak N, Segraves R et al. Assembly of microarrays for genome-wide measurement of DNA copy number. Nat Genet 2001;29:263-264.

115. Oostlander AE, Meijer GA, Ylstra B. Microarray-based comparative genomic hybridization and its applications in human genetics. Clin Genet 2004;66:488-495.

116. Veltman JA, Fridlyand J, Pejavar S et al. Array-based comparative genomic hybridization for genome-wide screening of DNA copy number in bladder tumors. Cancer Res 2003;63:2872-2880.

117. Fritz B, Schubert F, Wrobel G et al. Microarray-based copy number and expression profiling in dedifferentiated and pleomorphic liposarcoma. Cancer Res 2002;62:2993-2998.

118. Peiffer DA, Le JM, Steemers FJ et al. High-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping. Genome Res 2006;16:1136-1148.

119. Grubor V, Krasnitz A, Troge JE et al. Novel genomic alterations and clonal evolution in chronic lymphocytic leukemia revealed by representational oligonucleotide microarray analysis (ROMA). Blood 2009;113:1294-1303.

120. de Leeuw RJ, Davies JJ, Rosenwald A et al. Comprehensive whole genome array CGH profiling of mantle cell lymphoma model genomes. Hum Mol Genet 2004;13:1827-1837.

121. Kohlhammer H, Schwaenen C, Wessendorf S et al. Genomic DNA-chip hybridization in t(11;14)-positive mantle cell lymphomas shows a high frequency of aberrations and allows a refined characterization of consensus regions. Blood 2004;104:795-801.

122. Rubio-Moscardo F, Climent J, Siebert R et al. Mantle-cell lymphoma genotypes identified with CGH to BAC microarrays define a leukemic subgroup of disease and predict patient outcome. Blood 2005;105:4445-4454.

123. Tyybakinoja A, Saarinen-Pihkala U, Elonen E, Knuutila S. Amplified, lost, and fused genes in 11q23-25 amplicon in acute myeloid leukemia, an array-CGH study. Genes Chromosomes Cancer 2006;45:257-264.

124. Damle RN, Wasil T, Fais F et al. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. Blood 1999;94:1840-1847.

125. Puente XS, Pinyol M, Quesada V et al. Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. Nature 2011;475:101-105.

126. Quesada V, Conde L, Villamor N et al. Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. Nat Genet 2011;44:47-52.

127. Shaffer LG, Bejjani BA. A cytogeneticist's perspective on genomic microarrays. Hum Reprod Update 2004;10:221-226.

128. Pfeifer D, Pantic M, Skatulla I et al. Genome-wide analysis of DNA copy number changes and LOH in CLL using high-density SNP arrays. Blood 2007;109:1202-1210.

129. Gunn SR, Bolla AR, Barron LL et al. Array CGH analysis of chronic lymphocytic leukemia reveals frequent cryptic monoallelic and biallelic deletions of chromosome 22q11 that include the PRAME gene. Leuk Res 2009;33:1276-1281.

130. Patel A, Kang SH, Lennon PA et al. Validation of a targeted DNA microarray for the clinical evaluation of recurrent abnormalities in chronic lymphocytic leukemia. Am J Hematol 2008;83:540-546.

131. Tyybakinoja A, Vilpo J, Knuutila S. High-resolution oligonucleotide array-CGH pinpoints genes involved in cryptic losses in chronic lymphocytic leukemia. Cytogenet Genome Res 2007;118:8-12.

132. Tanner MM, Tirkkonen M, Kallioniemi A et al. Increased copy number at 20q13 in breast cancer: defining the critical region and exclusion of candidate genes. Cancer Res 1994;54:4257-4260.

133. Tanner MM, Tirkkonen M, Kallioniemi A et al. Independent amplification and frequent co-amplification of three nonsyntenic regions on the long arm of chromosome 20 in human breast cancer. Cancer Res 1996;56:3441-3445.

134. Kay NE, Eckel-Passow JE, Braggio E et al. Progressive but previously untreated CLL patients with greater array CGH complexity exhibit a less durable response to chemoimmunotherapy. Cancer Genet Cytogenet 2010;203:161-168.

135. Kujawski L, Ouillette P, Erba H et al. Genomic complexity identifies patients with aggressive chronic lymphocytic leukemia. Blood 2008;112:1993-2003.

136. Ouillette P, Fossum S, Parkin B et al. Aggressive chronic lymphocytic leukemia with elevated genomic complexity is associated with multiple gene defects in the response to DNA double-strand breaks. Clin Cancer Res 2010;16:835-847.

137. Carter SL, Eklund AC, Kohane IS, Harris LN, Szallasi Z. A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers. Nat Genet 2006;38:1043-1048.

138. Wu G, Guo Z, Chatterjee A et al. Overexpression of glycosylphosphatidylinositol (GPI) transamidase subunits phosphatidylinositol glycan class T and/or GPI anchor attachment 1 induces tumorigenesis and contributes to invasion in human breast cancer. Cancer Res 2006;66:9829-9836.

139. Scotto L, Narayan G, Nandula SV et al. Identification of copy number gain and overexpressed genes on chromosome arm 20q by an integrative genomic approach in cervical cancer: potential role in progression. Genes Chromosomes Cancer 2008;47:755-765.

140. Clauss A, Lilja H, Lundwall A. A locus on human chromosome 20 contains several genes expressing protease inhibitor domains with homology to whey acidic protein. Biochem J 2002;368:233-242.

141. Gutierrez NC, Lopez-Perez R, Hernandez JM et al. Gene expression profile reveals deregulation of genes with relevant functions in the different subclasses of acute myeloid leukemia. Leukemia 2005;19:402-409.

142. Valk PJ, Verhaak RG, Beijen MA et al. Prognostically useful gene-expression profiles in acute myeloid leukemia. N Engl J Med 2004;350:1617-1628.

143. Wiestner A, Rosenwald A, Barry TS et al. ZAP-70 expression identifies a chronic lymphocytic leukemia subtype with unmutated immunoglobulin genes, inferior clinical outcome, and distinct gene expression profile. Blood 2003;101:4944-4951.

144. Aalto Y, El-Rifa W, Vilpo L et al. Distinct gene expression profiling in chronic lymphocytic leukemia with 11q23 deletion. Leukemia 2001;15:1721-1728.

145. Dickinson JD, Joshi A, Iqbal J et al. Genomic abnormalities in chronic lymphocytic leukemia influence gene expression by a gene dosage effect. Int J Mol Med 2006;17:769-778.

146. Hanlon K, Ellard S, Rudin CE et al. Evaluation of 13q14 status in patients with chronic lymphocytic leukemia using single nucleotide polymorphism-based techniques. J Mol Diagn 2009;11:298-305.

147. Zenz T, Mertens D, Dohner H, Stilgenbauer S. Importance of genetics in chronic lymphocytic leukemia. Blood Rev 2011;25:131-137.

148. Catovsky D, Richards S, Matutes E et al. Assessment of fludarabine plus cyclophosphamide for patients with chronic lymphocytic leukaemia (the LRF CLL4 Trial): a randomised controlled trial. Lancet 2007;370:230-239.

149. Guarini A, Chiaretti S, Tavolaro S et al. BCR ligation induced by IgM stimulation results in gene expression and functional changes only in IgV H unmutated chronic lymphocytic leukemia (CLL) cells. Blood 2008;112:782-792.

150. Su TT, Guo B, Kawakami Y et al. PKC-beta controls I kappa B kinase lipid raft recruitment and activation in response to BCR signaling. Nat Immunol 2002;3:780-786.

151. Bernal A, Pastore RD, Asgary Z et al. Survival of leukemic B cells promoted by engagement of the antigen receptor. Blood 2001;98:3050-3057.

152. Calin GA, Ferracin M, Cimmino A et al. A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. N Engl J Med 2005;353:1793-1801.

153. Visone R, Rassenti LZ, Veronese A et al. Karyotype-specific microRNA signature in chronic lymphocytic leukemia. Blood 2009;114:3872-3879.

154. Austen B, Powell JE, Alvi A et al. Mutations in the ATM gene lead to impaired overall and treatment-free survival that is independent of IGVH mutation status in patients with B-CLL. Blood 2005;106:3175-3182.

155. 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;29:2223-2229.

156. Ritgen M, Lange A, Stilgenbauer S et al. Unmutated immunoglobulin variable heavy-chain gene status remains an adverse prognostic factor after autologous stem cell transplantation for chronic lymphocytic leukemia. Blood 2003;101:2049-2053.

157. Zenz T, Krober A, Scherer K et al. Monoallelic TP53 inactivation is associated with poor prognosis in chronic lymphocytic leukemia: results from a detailed genetic characterization with long-term follow-up. Blood 2008;112:3322-3329.

158. Zenz T, Habe S, Denzel T et al. Detailed analysis of p53 pathway defects in fludarabine-refractory chronic lymphocytic leukemia (CLL): dissecting the contribution of 17p deletion, TP53 mutation, p53-p21 dysfunction, and miR34a in a prospective clinical trial. Blood 2009;114:2589-2597.

159. Volinia S, Calin GA, Liu CG et al. A microRNA expression signature of human solid tumors defines cancer gene targets. Proc Natl Acad Sci U S A 2006;103:2257-2261.

160. Croce CM. Causes and consequences of microRNA dysregulation in cancer. Nat Rev Genet 2009;10:704-714.

161. Fabbri M, Garzon R, Andreeff M et al. MicroRNAs and noncoding RNAs in hematological malignancies: molecular, clinical and therapeutic implications. Leukemia 2008;22:1095-1105.

162. Hui AB, Lenarduzzi M, Krushel T et al. Comprehensive MicroRNA profiling for head and neck squamous cell carcinomas. Clin Cancer Res 2010;16:1129-1139.

163. Mishra PJ, Humeniuk R, Mishra PJ et al. A miR-24 microRNA binding-site polymorphism in dihydrofolate reductase gene leads to methotrexate resistance. Proc Natl Acad Sci U S A 2007;104:13513-13518.

164. Bertino JR, Banerjee D, Mishra PJ. Pharmacogenomics of microRNA: a miRSNP towards individualized therapy. Pharmacogenomics 2007;8:1625-1627.

165. Mayr C, Hemann MT, Bartel DP. Disrupting the pairing between let-7 and Hmga2 enhances oncogenic transformation. Science 2007;315:1576-1579.

166. Zhou X, Chen X, Hu L et al. Polymorphisms involved in the miR-218-LAMB3 pathway and susceptibility of cervical cancer, a case-control study in Chinese women. Gynecol Oncol 2010;117:287-290.

167. Zenz T, Mohr J, Eldering E et al. miR-34a as part of the resistance network in chronic lymphocytic leukemia. Blood 2009;113:3801-3808.

168. Stamatopoulos B, Meuleman N, Haibe-Kains B et al. microRNA-29c and microRNA-223 down-regulation has in vivo significance in chronic lymphocytic leukemia and improves disease risk stratification. Blood 2009;113:5237-5245.

169. Nylandsted J, Brand K, Jaattela M. Heat shock protein 70 is required for the survival of cancer cells. Ann N Y Acad Sci 2000;926:122-125.

170. Broemer M, Krappmann D, Scheidereit C. Requirement of Hsp90 activity for IkappaB kinase (IKK) biosynthesis and for constitutive and inducible IKK and NF-kappaB activation. Oncogene 2004;23:5378-5386.

171. Sato S, Fujita N, Tsuruo T. Modulation of Akt kinase activity by binding to Hsp90. Proc Natl Acad Sci U S A 2000;97:10832-10837.

172. Trepel J, Mollapour M, Giaccone G, Neckers L. Targeting the dynamic HSP90 complex in cancer. Nat Rev Cancer 2010;10:537-549.

173. Best OG, Che Y, Singh N et al. The Hsp90 inhibitor SNX-7081 synergizes with and restores sensitivity to fludarabine in chronic lymphocytic leukemia cells with lesions in the TP53 pathway: a potential treatment strategy for fludarabine refractory disease. Leuk Lymphoma 2012

174. Best OG, Singh N, Forsyth C, Mulligan SP. The novel Hsp-90 inhibitor SNX7081 is significantly more potent than 17-AAG against primary CLL cells and a range of haematological cell lines, irrespective of lesions in the TP53 pathway. Br J Haematol 2010;151:185-188.

175. Lin K, Rockliffe N, Johnson GG, Sherrington PD, Pettitt AR. Hsp90 inhibition has opposing effects on wild-type and mutant p53 and induces p21 expression and cytotoxicity irrespective of p53/ATM status in chronic lymphocytic leukaemia cells. Oncogene 2008;27:2445-2455.

176. Castro JE, Prada CE, Loria O et al. ZAP-70 is a novel conditional heat shock protein 90 (Hsp90) client: inhibition of Hsp90 leads to ZAP-70 degradation, apoptosis, and impaired signaling in chronic lymphocytic leukemia. Blood 2005;106:2506-2512.

177. Trentin L, Frasson M, Donella-Deana A et al. Geldanamycin-induced Lyn dissociation from aberrant Hsp90-stabilized cytosolic complex is an early event in apoptotic mechanisms in B-chronic lymphocytic leukemia. Blood 2008;112:4665-4674.

178. Hertlein E, Wagner AJ, Jones J et al. 17-DMAG targets the nuclear factor-kappaB family of proteins to induce apoptosis in chronic lymphocytic leukemia: clinical implications of HSP90 inhibition. Blood 2010;116:45-53.

179. Johnson AJ, Wagner AJ, Cheney CM et al. Rituximab and 17-allylamino-17-demethoxygeldanamycin induce synergistic apoptosis in B-cell chronic lymphocytic leukaemia. Br J Haematol 2007;139:837-844.

180. Cohen-Saidon C, Carmi I, Keren A, Razin E. Antiapoptotic function of Bcl-2 in mast cells is dependent on its association with heat shock protein 90beta. Blood 2006;107:1413-1420.