UNIVERSIDAD AUTÓNOMA DE MADRID

FACULTAD DE CIENCIAS

DEPARTAMENTO DE BIOLOGÍA MOLECULAR

Advanced Classical Hodgkin Lymphoma:

New insights in prognostic factors using gene and microRNA expression signatures

By

Beatriz Sánchez Espiridión

Tesis presentada en la Universidad Autónoma de Madrid para optar al grado de Doctor

Madrid 2011

This thesis, submitted for the degree of Doctor in Philosophy at the "Universidad Autónoma de Madrid", has been performed in the laboratory of Lymphoma Group at Spanish National Cancer Research Center (CNIO). The work was done under the guidance of Dr. Juan Fernando García García and Dr. Miguel Ángel Piris Pinilla.

El trabajo presentado en esta memoria para optar al grado de doctor en la Universidad Autónoma de Madrid ha sido realizado en el laboratorio de Linfomas del Centro Nacional de Investigaciones Oncológicas (CNIO) bajo la dirección del Dr. Juan Fernando García y del Dr. Miguel Ángel Piris Pinilla.

This work has been supported by grants from the Fondo de Investigaciones Sanitarias (PI08/1985, PI05/1623, PI05/2800, PI05/2327, RTICC RD06/0020/0107), and the Ministerio de Ciencia y Tecnología (SAF2008-03871). Beatriz Sánchez Espiridión was supported by a grant from the Ministerio de Ciencia e Innovación (BEFI/FIS).

El trabajo presentado en esta tesis ha sido posible gracias a la financiación del Fondo de Investigaciones Sanitarias (PI08/1985, PI05/1623, PI05/2800, PI05/2327, RTICC RD06/0020/0107), y el Ministerio de Ciencia y Tecnología (SAF2008-03871), España. Beatriz Sánchez Espiridión fue financiada por una beca del Ministerio de Ciencia e Innovación (BEFI/FIS).

DEDICADA A TODAS AQUELLAS PERSONAS QUE DE UNA MANERA U OTRA HAN FORMADO PARTE DE MI VIDA DURANTE ESTOS AÑOS.

ACKNOWLEDGEMENTS

¡GRACIAS!

No hay mejor palabra para comenzar una tesis; no hay nada mejor que un sincero sentimiento de agradecimiento para todas aquellas personas que me han acompañado a lo largo de estos años de doctorado y para los que me animaron a emprender este camino y los que me acompañarán en él.

Son muchos todos aquellos a quienes debería mencionar aquí, y seguro algunos no quedarán reflejados. Pido disculpas de antemano.

En primer lugar, GRACIAS a mis codirectores Miguel Ángel y Juan Fernando, por darme la oportunidad de trabajar con vosotros y de aprender a vuestro lado. GRACIAS Miguel Ángel por tus ideas, palabras y consejos sobre ciencia (y no sólo ciencia); por darme la oportunidad de conocer de cerca el mundo de la investigación traslacional y darme alas para volar. GRACIAS a Juan Fernando, por su entusiasmo, su contagiosa vitalidad, por aguantar a Beatriz diciendo "esto no saldrá", "esto no me lo creo"... todo ello sin perder la sonrisa y siempre con mucha paciencia. GRACIAS por darme la oportunidad de ir a Houston, ofreciéndome un proyecto para realizar allí y proponerme las personas de referencia y contacto adecuadas (¡GRACIAS Paco!), facilitándome todo lo necesario para esa aventura tan apasionante.

GRACIAS a todo linfomas y al CNIO, a mi casa de acogida durante estos últimos 5 años de vida, a mi pecera del 307A donde desde el 1 de septiembre de 2006 comencé a hacerme mi hueco. GRACIAS Nerea, por tu compañía diaria y en especial esos (suspirantes) viernes; GRACIAS también a Lore, (Lorelain) por nuestras conversaciones filosóficas y nuestros emails (a pesar de estar sentadas al lado); a Santiago, por saber de todo y estar "casi-siempre" dándome datos para "entretenerme". GRACIAS a Dani, por estar siempre ahí resolviendo cualquier duda ya desde aquel primer desayuno que compartimos como estudiantes de verano; a Lina por su sonriente manera de ser y su amabilidad; a Bea H. por su profundo análisis de todo; a Magda por contarnos sus aventuras y desventuras; a Espe (de Bilb…no, no… ¡perdón! Vitoria) por darme pie siempre a recalcar que yo… ¡"Soy Navarra"! ; a Piertxo, (uno de mis italianos favoritos) por sus "curiosos" comentarios acerca de todo (con especial mención a sus análisis no estrictamente científicos junto a Dani); a Helen (Helenai) por ser "la alegría de la huerta" y el "alma matter de la fiesta"; a Marién (nuestra supermami) y su inestimable ayuda con la ardua tarea de conseguir todos los datos de nuestra base cHL; a Mar, por estar siempre ahí conmigo desde el principio, ayudándome ya en mis primeras extracciones de RNA y siendo un gran apoyo (GRACIAS al pequeño Victor); a "nuestra inigualable y única Elena" (superteniente), por tenerme siempre a raya y ser el mejor equipo para "las cuantis" (echaré de menos tu dulzura) ; GRACIAS también a May, por su buen talante y gran corazón; a María (Socorro) Rodríguez por ser… ¡tal cual! ; a Cristina, por su sonriente, coherente y positiva manera de ser.

GRACIAS a Gonzalo, por ser "uno más" de linfomas y un gran punto de "consulta de información" durante todos estos años.

GRACIAS también a todos aquellos que , sin serlo oficialmente, han formado siempre parte de nuestro grupo: Mariaje y Lauri (aunque a veces nos olvidásemos de ellas para comer); a

nuestro colaborador Carlos Montalbán y "sus alfombras" ; a Luis Lombardía, a Manolo Morente (y sus comentarios…); a María Lozano y su microdisector ; a Elena (Elenita) Doménech, ¡una pena coincidir tan poco contigo!; a Rocío Ramos y sus continuas visitas a la pecera del 307A (aunque no vinieses a verme a mí) ; asimismo, GRACIAS a los que pasaron por linfomas a lo largo de estos años con los cuales pude compartir mis días: a Abel (por darme su testigo en el proyecto HL); David Álvarez; Antonio Carreño; Denise (fantástica Denisa) y nuestras consultas a revistas "científicas"; a todos nuestros estudiantes (Luis , Rubén (Rubito) , Bea D (Charlize) , Cristina…). GRACIAS a nuestra "campeona" Eva Mata… sí niña ¡por tu valentía y tu fuerza!, por darnos ejemplo.

GRACIAS a Houston, a Pilar (Pilarice) y Dani G, por todo lo compartido en tierras texanas.

Sin olvidarme de aquellos que dentro del CNIO y fuera de él me han acompañado haciendo que estos años en Madrid sean inolvidables (ellos saben quienes son).

Muchísimas gracias a también a todo mi mundo y mi "cuadrilla" de Pamplona por ser como son y escuchar "mis historias" (aunque no entendiesen nada). Especial mención a quienes me transmitieron la ilusión por la biología y la ciencia desde ya el instituto y la universidad (¡GRACIAS José Luis!).

Y cómo no, GRACIAS a mi familia, especialmente a mis queridísimos padres y mi hermano Alex quienes siempre me impulsaron y dieron libertad para elegir, transmitiendo total confianza en mí, y apoyándome en los momentos de flaqueza. GRACIAS por ser un ejemplo a seguir.

Beatriz

INDEX

SUMMARY/ RESUMEN

Classical Hodgkin lymphoma (cHL) is assumed to be a curable tumor, but an important fraction of patients with advanced disease do not respond favorably to the standard chemotherapy regimens based on adriamycin and current predictive systems, based on clinical and analytical parameters, fail to accurately identify these high risk patients. Thus, the identification of biomarkers associated with treatment response remains a critical challenge in this lymphoid malignancy.

cHL represents a distinctive model of histological complexity, with a minor population of the characteristic tumor cells, the Hodgkin and Reed Sternberg cells (HRS) diluted in a complex reactive inflammatory background. The intrinsic relationship between the HRS cells and their microenvironment is only partially understood but important advances in the understanding of the cHL pathophysiology are being made. Indeed, microarray expression studies of tumor samples have led to the identification of signatures and biological processes associated to Hodgkin´s lymphoma pathogenesis and treatment response, showing that outcome in cHL patients could be related with both biological characteristic components of cHL tumors. However, no steps towards the translation of gene profiling results into routine clinical practice have been made.

This study identified gene subsets expressed either by the tumoral HRS cells and the Hodgkin´s microenvironment showing that robust methodologies based on quantitative real-time-PCR- (RT-PCR) assays are suitable for expression profiling of tumors and can be easily applied to paraffin-embedded samples. By a sequential analysis process, functional gene signatures associated with treatment response were first identified and finally a quantitative RT-PCR assay for patients with advanced cHL was described, incorporating a limited number of genes and pathways from tumor and microenvironment cell components, designed for application to routine formalin-fixed paraffin-embedded samples. This model was able to identify patient subgroups with highly different probabilities of treatment failure. Additionally, it was able to incorporate one of the well established clinical variables (Stage IV), thus integrating the main biological and molecular characteristics of the tumors related with treatment response and the tumor burden estimation in a single scoring system.

MicroRNAs (miRNAs) have been described to be essential regulators of cell differentiation and biological processes, emerging as robust predictor and prognostic markers. Specific signatures from the HRS cells and their microenvironment have been proposed, but the potential prognostic role of them remains unclear. Thus, this study additionally investigated whether miRNA signatures could allow to a better understanding of cHL pathogenesis and identified a cHL miRNA signature related to treatment response.

El linfoma de Hodgkin clásico presenta en general una alta tasa de curación, sin embargo una importante proporción de pacientes con estadios avanzados presentan quimiorresistencia a los protocolos quimioterapéuticos actuales basados en adriamicina (ABVD y variantes). Los sistemas pronósticos disponibles en la actualidad, basados en parámetros clínicos y analíticos como el IPS (International Prognostic Score, Índice Pronóstico Internacional), fallan a la hora de identificar una importante fracción de estos pacientes. Así, el empleo de técnicas moleculares robustas y la identificación de marcadores asociados con fallo terapéutico es esencial para el desarrollo de nuevas herramientas que permitan un diagnóstico de mayor precisión.

Asimismo, esta neoplasia linfoide representa un modelo único de complejidad histológica, se diferencia de la mayoría de las neoplasias malignas en su especial composición celular, de forma que en la masa tumoral las células neoplásicas (células de Hodgkin y Reed-Sternberg, HRS) son minoritarias mientras que el componente mayoritario está constituido por células inflamatorias .La compleja interacción existente entre las células neoplásicas y el microambiente tumoral no ha sido del todo elucidada si bien importantes avances en su conocimiento han sido realizados recientemente. De hecho, estudios de expresión génica han permitido la identificación de firmas moleculares asociadas a los fenómenos biológicos característicos de su patogenia y respuesta terapéutica, señalando la importancia de ambos componentes tumorales en la respuesta a terapia. Sin embargo, estos descubrimientos no han sido trasladados a la práctica clínica.

En el presente estudio, se identificaron firmas moleculares atribuibles tanto a las propias células tumorales como a su correspondiente microambiente, demostrando la idoneidad y potencial aplicación de ensayos basados en PCR cuantitativa para el análisis de muestras parafinadas. Se identificaron las rutas biológicas y genes característicos de cada uno de los componentes tumorales asociados a fallo terapéutico, y se desarrolló un ensayo de PCR cuantitativa capaz de identificar grupos de pacientes con distintas probabilidades de fallo. Asimismo, la inclusión de variables clínicas (Estadio IV) permitió el desarrollo de un modelo predictivo final combinando en un único sistema tanto características moleculares propias del tumor (genes) como factores relacionados con la masa tumoral.

Por otro lado, el estudio de microRNAs dentro del linfoma del Hodgkin ha permitido la identificación de firmas de microRNAs característicos sin estudiar su potencial predictivo. Dado su crucial papel en regulación se estudió la relevancia de su expresión dentro del contexto de

la enfermedad y su potencial poder predictivo identificándose finalmente una firma de microRNAS asociados a fallo terapéutico.

ABBREVIATIONS

ABVD Doxorubicin/adryamicin, bleomycin, vinblastine and dacarbazine

cHL Classical Hodgkin lymphoma

Ct Treshold cycle

CGH Comparative genomic hybridization

CR Complete response/ remission

CV Cross validated

DFS Disease free survival

DLDA Diagonal linear discriminant analysis

DNA Desoxiribonucleic acid

DSS Disease specific survival

EBV Epstein Barr Virus

ES Enrichment score

F Favorable

FDR False discovery rate

FFPE Formalin fixed paraffin embedded

FFS Failure free survival

GEP Gene expression

GSEA Gene set enrichment analysis

HL Hodgkin lymphomas

HRS Hodgkin and Red Sternberg

IARC International Agency for Research on Cancer

IPS International Prognostic Score

IRB Institutional review board

KNN K-nearest neighbor

LCM Laser capture microdissection

LD Lymphocyte depleted

LR Lymphocyte rich

MC Mixed cellularity

MDB Microenvironment Database.

MiRNA Micro ribonucleotic acid

MRS Molecular Risk Score

NCI National Cancer Institute

NES Normalized enrichment score

NHL Non-Hodgkin lymphomas

NK Natural Killer

NLPHL Nodular lymphocyte predominant Hodgkin lymphoma

NS Nodular sclerosis

OS Overall survival

P Progression

PFS Progression free survival

PR Partial response/ remission

RNA Ribonucleic acid

RT Retrotranscription

RT-qPCR Real time quantitative polymerase chain reaction

ROC Receiver operating curves

SEER Surveillance, Epidemiology and End Results

SNP Single nucleotide polymorphism

SVM Support Vector Machines

TDB Tumor Database

Th T helper

TLDA Taqman Low Density array

TMA Tissue microarray

T/NK T/ natural killer

U Unfavorable

UTR Untranslated region

USA United States of America

WHO World Health Organization

INTRODUCTION

1.1 Lymphomas.

Lymphomas are neoplasms derived from either B or T lymphocytes. They represent around 4-5 % of all diagnosed cancer cases in industrialized countries with an annual incidence rate of 120,000 new cases in Europe (7.000 in Spain) and 70.000 in United States of America (USA) based on data from the Surveillance Research Program of the National Cancer Institute of United States (NCI). Noteworthy, this incidence has grown since the last decades and according to the SEER (Surveillance, Epidemiology and End Results, http://seer.cancer.gov/), it´s estimated that 74.030 men and women (40.050 men and 33.980 women) were diagnosed and 21.530 men and women died of lymphoma in 2010 in USA.

The term ***"lymphoma"*** comprises a complex group of malignancies that includes more than 40 subtypes classified by the World Health Organization (WHO) attending to their morphology, immunophenotype and genetic characteristics. This distinction is not only relevant in terms of lymphoma pathogenesis; it also defines different clinical behaviors therefore determining treatment strategies. Although lymphoma classification is dynamic and frequent revisions are published, there are two major categories recognized:

A. **Non-Hodgkin lymphoma** (NHL), that constitute around 87.5% of all lymphomas and include neoplasms derived from either B or T/Natural killers cells (T/NK); with B-cell lymphomas representing around 90% of the total NHL cases.

B. **Hodgkin lymphoma** (HL), the first type of lymphoma described (Thomas Hodgkin, 1832) in which this thesis work has been focused.

Lymphoma classification is complex with an extension out from the scope of this work. This doctoral thesis is focused on HL. Thus, its main biological and clinical characteristics are described along this chapter.

1.2 Hodgkin Lymphoma.

Hodgkin lymphoma (HL) is a lymphoid neoplasm first described in patients in 1832 by Thomas Hodgkin. It´s relatively rare and according to the Leukemia and Lymphoma Society, Hodgkin Lymphoma represents about 11% of all lymphoma diagnoses. Approximately 8.000 cases (4.400 males and 3.820 females) of Hodgkin's Lymphoma are detected per year, thus representing less than 1% of all cancers. Nevertheless, this is a still a significant disease worldwide and around 1.000 new cases per year are diagnosed in Spain (National Spanish Statistic Institute data, www.ine.com).

The WHO Classification of HL (Table 1.1) defines two clinical, pathological and biological disease entities: classical Hodgkin lymphoma (cHL) (representing 95% of all cases) and nodular lymphocyte predominant Hodgkin lymphoma (NLPHL). So-called classical HL is the one further studied in this work and includes different histological subtypes: Nodular sclerosis (NS), mixed cellularity (MC), Lymphocyte rich (LR) and Lymphocyte depleted (LD) with nodular sclerosis and mixed cellularity accounting for the majority of classical HL cases.

Table 1.1 WHO 2008 classification of Hodgkin lymphoma

Nodular lymphocyte-predominant Hodgkin lymphoma (NLPHL)
Classical Hodgkin lymphoma (cHL)
Nodular sclerosis (NS)
Mixed cellularity (MC)
Lymphocyte rich (LR)
Lymphocyte depleted (LD)

1.2.1 Epidemiology

The distinguishing epidemiological feature of Hodgkin lymphoma is its bimodal age incidence curve characteristically seen in economically advantaged populations such as in United States, where the majority of patients are young adults. With a relatively few number of cases in children, there is a rapid increase of incidence rates among teenagers peaking at about 25 years. Incidence rates decline through middle age, after which they increase again with advancing age. However, in developing countries this bimodality disappears and a high incidence rate during childhood is observed.

A large number of studies point out a possible viral infectious etiology for HL. Several clinical features of the disease, such as fever, night sweats, and lymphoadenopathy, have suggested an infectious etiology to investigators from the time of its first description (Gutensohn, 1982; Mueller, 1992). However, most of them are not conclusive with exception of Epstein-Bar virus (EBV). Great interest in the disease has been generated by the use of molecular methods to identify the frequent presence of the Epstein-Barr virus genetic material in tumor specimens (Weiss et al., 1987). The Epstein-Barr virus, a human herpesvirus that is the cause of infectious mononucleosis, was first isolated from a Burkitt lymphoma tumor specimen. Scientific data linking this virus with Hodgkin lymphoma have become strong enough for the International Agency for Research on Cancer (IARC) in 1997 to classify the

Epstein- Barr virus as a group 1 carcinogen (that is, " the agent is carcinogenic to humans") for HL, thus validating the intuition of previous generations of scientific investigators.

Other data suggest genetic epidemiological factors. Although not recurrent genetic abnormalities have been described, familial aggregation and genetic susceptibility play important roles in the causation of Hodgkin lymphoma. Thus, it´s well established that first-degree relatives of patients with HL have about three fold increased risk of developing the disease (with a higher risk in identical twins, 100 fold higher (Mack et al., 1995))

1.2.2 Classical Hodgkin lymphoma. Histopathology and immunophenotype

cHL is a tumor composed by a heterogeneous cell composition with characteristic neoplastic cells (so called Hodgkin and Red Sternberg cells, HRS) diluted in a reactive inflammatory background composed of non-neoplastic B and T cells, macrophages, eosinophils, neutrophils, plasma cells and fibroblasts (Jaffe et al., 2001; Pileri et al., 2002).These tumoral cells are present in a low percentage (5-10%) within the reactive inflammatory background.

The diagnosis of cHL is established with the identification of these characteristic HRS cells (large cells, either with a large poliploid nucleus or multinucleated) in the appropriate cellular milieu. At the current time, definitive diagnosis can be made either by morphologic assessment alone, or by morphologic assessment combined with immunohistochemical studies, thus diminishing the importance of the definitive identification of diagnostic HRS cells.

From an immunohistochemical point of view, HRS cells harbor an immunophenotype not similar to any kind of normal hematopoietic cells. They express proteins that are characteristic to several cell types, such us granulocytes (CD15), T lymphocytes (perforin, Notch-1), memory B cells (syndecan) and dendritic cells (fascin, CCL17) whereas typical markers of B cell lineage are not observed with the exception of PAX5 (paired box 5, B cell specific trasncription factor) which is expressed in almost all cases. Most of the cases (around 89%) overexpress CD30 which is commonly used as a diagnostic marker (Figure1.1). In contrast, tumoral cells for the NLPHL type express specific B cells markers (immunoglobulin, CD19, CD20, CD79A and the B cell specific transcription factors PAX5, OCT2 and BOB1), including some for the germinal center markers (BCL6, AICDA).(Jaffe et al., 2001; Pileri et al., 2002).

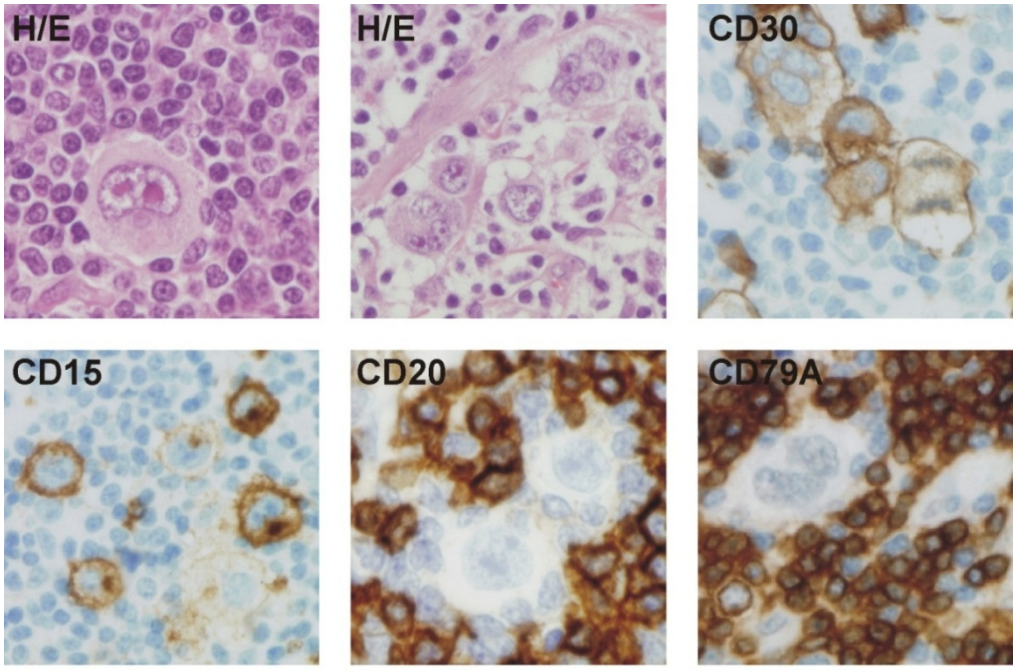


Figure 1.1 Immunophenotype of HRS cells. Hematoxilin-eosin stained paraffin sections (H/E; either mononucleated or multinucleated cells) and immunohistochemical markers for CD30, CD15 and B cell markers (CD20 and CD79A).As commented, HRS cells don't express B specific markers (CD20 and CD79).

1.2.3 Biology of classical Hodgkin lymphoma.

This lymphoid malignancy is considered to be a monoclonal proliferation of the characteristic HRS cells. It has a defective B-cell inmunophenotype and a characteristic paucity of the HRS cells within the tumor. The scarcity of HRS cells (5-10%) accounting for the majority of cases has limited molecular studies by conventional techniques, thus being cellular origin of HRS cells an intriguing issue during a long time. However, the B-cell origin and the monoclonality of the HRS cells have been clearly established in the last two decades by several studies (Kanzler et al., 1996a; Kuppers et al., 1994; Marafioti et al., 2000)

1.2.3.1 Cellular origin of HRS cells.

The origin of the HRS cells was finally clarified when microdisecction and single-cell PCR methods were applied to individual HRS cells to analyze them for immunoglobulin or TCR rearrangements. This rearrangements are highly diverse and specific for either B or T cells, thus being ideal markers for the elucidation of either B- or T-cell origin (Rajewsky, 1996). HRS-Single cells studies allowed the identification of clonal immunoglobulin gene rearrangements in nearly all cHL cases, thus establishing their clonality and B-cell lineage origin(Kanzler et al., 1996b; Kuppers et al., 1994; Marafioti et al., 2000). Nearly all cases carried somatically

mutated immunoglobulin variable (V) gene rearrangements whereas just a few cases presented clonal TCR rearrangements (Muschen et al., 2000).

Also, in nodular lymphocyte predominant Hodgkin lymphoma (NLPHL) a B-cell origin was demonstrated by the presence of clonal immunoglobulin gene rearrangements. In many of the cases, intraclonal V-gene diversification was observed, indicating ongoing hypermutation activity during clonal expansion. As active hypermutation is a hallmark of germinal center B-cells (Kuppers et al., 1993), this fact indicated a germinal-center B-cell origin of the tumoral cells in NLPHL.

1.2.3.2 Loss of the B-cell phenotype of HRS cells.

Despite the presence of a B cell genotype in HRS cells (immunoglobulin rearrangements and hypermutations), immunohistochemical studies showed down-regulation of B cell transcriptional program (numerous B-cell markers ,such as CD20,CD79, Oct-2,Bo and Pu.1). The lack of expression of transcription factors regulating immunoglobulin gene transcription seen in almost all of cHL cases (Oct-2, Bo and Pu.1) could explain the absence or the very low level of immunoglobulin gene transcription observed in HRS cells.(Re et al., 2001; Stein et al., 2001; Torlakovic et al., 2001)

However, there is now an indication that the immunoglobulin loci could also be silenced by epigenetic mechanisms (DNA methylation) (Ushmorov et al., 2006; Ushmorov et al., 2004). In fact, this epigenetic silencing by promoter methylation not only affects the immunoglobulin genes but also other B-cell specific genes, such as Pu.1, CD19, CD79b, and Bob1. Other mechanisms may contribute to the loss of B-cell phenotype: immunoglobulin rearrangements in the VDJ region, thus resulting in truncated proteins (25% out of the total cases)(Kanzler et al., 1996b); and mutations in regulatory area of the genes (despite being rare).

Interestingly, gene expression profiling studies of HL derived cell lines (HDLM2, L1236, L428, KMH2) and normal mature B cells revealed that B-cell-specific genes downregulation is more general than previously thought(Schwering et al., 2003b). Decreased mRNA levels from the vast majority of known B-cell lineage-specific genes were identified, including components of the BCR pathway (Blk, Syk, and SLP-65), transcription factors (Pu.B, A-myb, Spi-B),and surface markers (CD37, CD53). Taking into account that normal B cells are selected for expression of a BCR throughout their life, it might be a connection between the proposed origin of HRS cells from crippled germinal-center B cells and their global loss of B-cell

phenotype (Kuppers et al, 1998). A germinal-center B cell that, because of unfavorable immunoglobulin V-gene mutations fails to express a high–affinity BCR and would undergo apoptosis in normal conditions, could escape the selection for expression of a (high affinity) BCR by losing its B-cell identity. However, this loss of identity could also reflect an involvement of transforming events in the cHL pathogenesis rendering cells independent from the normal survival signals of B-cells and enabling them to adopt another phenotype.

Table 1.2 Genetic and phenotypic characteristics of HRS cells in relation to their cellular origin

Feature	
Somatically mutated Ig V genes	Yes
Destructive somatic mutations*	Yes (25%)
Ongoing somatic mutations	No
Proposed cellular origin	Pre-apoptotic germinal-center B cell
Expression of a BCR	No
Expression of B-cell-specific transcription factors (Oct-2,Bob1, Pu1)	Very rarely
Expression of B-lineage commitment factor Pax-5	Yes (low level)
Expression of B-cell surface molecules (CD20, CD79)	No or rarely
Expression of germinal center B-cell markers (BCL6, AID)	Rarely
Expression of plasma cell markers (Mum1, CD38)	Often
Expression of molecules involved in antigen presentation (MHC class II, CD40, CD80, CD86)	Yes
Expression of markers for non B-cells (TARC, granzyme B, CD15, CD3)	Occasionally/ frequently
Rare cases with T-cell origin	Yes (<2%)

Note : Ig_ immunoglobulin; BCR_ B cell receptor;MHC_ Major histocompatibility complex

****crippling mutations that can be easily identified (nonsense mutations)***

Finally, it must be pointed out that plasma cell origin of HRS cells has been already proposed (Carbone et al., 1998; Carbone et al., 1999) because down-regulation of many B-cell markers is also a feature of post-germinal-center (plasma) cells and expression of the plasma cell marker CD138 was reported in HRS cells. However, CD138-positivity of HRS cells could not be confirmed in other studies and there are major differences in key features of their gene expression patterns (lack of expression of immunoglobulin and plasma cell marker BLIMP1 and retained expression of Pax-5 and MHC Class II). In conclusion, despite derived from germinal

center B lymphocytes, HRS cells present a defective B cell transcription program (Hertel et al., 2002; Jundt et al., 2002; Schwering et al., 2003b) (Table 1.2).

1.2.3.3. Apoptosis

In B normal lymphocytes, the absence of functional immunoglobulins would make them to undergo apoptosis (FAS mediated) in germinal center; however in HRS cells multiple signaling pathways are aberrantly activated and many other alterations have been described to contribute to their proliferation and survival.

A. FAS gene mutations; either somatic mutations (Muschen et al., 2000) (less frequent than in other germinal center derived lymphomas)or germ line mutations(Straus et al., 2001).

B.c-FLIP expression;(observed in most cases)(Thomas et al., 2002). There is now a strong evidence that the resistance of HRS cells against CD95-mediated apoptosis induction is indeed due to strong expression of cFLIP, a negative regulator of CD95 signaling (Mathas et al., 2004).

C. Deregulated expression levels of the apoptosis protein regulators: IAPs (XIAP, cIAP1, cIAP2, survivin) and BCL2 proteins. (Garcia et al., 2003; Kashkar et al., 2003).

D. Constitutive activation of signaling pathways in HRS cells:

- NF-κB, a heterodimeric key transcription factor involved in many inflammatory and immunologic processes, was first described in 1996 to be constitutively active in Hodgkin lymphoma cell lines (Bargou et al., 1996). It has been shown to mediate both proliferative and antiapoptotic gene transcription programmes in HRS cells.

There are various explanations for NF-κB upregulation in HRS cells: c-REL amplification (Martin-Subero et al., 2002); activation of cell-surface receptors such us CD30, CD40, RANK or Notch1 (Cabannes et al., 1999; Horie et al., 2002; Joos et al., 2002); expression of viral latent membrane proteins LMP1 and LMP2a in EBV-positive cases; or loss of function mutations of the NF-κB suppressor IκBα or IκBε.

- JAK/STAT. In HRS cells STAT3, STAT5 and STAT6 are expressed and constitutively activated. Upon cytokine activation, Jak kinases become activated, and these can then phosporilate members of the Stat family transcription factors. Phosphorylated STATs can enter the nucleus and activate transcription of target genes. Among factors contributing to this activation,

recurrent amplifications of JAK2 locus and autocrine stimulation loop through IL-13 and IL-13R can be mentioned.(Joos et al., 2000; Kapp et al., 1999).

-There are also other signaling pathways frequently activated in HRS cells such as PI3K/AKT, MEK/ERK and AP1 pathways(Dutton et al., 2005; Mathas et al., 2002; Zheng et al., 2003).

1.2.3.4. Role of Epstein Barr Virus (EBV) in Hodgkin lymphoma.

In about 40-50% of cHL cases EBV genome has been detected (Weiss et al., 1987), with a higher incidence in mixed cellularity histological subtype, elderly patients , childhood and cases diagnosed in developing countries. HRS cells from EBV positive cases show a viral gene expression pattern (LMP1, LMP2A y EBNA1) that persists throughout the course of the disease and is characteristic of a latent infection ("type2 latency") (Herbst et al., 1991). EBV plays a central role in cHL pathogenesis by rescuing BCR deficient, pre-apoptotic germinal center B cells from apoptosis, thus rendering to the HRS cells independence of survival signals normally supplied by the BCR (Bechtel et al., 2005). LMP1 and LMP2A contribute to the constitutive activation of CD40/NF-κB and B cell receptor signaling pathways and thereby mimic survival signals provided by either T lymphocytes or follicular dendritic cells.

1.2.3.5 Other genes

It has not been identified a consistent mutation pattern neither in oncogenes nor in tumor suppressor genes, being mutations in p53, NRAS or BCL10 very rare or inexistent. Numerical and structural chromosome abnormalities are detected in most cases of cHL in the HRS cell population (Falzetti et al., 1999; Joos et al., 2000; Martin-Subero et al., 2003). Thus, it is known from classical cytogenetic studies the presence of several recurrent chromosomal breakpoints affecting the immunoglobin loci (Martin-Subero et al., 2006). Additionally, known translocation partners such as BCL2 (Szymanowska et al., 2008), BCL6, MYC and MALT1 can be also found in HRS cells.

Applying comparative genomic hybridization (CGH), recurrent genomic imbalances were found. Among them, 2p13-16 (REL); 9p23-24 (JAK) and 12q14 (MDM2) were involved most frequently (Joos et al., 2000; Joos et al., 2002; Kupper et al., 2001). In a high percentage of cases, HRS cells are characterized by deregulation or abnormal expression of multiple proteins implicated in the regulation of cell cycle, such as CDK1, CDK2, cyclin A, B1 and E and MDM2 (Figure 1.2). In addition, it is frequently found the lost of expression of tumor

suppressor genes (p16, p15, p27) thus leading to an abnormal integrity of S, G1 and G2/M mitotic checkpoints (Garcia et al., 2003; Garcia et al., 1999).

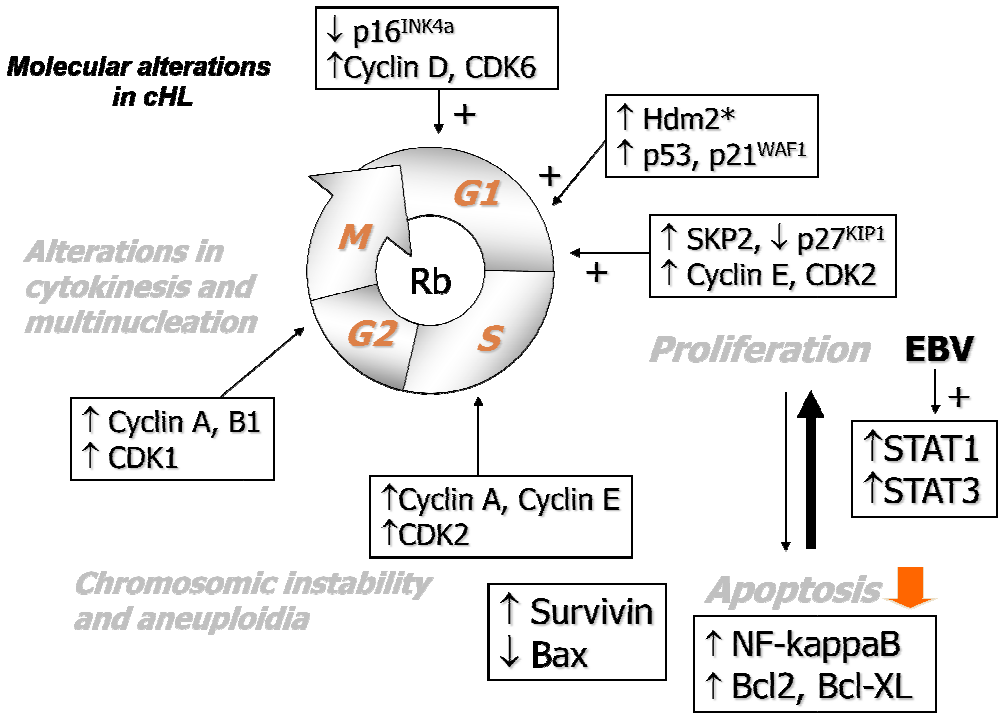


Figure 1.2 Molecular alterations in cHL (reproduced from García JF et al, *2003)*

1.2.4 Hodgkin lymphoma microenvironment.

Most of the tissue affected by cHL is comprised of a complex inflammatory infiltrate. Multiple evidences suggest an active role of this "microenvironment" in the tumor biology, through a bidirectional signaling between the inflammatory cells that promotes proliferation and survival of neoplastic cells being also reciprocally attracted and modulated by tumoral cells. A complex network of cytokines and cell contact-mediated interactions between tumor and inflammatory cells are thought to be implicated. It may also rescue HRS cells from the proapoptotic state arising from their characteristic BCR deficiency by providing alternative survival signals.

The cellular composition and biological role of tumoral microenvironment is determined by the already mentioned complex crosstalk between HRS cells and the infiltrating cells (Fig 1.3). Although much remains unsolved, some clear interactions have been observed with cytokines, and cell adhesion molecules largely accounting for the attraction of various infiltrating cells. Several chemokines secreted either for HRS cells (CCL17, CCL22) or

inflammatory cells (CXCL10, CCL3, CCL4, CCL5) induce migration and activation of T lymphocytes, monocytes and other cell populations. Likewise, HRS cells express receptors (CCR7, CXCR4) for chemokines already secreted by cells from the microenvironment (CCL19/21 and CXCL12) (Skinnider and Mak, 2002).

HRS cells constitute a small percentage of the tumoral mass, which presents a characteristic infiltrate of inflammatory cells comprised mainly of activated T helper 2 (Th2) - like lymphocytes, although these neither produce the Th2 cytokines IL-4 and IL-13 nor express functional IL-2 receptor. However, they express CTLA-4 (a co-stimulatory molecule expressed in regulatory T cells), are anergic (hyporesponsive to stimulation by mitogens and antigens) and can inhibit Th-cell responses.

T cells, predominantly $CD4^+$, are a significant component of the infiltrate enriched in regulatory T cells (Marshall et al., 2004). These cells are not tumor-antigen-specific T cells, and may be recruited non-specifically by CCL17, CCL22 and CCL11, produced by HL tissues. They secrete cytokines like IL 10 and TGF-β, thus creating an immunosuppressive environment that explains the lack of an effective anti-tumor response to antigens expressed by HRS cells (like EBV proteins). CD8+ cells represent a small proportion of the T-cell infiltrate and are typically not in close contact with HRS cells and, similarly to $CD4^+$ T cells, do not seem to be directed against a common antigen. These T cells express ligand receptors (CD40L, CD27, CD4, LFA-1) for a variety of surface molecules expressed by HRS cells (CD40, CD70, HLA-II, ICAM1) (Poppema and van den Berg, 2000), underlining the presence of the mentioned crosstalk between HL and T cells.

Other cell populations present are histiocytes, plasma cells, esosinophils, mast cells, neutrophils, fibroblasts and other stroma cells such as macrophages (currently becoming as a hot topic in cHL biology, as it will be discussed in this work).

• **Mast cells** are present in mostly all LH and express CD30L being able to stimulate HRS cell proliferation *in vitro* (Molin et al., 2001).

• **Eosinophils** also express ligands for CD30 and CD40 (receptors expressed by HRS), providing proliferative and antiapoptotic signals for HRS cells. Furthermore, eosinophils produce TGF- and may therefore be involved in fibroblast activation and fibrosis.

• **Fibrosis** is mainly observed in nodular sclerosis subtype and can be due to the unbalanced production of pro-fibrogenic cytokines like IL13, TGF-β, b-FGF and CD40. Primary fibroblasts from HL also produce growth factors for HRS cells, such as IL6 and IL7(Aldinucci et al., 2004).

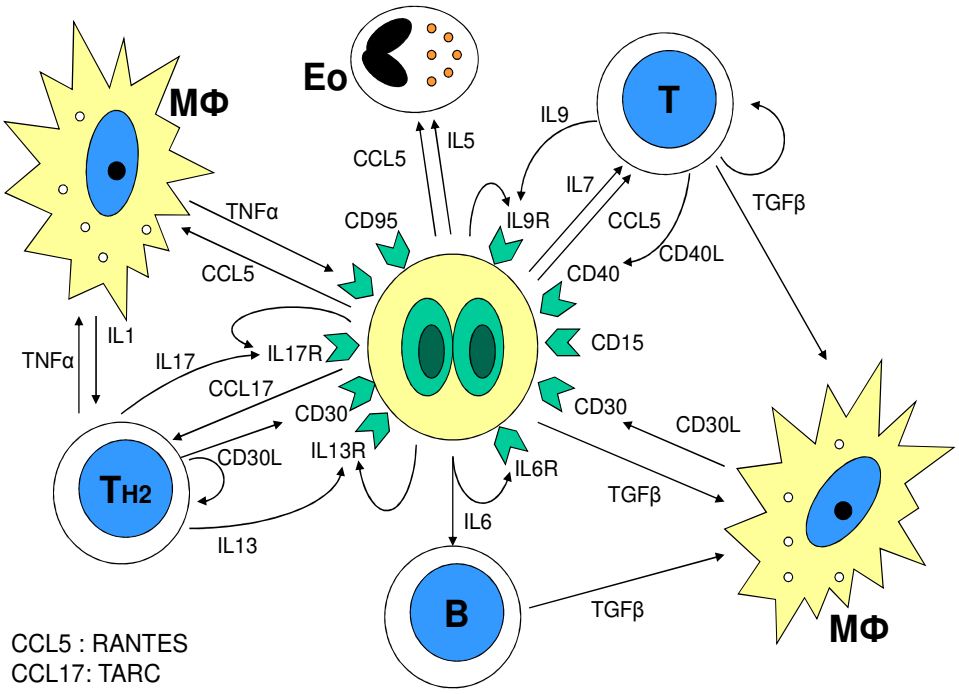


Fig 1.3 Chemical crosstalk between HRS cells and the microenvironment. B_ Bcell; T_T-cells; Th2_T helper 2 cells, MΦ_ Macrophage; Eo_ eosinophil.

A detailed description of all the interactions found is out of the scope of this thesis work , thus just a general overview has been given in order to remark the complexity and thus important role of the cHL microenvironment for the biology and pathogenesis of the cHL. Interestingly, now there are so many evidences pointing out that the microenvironment is also playing an important role in the clinical outcome of cHL and some findings regarding the function of macrophages and different T-cell subsets are proving to be of interest. In this context, a complete analysis of cHL should pay attention to both cellular components.

1.2.5 In vitro models of HL.

The scarcity of HRS cells within the infiltrate tissue has limited molecular studies and the progress of Hodgkin´s lymphoma research for a long time. The establishment of HL derived cell lines represented a very important step, as these cell lines were used for the successful

discovery of HRS cell-associated antigens (CD30, CD70) as well as for the analysis of multiple signaling pathways in HRS cells. Most remarkably, the wide use of gene expression techniques for cHL cell lines profiling provided new clues regarding pathogenic mechanisms of the disease,(Kuppers et al., 2003; Schwering et al., 2003a);(Cossman et al., 1999; Kapp et al., 1999).

Most attempts to establish permanently growing cell lines derived from a pure HRS cell population failed, but finally more than 10 bona fide Hodgkin lymphoma cell lines have been established, being 5 of them the most widely used and accepted for molecular studies in cHL (Table 1.3).

Table 1.3 Immunophenotype and genotype of the most widely used Hodgkin lymphoma cell lines

Cell line	CD30	CD15	B Ag	T-Ag	Rearranged Ig or TCR gene	BCR or TCR expression	EBV
L428	+	+	-	-	IgH,L	-	-
L540	+	+	-	-	TCRαβδ	-	-
L1236	+	+	-	-	IgH,L	-	-
HDLM2	+	+	-	-	TCRαβδ	-	-
KMH2	+	+	-	-	IgH,L	-	-

B-Ag, B Antigens; T-Ag, T Antigens

cHL derived cell lines constitute a good approach for cHL *in vitro* studies and elucidation of HRS cells biology as presenting a close resemblance to tumoral cells in immunophenotype. In fact, L1236 cell line has been proven to derive directly from HRS cells.

1.3 Current challenges in classical Hodgkin lymphoma

Although cHL is a curable tumor, an important fraction of patients in advanced stages are refractory to standard chemotherapy regimens based on adryamicin and other variants. Despite major advances, treatment approaches for cHL patients are rather uniform without taking into account this clinical problem and approximately 20-30% may relapse or eventually die due to progressive disease or complications of the therapy (Bonadonna et al., 2005; Canellos and Niedzwiecki, 2002). Patients with advanced disease and clinical indicators of poor prognosis, and those with disease that persists despite optimized primary treatment, may need intensified chemotherapy (Connors, 2005) while others could benefit from reduced treatment approaches (Hasenclever and Diehl, 1998).

1.3.1 Prognostic factors, definition and end-points.

Prognostic factors are variables measured in individual patients that offer a partial explanation of the heterogeneity observed in the outcome of a given disease, important in patient care for selecting management. Although we cannot prognosticate exactly for individual patients, we can make statements of probability that will be more accurate for groups of patients than for individuals.

Prognostic factors can be divided into tumor-related factors (directly related to tumor characteristics such as tumor pathology, biology, extension); host related factors (age and gender) and environmental-related factors (factors outside the patient).

Different outcomes may be of interest in analyses of prognostic factors, including overall survival (OS), Disease specific survival (DSS), Progression free survival (PFS), Failure free survival (FFS) and duration of response. For each of these end-points it must be clearly specified the starting point in time for the analysis and the clinical characteristics of events and censoring. Regarding this work, it has been mainly focused on directly tumor related factors (either related to HRS cells or microenvironment), looking for new clues and biological processes underlying cHL pathogenesis (with special attention of treatment response, FFS and OS events).

1.3.2 Clinical aspects and prognostic factors in Hodgkin lymphoma.

The most common clinical presentation of Hodgkin lymphoma is the presence of an enlarged lymph node in cervical region. Another common presentation is the discovery of a mediastinal mass on imaging, and approximately 60% of patients have mediastinal involvement at diagnosis with 50% of the patients presenting localized disease (early stages I-II). Mediastinal tumor masses are frequent, especially in Nodular sclerosis subtype (NS) and around 25-40% of patients present systemic symptoms such as fever, night sweats and weight loss (known as “B symptoms”). Staging of the disease is based on the Ann Harbor staging system modified by Costwolds (Table 1.4).

Current predictive systems are based on clinical and analytical parameters such as the International Prognostic Score (IPS) (Hasenclever and Diehl, 1998) that still fails to accurately identify a significant fraction of patients with very short failure free survival (FFS) (Gobbi et al., 2001). Thus, outcomes in cHL patients have been classically prognosticate using recognized

clinical variables like bulk of the disease , patient age, number of nodal sites and erythrocyte sedimentation or the ones present in the IPS .

Table 1.4 Ann Arbor staging classification for Hodgkin lymphomas

Stage I	Involvement of a single lymph node region (I) or a single extra lymphatic organ or site (IE)
Stage II	Involvement of two or more lymph node regions or lymphatic structures on the same side of the diaphragm alone (II) or with involvement of limited, contiguous extralymphatic organ or tissue (IIE)
Stage III	Involvement of lymph node regions on both sides of the diaphragm (III) which may include the spleen (IIIS) or limited, contiguous extralymphatic organ or site (IIIE) or both (IIES)
Stage IV	Diffuse or disseminated foci of involvement of one or more extralymphatic organs or tissues, with or without associated lymphatic involvement.

This prognostic mode (IPS), the most classic and widely used prognostic index for advanced cHL, consists of seven clinical variables, including serum albumin less than 4g/dL, hemoglobin less than 10.5 g/dL, male gender, age 45 years or higher, stage IV disease, white blood cell count at least 15,000/mm^3, and absolute lymphocyte count less than 600/mm^3 (or less than 8% of the total white cell count). With improvements in therapy, increasing dose intensity, early recognition of toxicity, and enhanced supportive care, current outcomes in patients may be superior to the ones predicted by IPS. In fact, recent studies show that IPS doesn´t adequately risk-stratify subgroups of patients with excellent, intermediate or poor outcomes, because patients with 0 to 3 risk factors demonstrated similar 5-year FFS rates ranging from 80% to 86% and patients with 4 or more risk factors had similar outcomes ((Blum, 2010).

Therefore, a re-evaluation of the IPS in a larger population of patients treated at multiple centers is needed to confirm these findings and to determine if the IPS in the current era segregates patients into clearly defined risk groups suitable for risk-adapted therapeutic strategies designed to minimize treatment for patients with the most favorable outcomes and intensify or utilize novel therapy for those ones that most likely will relapse.

1.3.2. Biological prognostic factors.

In the background previously described, we could hopefully think in the future; in a greater understanding of cHL´s biology that could determine and individualize patient therapy. Biologic factors underlying this malignancy may serve as prognostic markers and therapeutic targets. Nevertheless, numerous biologic markers have been recently identified including surface receptors, intracellular proteins, cytokines, and genetic abnormalities comprised of

amplifications, deletions, epigenetic silencing, or alterations in microRNA in the tumoral HRS cells or their surrounding reactive infiltrate.

However, the identification of biological markers has been problematic due to several reasons:

- 1. Lack of large and confirmatory prospective trials.
- 2. Failure in the reproducibility and feasibility of the assays. Many of the techniques used to detect the markers are not easily reproducible in other laboratories.
- 3. Failure to improve upon already clinical risk factors (IPS or others).

Technical advances such us DNA microarray-based gene and microRNA expression profiling, RT-PCR platforms, comparative genomic hibridyization (CGH) , SNPs arrays, microdisecction and immunohistochemical indexes have led to new discoveries and there is no doubt that their integration with the classical parameters will further improve the understanding of disease, and likely the management of patients. Recently, some markers have been recognized as putatively useful prognostic factors in HL as described below.

- **MMP11**

MMP1 is a marker of matrix metallopeptidases expressed by tumor-associated macrophages, found to be associated with treatment failure by both gene expression profiling and immunohistochemistry. This gene had a good discriminate power with respect to patient outcome and, more than 1% MMP1 staining by immunohistochemistry predicted a reduced progression-free survival in both univariate ($p=0.008$) and multivariate ($p=0.009$) analyses. However, this marker stains many type cells in classical HL specimens (endothelial cells, macrophages) thus being very difficult to determine the biologic mechanism responsible for this protein´s association with outcome.

- **CD20**

CD20 is a gene that encodes a B-lymphocyte cell-surface antigen has been evaluated by immunohistochemistry in several studies (Chetaille et al., 2009; Steidl et al.) finding a positive association of an increased number of tumor-infiltrating CD20+ cells with either prolonged progression-free ($p=0.02$) and disease-free ($p=0.02$) survivals. Similar findings, in which a high number of CD20+ cells in the inflammatory infiltrated correlated with improved event-free and OS have been reported by Chetaille (Chetaille et al., 2009). No correlation with the number of

CD20+ HRS cells and patient outcome was found neither by the previous cited Steidl´s study nor other one of 598 patients(Rassidakis et al., 2002b). However, a small study performed by Portlock et al (Portlock et al., 2004)in a set of 248 samples showed that both time to treatment and OS were significantly lower in patients treated with ABVD with CD20+ cells. Interestingly, in a fourth study, greater than 10% CD20+ HRS cells was associated with inferior patient outcomes in patients treated from 1974 to 1980 whereas failed to impact failure-free survival in patients treated from 1981 to 1999(Tzankov et al., 2003). Thus, the prognostic relevance of CD20 in HL might depend on its location and other factors (such as therapeutic modality, and the criteria used to define CD20 positivity). As it happens to many other identified markers, these results need further validation in prospective studies to determine their reproducibility.

- **CD68**

CD68 is a macrophage marker. This gene encodes a 110-kD trans-membrane glycoprotein that is highly expressed by human monocytes and tissue macrophages. It is a member of the lysosomal/endosomal-associated membrane glycoprotein (LAMP) family. In the previously mentioned study (Steidl et al, 2010) the initial finding of a tumor infiltrating macrophages gene signature associated with treatment failure prompted to an additional immunohistochemical validation of CD68 to assess if this marker could be easily studied and expression levels correlated with outcome. In univariate analysis, the presence of a high number of infiltrating CD68+ cells was correlated with reduced progression free (p=0.003) and disease-free (p=0.003) survivals. Noteworthy, this marker was superior to the IPS in predicting disease free-survival in multivariate analysis (p=0.003 vs p=0.03) and indeed is being currently discussed in so many works validating its predictive value (Kamper et al.).

Similar findings correlating the presence of tumor-infiltrating macrophages with adverse patient outcomes have been also been described in follicular non-Hodgkin's lymphoma(Dave et al., 2004; Farinha et al., 2005) , suggesting that interactions between the malignant lymphoma cells and the microenvironment are critical to the pathogenesis and progression of these diseases. However, the use of CD68 alone as a marker is being currently discussed in so many works some of them with contradictory results (Azambuja et al., 2011)

- CD163.

CD163 is a monocyte/macrophage specific protein that seems to be involved in anti-inflammatory functions marker that has been found to have similar correlations with outcome

as for CD68 in several studies (Steidl et al., 2010) (Kamper et al., 2011) and is considered to be more specific and thus more accurate for macrophage identification than CD68.

- Others. A variety of other prognostic markers have been described in HL.

Increased serum levels of soluble CD30 (sCD30)(Zanotti et al., 2002), interleukin (IL)-10, B-cell-activating factor (BAFF), tumor necrosis factor-alpha (TNFα), thymus and activation regulated chemokine (TARC) and vascular endothelial growth factor (VEGF) have all been correlated with adverse patient outcomes (Casasnovas et al., 2007). In several studies, combining sCD30 level with the IPS improved its predictive value (Ma et al., 2009; Renna et al., 2009).

-BCL2. The relationship between BCL2 expression and patient outcome in HL is a controversial issue with studies reporting that BCL2 expression was independently associated with reduced failure free survival in addition to other clinical prognostic variables (age over 45, stage IV disease, low albumin and elevated lactate dehydrogenase)(Rassidakis et al., 2002a) whereas other have not demonstrated the same correlation (Montalban et al., 2004).

Similarly to BCL2, other markers like p21 and p53 have been associated with treatment outcomes (Sup et al., 2005) with controversial results, although more studies suggest a prognostic role for BCL2 than for p53 and p21.

-T- cells. The presence of T cells in the inflammatory infiltrate also seems to influence patient outcomes and several studies have examined the ratio of regulatory T cells ($CD4^+$ $CD25^+$ and FOXP3 expression) and cytotoxic T cells ($TIA\text{-}1^+$ or granzyme B expression) in relationship to overall and disease free survival rates (Alvaro et al., 2005). The ratio of regulatory $FOXP3^+$ T cells to cytotoxic T/NK (natural killer cells) with granzyme B expression independently predicted patient survival.

Noteworthy, these results showed that not only biological markers from the HRS cells are prognostic in HL; also the composition of the surrounding cellular matrix and microenvironment is playing a key role in the treatment outcome process, thus setting the basis for the interest of this thesis work:

Elucidation of biological processes underlying treatment outcome in both tumoral components.

1.4 Hodgkin Lymphoma, expression analyses of whole tissue sections. Rationale.

Scientific advances in the field of genetics, gene and micro-RNA expression profiling have revolutionized the concept of molecular biology, patient management and treatment. Analysis of differential gene-expression patterns across thousands of biological samples in a single experiment, and extrapolation of these data to answer pertinent questions such as those relating to tumor classification, metastatic potential and treatment outcome , can help define clinical decisions and the best therapeutic regimens for particular subgroups of patients.

Gene-expression profiling has been used to study all kind of diseases but they have been predominantly focused on the study of cancer. Microarray analysis has made possible to identify groups of genes according to their expression pattern or “genetic signature” for the most common types of neoplastic malignancies including lymphomas. These studies have demonstrated the ability to identify pathogenic mechanisms, new molecular targets and biological processes involved in lymphomagenesis.

The main strategies to identify categories of tumors by gene-expression profiling are unsupervised and supervised classification methods. The main difference between these two methods is that for supervised classification clinical or pathological information is used to find correlations with gene expression patterns, whereas with supervised methods tumors are grouped on the basis of gene expression pattern independently of the clinicopathological status. Thus, unsupervised methods are useful when looking for genes responsible for cellular processes and specific properties from tumor and non-tumor cells. However, to find gene expression patterns that can predict the clinical behavior of tumors, it is more appropriate to use a supervised classification able to distinguish on the basis of predefined clinical and pathological information, because classes are already known (T-test, Anova and other traditional statistical test can be used for this aim).

1.4.1 Gene signatures in classical Hodgkin lymphoma

In Hodgkin lymphoma, gene expression profiling of cHL cell lines has given new clues for the pathogenic mechanisms underlying this malignancy, such as the complex crosstalk between HRS cells and microenvironment previously mentioned and other important discoveries such as the inactivation of B-cell receptor and differentiation program (Kuppers et al., 1994).

In addition, gene expression analyses using whole tumor sections helped into the identification of different molecular subtypes of cHL disease (Devilard et al., 2002) and the elucidation of gene patterns and processes associated to treatment response (outcome). Interestingly, by unsupervised methods, Devilard identified three main molecular subgroups of cHL tumors with respect to histology and clinical outcome (response to therapy and survival), thus establishing the first molecular taxonomy of cHL in correlation with clinical variables and suggesting the possibility of improving the current prognostic signification. This study found that samples from all bad outcome patients clustered in one group whereas the two other groups contained most good outcome (related to NS subtype) cases and revealed the upregulation of genes implicated in fibroblast activation, angiogenesis, extracellular matrix remodeling, cell proliferation and downregulation of tumor suppressor genes.

Other works have also identified specific gene patterns by comparing data from two completely opposed biological situations (treatment responders vs non-responders) (Sanchez-Aguilera et al., 2006) all of them pointing out the important role of genes involved in cell proliferation, cellular matrix remodeling and apoptosis in treatment outcome of patients. These initial studies based on DNA microarrays demonstrated their ability to define the interaction pathways between neoplastic and nonmalignant cells in cancer tissues. Taking into account the special composition of cHL, expression profiling on whole tissue sections can detect specific transcriptional patterns from both the tumoral cells and the non-tumoral microenvironment. During the past ten years, a number of approaches including microdissection have tried to resolve the variability in measurements derived from this heterogeneity. Another approach that has also been used in thesis work and is designated as virtual or *in silico* microdissection (Alizadeh et al., 2001), avoids the laborious and time-consuming step of anatomic microdisecction and consists of confronting the gene expression profiles of whole tissue samples to those of cell lines representative of different cell lineages, differentiation stages, or different signaling pathways. This strategy has been used in recent studies, thus aiming the analysis of whole tissue sections as a good alternative to identify genes that could ideally reveal as useful prognostic markers and be added to the already known prognostic factors.

1.4.2 MicroRNA signatures in classical Hodgkin lymphoma.

MicroRNAs (miRNA) are small (approximately 22 nucleotides length) non-coding RNA molecules that regulate gene expression at post-transcriptional level through messenger RNA interference by binding to complementary sequence at 3´ untranslated region (UTR) of target

genes. These molecules were described first time in 1993 by Ambros and colleagues in *Caenorhabditis elegans* and to date hundreds of miRNAs have been identified in other species, including viruses (such as EBV and others). They are encoded by intronic or intergenic DNA regions, first as large molecules that are cleaved and processed by an RNAase complex called Dicer to generate the finally mature miRNA that binds a protein complex called the RNA-induced silencing complex (RISC). Nucleotides 2 to 7 from the 5´ end of the miRNA, the seed region, are crucial for miRNA-target gene interaction (Grimson et al., 2007). If the miRNA has perfect or nearly perfect complementarity to the 3´UTR region, induces messenger RNA cleavage (Llave et al., 2002; Yekta et al., 2004) whereas when complementarity is not perfect translational silencing of the target gene occurs(Barrington et al., 2009).

MiRNA expression has been shown to be tissue specific as well as temporally regulated. They are emerging as important regulators of biological processes such as proliferation, differentiation , and they also act as oncogenes or tumor suppressor genes (Hammond, 2006, 2007; Kent and Mendell, 2006), thus suggesting an important role for miRNAs in human tumorigenesis. In fact, their role in cancer pathogenesis has been demonstrated in several malignancies, including lymphomas (Calin et al., 2004). Remarkably, the majority of miRNAs are found in cancer –associated genomic regions or in chromosome-fragile sites.

Some attempts have been made to elucidate specific miRNA signatures for the characteristic HRS cells and their microenvironment. Overexpression of miR-155 has been found in cHL cell lines by qRT-PCR analysis (Kluiver et al., 2005), being also differentially expressed between HL and Burkit lymphoma. MiRNA profiling of cHL derived cell lines has revealed HL-specific miRNAs that includes miR-17-92 cluster members which is frequently amplified in B cell lymphomas(Ota et al., 2004), and others miRNAs such as miR-16, miR-21, miR-24 and miR-155. A significant downregulation in HL compared to non - Hodgkin lymphoma has been observed only for miR-150 (Wong and Altekruse, 2009). At the same time, comparison of miRNA profiles of microdiseccted HRS cells from nine cHL patients , four common cell lines (HDLM2, L540, KMH2 and L1236) and CD77+ germinal center B-cells (used as normal counterparts) yielded a distinct cHL signature of 12 over- and three underexpressed miRNAs (Van Vlierberghe et al., 2009).

Despite these studies, the potential prognostic role of miRNA signatures in cHL remains unclear. Overexpression of miR-328 has been found in advanced cHL stages (III-IV stages) (Navarro et al., 2007)and low miR-135 levels are associated with higher relapse rates

and shorter disease free survival (DFS). Thus, it has not been yet well investigated the relevance of microRNA expression in cHL and their putative value for patient outcome prediction.

1.4.3 MicroRNA signatures and EBV

Additionally, several oncogenic viruses also express miRNAs although the function of most of these miRNAs is largely unknown. This is especially important in oncogenic viruses like EBV that expresses more than 20 miRNAs organized in two clusters within the EBV genome: one in the intronic region of the BART (BamH1-A Rightward transcript) gene and another in the untranslated regions (UTRs) of the BHRF1 (BamHI-H right reading frame) gene (Cai et al., 2006). These miRNAs are important for the biological cycle of the virus and control several viral genes such as BALF5, LMP1 and LMP2a (Barth et al., 2008; Lo et al., 2007; Lung et al., 2009). Some of these viral miRNAs promote survival of the host cell by targeting cellular transcripts. An example is ebv- BART5 which targets p53 up-regulated modulator of apoptosis (PUMA), thus rendering the infected cells less sensitivity to apoptosis (Choy et al., 2008). However, the exact role that EBV encoded miRNAs are playing in cHL is still and intriguing issue.

1.5 Translating gene and microRNA signatures to the clinical care.

Currently, microarray technology is in a transition phase where scientific information is beginning to guide clinical care decisions although they still need to demonstrate reliability and reproducibility. In this context, techniques such as real time quantitative polymerase chain reaction (qPCR, RT-qPCR) and immunohistochemistry are becoming necessary and widely used to validate hypothesis generated by microarray analysis. In addition, it is important to remark that microarray techniques present today important limitations, such as the complexity of the biostatistical analysis methods used, the high cost of technique and the sample acquisition among others.

1.5.1 The sample problem.

Microarray technology needs from frozen available samples that are not easily to find. In contrast, formalin fixed paraffin embedded samples (FFPE) are extensively available and easily accessible. Thus, a technique that could be applied to FFPE tissues would be ideal for profiling of genes without limitations of sample availability.

Quality of RNA from formalin fixed paraffin embedded tissues is not usually good due to the fixation and paraffination process (that causes severe degradation of the RNA). This RNA

is shorter (primarily with fragments of less than 300 bases in length), and extraction yields in general are lower in comparison with the ones from frozen tissues. However, RNA from FFPE tissues is still suitable for some applications. Interestingly, microarray hybridization of FFPE tissue samples has been already performed by some groups, although the reliability is still not enough clear (April et al., 2009; Grenert et al., 2011).

However, several studies have shown that RT-qPCR, the gold standard technique for microarray hits validation, can be used to quantify gene expression of RNA isolated from FFPE tissues even after laser capture microdisecction (LCM) (Koch et al., 2006; Specht et al., 2001) obtaining good quality data results. Moreover, in the past several years great advances have been made in the development of sensitive and quantitative gene expression assays using RT-PCR and FFPE tissue. Limitations due to the sample origin (low quality and amount of the extracted RNA) can be overcome in order to set assays based on this technique either for validation of microarray hits or the development of tests able to be applied to clinical routine.

1.5.2 Low density arrays (TLDAs)

As mentioned, researchers have been frustrated in the past by attempts to profile gene expression in archived cancer tissue because almost all archived samples were embedded in paraffin. In these samples RNA is degraded and cross-linked, thus significantly limiting the ability to examine changes in gene expression.

A technology PCR-based, sensitive enough, and reproducible enough was needed. Recently, a commercially available platform known as Taqman® Low density arrays (TLDAS, from Applied Biosystems AB) has been developed. They are 384-well micro fluidic cards that allow performing simultaneous RT-PCR reactions for sample without the need of robots of multi-channel pipettes to load samples. Taqman® Gene Expression Assays are preloaded into these low to medium-throughput arrays and consist on an ideal screening technology and a perfect tool for validating hits that come from microarrays. (See Material and Methods for further description and details). Their accuracy and reliability has been already demonstrated and are able to be applied to FFPE tissues as done in the first part of this work.

Additionally, the inclusion of novel pre-amplification systems (Taqman® PreAmp) designed to enhance expression analysis from limited amounts of RNA has successfully overcome the difficulties usually caused by FFPE tissue samples (Ciotti et al., 2009).

OBJECTIVES

Classical Hodgkin lymphoma (cHL) is assumed to be a curable tumor, but an important fraction of patients with advanced disease don´t respond favorably to the current standard chemotherapy regimens based on Adriamycin. The most widely used prognostic score is based on clinical and analytical parameters integrated in the International Prognostic Score (IPS), but this still fails to accurately identify at the moment of diagnosis, a significant fraction of patients with very poor prognosis. (Canellos GP, et al.1992).

Therefore, the identification of biomarkers that may consistently predict failure at diagnosis is essential for the recognition of patients at high risk of treatment failure, thereby to establish a more rational risk-adapted treatment strategy. In addition, gene expression studies of tumoral samples have led to the identification of gene signatures and biological processes associated to Hodgkin´s lymphoma pathogenesis and treatment response showing that outcome in cHL patients could be related with the both biological characteristic components of cHL tumors. However, no steps toward the application of this increasing knowledge in the development of assay tests feasible to be applied at the moment of diagnosis in a routine setting have been made. Hence, the objectives of this work were the following:

a. To identify gene expression signatures from tumoral (HRS cells) and their non-neoplastic microenvironment.

b. To elucidate biological processes and pathways related to patient outcomes.

c. To design an RT-PCR assay and derive a Molecular Risk Score using formalin fixed paraffin embedded (FFPE) cHL samples able to identify subgroups of patients with different risk of treatment failure.

In a similar manner, due to the emerging role of miRNAs in cancer, the objectives of the second part of this work were the following:

d. To investigate the relevance of microRNA expression in cHL and identify specific miRNA profiles from the tumoral cells and their non-neoplastic microenvironment.

e. To derive a miRNA signature associated to patient outcome, identifying miRNAs potentially useful for prognosis.

f. To explore the relationship among genes and microRNAs implicated in cHL pathogenesis and chemosensitivity.

MATERIAL AND METHODS

2.1 Patients and samples

2.1.1 Patient selection criteria

All patients samples included in this work fulfilled the same following stringent criteria: ages higher than 15 years, advanced cHL (Ann Harbor stage IV, III or IIB with Bulky masses), proven HIV-negative status, who had received a first line standard chemotherapy regimen that included ABVD (doxorubicin/ adryamicin, bleomycin, vinblastine and dacarbazine) or ABVD variants and with clinical information available about the achievement of complete remission (CR) and a follow-up of at least 12 months thereafter.

The protocol used to data collection comprised the following clinical and analytical parameters: birth date; gender; hospital of origin; diagnosis date; histological subtype; bone marrow infiltration; presence of Bulky masses; extra nodal and ganglionar affectation, B symptoms, lung involvement, stage, hemoglobin, albumin, white blood cell count, lymphopenia (lymphocytes<600), leukocytosis (leukocytes >15000), IPS, chemotherapy ,radiotherapy ,type of response (complete response (CR) , partial response (PR) or progression (P)) ,and data regarding last follow up (dead, alive with disease or without disease). Additionally, IPS was codified according to low versus High: IPS codified (low (< 3) versus High(≥ 3) .

All tissue samples consisted of representative specimens of pretreatment lymph node biopsies and were collected after revision by the institutional review board (IRB) of the participant's institutions from the Spanish Hodgkin Lymphoma Study Group and the MD Anderson Cancer Centre (Houston, Texas).

2.1.2 Patient sample series / Datasets

Different patient series were used in this work and consisted as follows:

2.1.2.1 Gene expression project

For the first project, an initial gene dataset previously generated in our laboratory by gene-expression profiling of frozen tumor samples from 29 patients was used. This dataset included 14 responders and 15 non-responders to the standard ABVD treatment and 5 cHL-derived cell lines (L428, L540, L1236, HDLM2 and HDMYZ). These data were used to define tumor and microenvironment databases and in subsequently bioinformatic analyses, looking for genes and pathways associated to treatment failure.

To validate specific genes identified by microarray data that were subsequently analyzed in FFPE samples, an independent series of 52 advanced cHL cases without overlapping with other datasets was used. Finally, the study leading to the Molecular Risk score (MRS) development enrolled 282 FFPE patient samples obtaining suitable analyzable data for 262 cases (92.90 percent) that were randomly split and assigned to either estimation or validation sets. A summary of the clinical characteristics of patients can be found in Table 2.1.

Differences in distribution of standard clinical parameters (age, gender, stage, IPS and outcome) between estimation and validation datasets were tested by Pearson chi-square test (χ^2 test) without any observed statistically significance (IPS values < 3 classified as low IPS; IPS values ≥ 3 classified as high IPS).

Table 2. 1. Clinical characteristics of patients with adequate RT-PCR profiles used for MRS development.

Characteristic		Estimation %	Validation %	Total %	p value (χ^2 test)
Age (yr)		**(N = 183)**	**(N = 79)**	**(N = 262)**	
	< 45	133 (72.67)	56 (70.88)	189 (72.14)	0.883
	≥ 45	50 (27.32)	23 (29.11)	73 (27.86)	
Gender		**(N = 183)**	**(N = 79)**	**(N = 262)**	
	Male	99 (54.10)	51 (64.55)	150 (57.25)	0.151
	Female	84 (45.90)	28 (35.44)	112 (42.75)	
Stage IV		**(N = 182)**	**(N = 79)**	**(N = 261)**	
	No	132 (72.52)	47 (59.49)	179	0.053
	Yes	50 (37.87)	32 (40.50)	82	
IPS		**(N = 182)**	**(N = 79)**	**(N = 261)**	
	< 3	109 (59.89)	41 (51.90)	150 (57.47)	0.288
	≥ 3	73 (40.11)	38 (48.10)	111 (42.53)	
Outcome		**(N = 183)**	**(N = 79)**	**(N = 262)**	
	F	132 (72.14)	57 (72.16)	189 (72.14)	1.000
	U	51 (27.86)	22 (27.84)	73 (27.86)	

2.1.2.2. MicroRNA expression project

Following a similar schema as the one described for gene expression project , for miRNA study an initial dataset was used for miRNA microarray hybridization and included frozen samples from 29 patients with advanced cHL, 24 responders (F) and 6 non-responders

(U) to ABVD treatment and 5 derived cell lines (L428, L540, L1236 , HDLM2 and HDMYZ). This set of samples had not overlap with the previous one used for the gene expression project and was used first to define tumor and microenvironment miRNA signatures and secondly to identify a set of miRNAS for further study in FFPE samples. As controls, 4 fresh frozen lymph nodes and 4 fresh frozen reactive tonsillectomy specimens were included.

A summary of the clinical characteristics of patients can be found in Table 2.2.

Table 2.2.Clinical characteristic of patients used for microRNA microarray hybridization (N=29)

Characteristic		**N**	**%**
Age	<45	24	82,76
	≥45	5	17,24
Gender	Male	14	48,28
	Female	15	51,72
Stage IV	No	21	72,41
	Yes	8	27,59
IPS code	<3	21	72,41
	≥3	8	27,59
Histologic Type	Nodular Sclerosis	22	75,86
	Mixed Cellularity	6	20,69
	Other	1	3,45
Outcome	F	23	79,31
	U	6	20,69

Additionally, 4 frozen cHL samples (without available clinical data) were used to perform laser capture microdisecction (LCM) in order to validate some miRNAs attributed to be expressed by HRS cells. (See Results II)

Finally, the second part of the study enrolled a set of 220 formalin fixed paraffin embedded samples (FFPE) recruited from different institutions of the Spanish Hodgkin lymphoma group and from MD Anderson Cancer Centre (Houston, TX) that were used as an entire dataset for further elucidation of miRNA prognostic signatures. A summary of the clinical characteristics of the FFPE series used for miRNA RT-qPCR is represented in Table 2.3.

Table2.3.Clinical characteristic of patients series included in RT_qPCR analyses for miRNA study (N=220)

Characteristic		N	%
Age	<45	163	74,09
	≥45	57	25,91
Gender	Male	116	52,73
	Female	104	47,27
Stage IV	NA	1	0,45
	No	149	67,73
	Yes	70	31,82
IPS code	NA	17	7,73
	<3	140	63,64
	≥3	63	28,64
Histologic Type	Nodular Sclerosis	158	71,82
	Mixed Cellularity	54	24,55
	Other	8	3,64
Outcome	F	143	65,00
	U	77	35,00

NA_ Not available data

2.1.2.3 TMA analysis

An independent series of 142 cHL cases available in our lab were analyzed by immunohistochemistry using tissue microarrays (TMAs) looking for specific genes identified as expressed either for HRS cells and their surrounding microenvironment.

2.2 Cell lines

cHL- derived cell lines were used in the different projects looking for specific signatures attributable to the HRS cells and are described in Table 2.4

Table 2.4 Description of cell lines used in this study

Cell line	Species	Description	Culture medium	Origin
HDLM-2	Human	Hodgkin lymphoma (T type)	RPMI-1640 + 20% FBS	DSMZ
HD-MYZ	Human	Hodgkin lymphoma (B type)	RPMI-1640 + 10% FBS	DSMZ
KM-H2	Human	Hidgkin lymphoma (B type)	RPMI-1640 + 10% FBS	DSMZ
L1236	Human	Hodgkin lymphoma(B type)	RPMI-1640 + 10% FBS	DSMZ
L428	Human	Hodgkin lymphoma (B type)	RPMI-1640 + 10% FBS	DSMZ
L540	Human	Hodgkin lymphoma (B type)	RPMI-1640 + 20% FBS	DSMZ

2.3 RNA extraction

2.3.1 Cell lines

2. 3.1.1 mRNA extraction

For mRNA extraction, 5 - 10 millions cells were stored in PBS media at -80ºC. After defreezing, they were disaggregated using a 1mL of Trizol reagent with the help of a syringe and a needle. The suspension was kept 10 minutes at room temperature an then, 200µl of chloroform were added. Samples were mixed thoroughly 15 seconds and maintained 10 additional minutes at room temperature. After 15 minutes of centrifugation at 8000g, the aqueous phase was recovered in a fresh tube. The same volume of isopropanol and 2µl of lineal acrilamide were added and samples were left at -20ºC overnight. Next day, samples were centrifuged at 12.000 g for 15 minutes at 4ºC. Pellets were washed twice with 70% cold ethanol and resuspended in RNAase free water. For a higher purity, the RNAeasy® Mini Kit from Qiagen (Hilden, Germany) was used following manufacturer instructions, including the DNAase treatment. Finally RNA was resuspended in 20-30µl of RNAase free water.

If RNA concentration or quality was not optimal, RNA was precipitated by addition of 1 µl of lineal acrylamide, 0.5 volumes of AcNH4 (ammonium acetate) 7.5 M and 2.5 volumes of ethanol overnight. After 20 minutes centrifugation, the pellet was washed twice with 70% cold ethanol, air-dried and resuspended in 20-30 µl of RNAse-free water. RNA was quantified using the Nanodrop 1000 device (ND-1000) (Nanodrop Technologies, Wilmington, DE, USA).

2.3.1.2 MiRNA extraction

RNA extraction protocol described in previous section is not suitable for small RNAs because small fractions are lost through the RNAeasy column. Thus, for analysis of miRNA expression, a different protocol was used. Since the detection of microRNA expression was performed using Taqman miRNA probes (Applied Biosystems) able to discriminate mature from inmature miRNAs without binding the microRNA gene, no DNA removal step was done**.**

First, 1- 2 millions cells were washed twice with 1X PBS and the pellet was disrupted by addition of 1 mL of Trizol reagent (Invitrogen) and 5 minutes incubation at room temperature. Trizol is a monophasic solution of phenol and guanidine isothiocyanate that maintains the integrity of the RNA, while disrupting cells and dissolving cell components. Then, 200 µl of chloroform were added, mixed and the sample was centrifuged at 1.200 rpm for 15 minutes at 4ºC. Aqueous phase was recovered and the same volume of isopropanol added for RNA

precipitation. The sample was then centrifuged and washed with 70% ethanol. Finally, pellets were dried and resuspended in 50 µL of RNAse-free water.

2.3.2 Tissues

2.3. 2.1 Frozen tissues.

Frozen tissues were cut with a cryostat to get 20-30 slides depending on the size of the frozen piece. The slides were homogenized using a Polytron homogenizer (Capitol scientific, Inc, Austin, TX).The homogenized samples were lyzed by addition of 1 mL of Trizol. Following steps are common to the protocol described in section 2.3.1.1 including column purification.

2.3.2. 2 FFPE tissues

Total mRNA was extracted from 3-10 FFPE sections of 10 µm thickness (depending on the size of the fixed piece).Paraffin sections were deparaffinized by incubation with 1 ml of xylene for 10 minutes at 65ºC. Samples were then centrifuged at 10,000 rpm for 10 minutes and the supernatant was removed. After an ethanol wash, tissue pellets were dried and resuspended in 200 µl of RNA lysis buffer containing 50 mM Tris, 0.5 mM EDTA (pH 8.0) and 10% sodium dodecylsulfate and incubated overnight at 65ºC with 10 µl of proteinase K (20 mg/ml, Qiagen Inc., Hilden, Germany). RNA was purified by phenol-chloroform extraction followed by precipitation in an equal volume of isopropanol in the presence of 1 µl of lineal acrylamide (Ambion, Austin, TX) at –20ºC. The RNA pellet was washed once in 70% ethanol, dried, and resuspended in 30 µl of of RNAse-free water. For genomic DNA removal, DNAse digestion was carried out by treating the total RNA with 5 µl DNaseI (1 U/µl, Epicentre® Biotechnologies, Madison, WI).

The final RNA concentration (A260:0.025) and purity (A260:A280 ratio) was measured using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA).

2.3.2.3 Microdissected tissues

RNA from microdiseccted samples was extracted from 800 - 1000 cells. Briefly, 350ul of RLT buffer from Quiagen (Hilden, Germany) and 3,5ul of ß -mercaptoetanol were added to the samples that were homogenized with a syringe. After, 150 µl of chloroform was added and samples were then centrifuged at 13,200 rpm for 10 minutes taking supernatant into a new eppendorf tube for RNA precipitation in 400 µl of isopropanol in the presence of 4 µl of

lineal acrylamide (Ambion, Austin, TX) at -20ºC overnight. The RNA pellet was washed twice in 70% ethanol, dried, and resuspended in 10-15 µl of RNAse-free water.

2.4 Retrotranscription (RT)

2.4.1 mRNA

First-strand cDNA was synthesized from total RNA using the High Capacity cDNA archive kit (Applied Biosystems, Foster City, CA), in 50 µl reactions using these cDNA samples for qPCR analyses by TLDA platforms according to the manufacturer's instructions. Briefly, 500ng of RNA were mixed with 25µl of RT Master Mix containing 5µl of 10x RT buffer and 10x Random Primers, 2 µl 25x dNTPs, 2.5 µl of 50U/µl MultiScribe RT, 9.5 µl of RNAse free water and 1 µl of RNAse inhibitor. The samples were put in a thermocycler with the following program:

Table 2.5 RT program conditions used for cDNA synthesis

Program	**Temperature**	**Time (min)**
HOLD	25ºC	10
HOLD	37ºC	120
HOLD	85ºC	5
HOLD	4ºC	∞

RT_ Retrotransciption; min_minutes

2.4.2 miRNA

Taqman® microRNA assays were used to specifically measure the mature form of the miRNAs. cDNA synthesis step was carried out with a specific primer for e This primer contains a specific sequence that anneals on the miRNA in the 3´end and a general sequence in the 5´end. This general sequence impedes the annealing of the immature forms. In the other hand, the primer yields a product much larger than the original miRNA allowing for an easier design of the PCR primers.

Retrotranscription was carried out using the TaqMan® MicroRNA Reverse Transcription Kit (Applied Biosystems). Briefly, 5µl of 250ng/µl RNA solution was mixed with 0.25 µl of 100mM dNTPs, 2µl of 50U/µ MultiScribe Reverse Transciptase, 3 µl of Reverse Transciption Buffer, 025 µl of 20 U/µl of RNAase inhibitor and 1µl of each one of the specific RT

primers (mixed in a pool) in a final reaction volume of 25µl. The samples were put in a thermocycler with the following program.

Table 2.6 RT conditions used for cDNA miRNA synthesis

Program	Temperature	Time (min)
HOLD	25ºC	10
HOLD	42ºC	50
HOLD	70ºC	15
HOLD	4ºC	∞

RT_ Retrotransciption; min_minutes

2.5 Preamplification

Taqman® RT-qPCR technique is based on the 5´nuclease activity of Taq DNA polymerase and involves cleavage of a specific fluorigenic hybridization probe that is flanked nby the PCR primers. Limited amounts of tissue and RNA and cDNA were available from FFPE samples, thus restricting the number of genes and cases that could be analyzed in some samples. In order to ensure high data quality a preamplification step was added to the sample processing.

The principle of TaqMan® PreAmp technique is to amplify target cDNA prior to RT-qPCR analysis. Briefly, cDNA is synthesized from total RNA using random priming; The cDNA for the specific target assays is then amplified by pre-amplification reaction using pooled gene-specific primers to increase the number of targeted copies. The pre-amplification product is diluted and finally analyzed by RT-qPCR using single assay

TaqMan® PreAmp Master Mix (T-PreAmp) (Applied Biosystems, Foster City, CA) that preamplifies small amounts of cDNA without introducing amplification bias to the sample was used. By incorporating a preamplification step into the sample preparation workflow, adverse sample effects due to the low quality of the RNA material were significantly reduced.

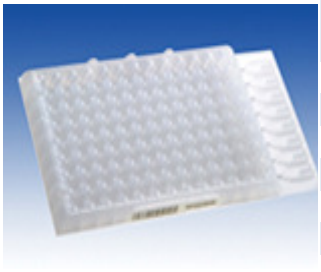
First, it was tested in an initial small set of 13 cases and 32 genes. Briefly, the pooled assays were diluted to a final concentration of one-fifth that of the PreAmp primer/assay pool. Initial experiments comparing the volume of PreAmp reaction recommended by the manufacturer (total PreAmp reaction volume: 50 µl) and a tenth of that volume (5 µl PreAmp reaction volume) gave identical PreAmp results, so we subsequently used the latter volume due to reasons of cost-effectiveness. The chosen PreAmp reaction involved 14 cycles of

preamplification with 15 s at 95ºC, and 4 min at 60ºC. Preamplified products were diluted at a ratio 1:10 and then used as templates for RT-PCR analysis. Due to the good results obtained, preamplification was included as a standard step prior to qPCR in the FFPE samples used in the gene expression study. No preamplification step was added into miRNA processing.

2.6 Real-time quantitative PCR

2.6.1 TaqMan low-density array assays (TLDAS).

To explore the different pathways and genes, a series of quantitative RT-PCR assays based on TaqMan® low-density array (TLDA) technology (Applied Biosystems, Foster City, CA), were designed , and measured the expression of each selected gene in triplicate. TLDAs are a very useful tool for gene expression studies. They are based on RT-qPCR technology using the relative quantification method (DeltaCt) .Depending of the chosen configuration, they allow to study from 12 to 384 genes in 1 to 8 samples,. In this work we used 2 different TLDA formats (64, 32) in a sequential gene selection process.



		Number of samples per TaqMan Array			
Format	**Number of assays***	**1 replicate****	**2 replicates**	**3 replicates**	**4 replicates**
Format 12	11				**8**
Format 16	15			**8**	
Format 24	23		**8**		**4**
Format 32	31			**4**	
Format 48	27		**4**		**2**
Format 64	63			**2**	
Format 96a	95		**2**		**1**
Format 96b	95		**2**		**1**
Format 384	380	**1**			

Figure 2.1 Taqman® Low Density array platform and formats. First column indicates number of assays included in each platform; following columns show number of samples analyzed depending on replicates.

2.6.1.1 TLDA models and gene study selection.

A multistep approach was applied to design the subsequent TLDA platforms used in this work, integrating genes known to be expressed either by the tumor HRS cells and their reactive microenvironment, and related with clinical response to adriamycin-based chemotherapy. A total of 2 different TLDA platforms were used. Their different gene composition of each

platform is described in this section. All assay details of genes included in the different TLDA platforms are described in Table 2.8.

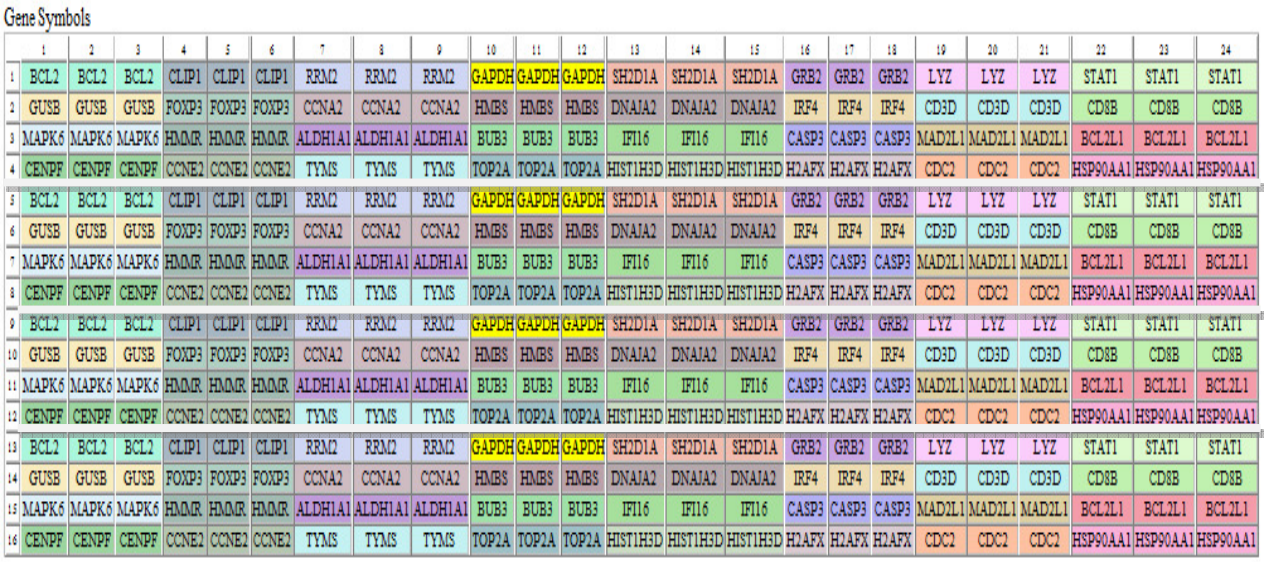
a. TLDA model _ Format 64 _ 2 samples per card, 60 genes study. This initial model included 60 genes from pathways related to cHL outcome that had been previously identified by bioinformatic analysis from gene expression data. It allowed the analysis of 2 samples per array and was applied to an initial set of 52 samples. The main aim of this platform was the technical validation of RT-qPCR technology on FFPE tissue samples. It included 60 study genes and 4 control endogenous genes (HMBS, GUSB, TBP and GADPH).

Gene Symbols

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	GUSB	GUSB	GUSB	IFI16	IFI16	IFI16	HMBS	HMBS	HMBS	18S	18S	18S	ALDH1A1	ALDH1A1	ALDH1A1	MAD2L1	MAD2L1	MAD2L1	DNAJA2	DNAJA2	DNAJA2	IRF4	IRF4	IRF4
2	CENPF	CENPF	CENPF	DPP4	DPP4	DPP4	CDC2	CDC2	CDC2	NBN	NBN	NBN	FOXP3	FOXP3	FOXP3	MAPRE1	MAPRE1	MAPRE1	ITGA4	ITGA4	ITGA4	H1F0	H1F0	H1F0
3	CCNE2	CCNE2	CCNE2	ITMΩA	ITMΩA	ITMΩA	CCNA2	CCNA2	CCNA2	CTSL	CTSL	CTSL	BUB3	BUB3	BUB3	H2AFX	H2AFX	H2AFX	TYMS	TYMS	TYMS	HIST1H3D	HIST1H3D	HIST1H3D
4	CASP3	CASP3	CASP3	CSE1L	CSE1L	CSE1L	TBP	TBP	TBP	CDC6	CDC6	CDC6	BCCIP	BCCIP	BCCIP	CHEK1	CHEK1	CHEK1	HMMR	HMMR	HMMR	MAPK6	MAPK6	MAPK6
5	CDK7	CDK7	CDK7	NUMA1	NUMA1	NUMA1	MAPK14	MAPK14	MAPK14	EMR3	EMR3	EMR3	RSN	RSN	RSN	CYCS	CYCS	CYCS	BUB1B	BUB1B	BUB1B	AURKA	AURKA	AURKA
6	GRB2	GRB2	GRB2	TNFRSF8	TNFRSF8	TNFRSF8	MLH1	MLH1	MLH1	CENPE	CENPE	CENPE	MAPK9	MAPK9	MAPK9	STAT1	STAT1	STAT1	DCK	DCK	DCK	HSP90AA1	HSP90AA1	HSP90AA1
7	CD3D	CD3D	CD3D	MAP3K7	MAP3K7	MAP3K7	CD8B	CD8B	CD8B	TOP2A	TOP2A	TOP2A	DTL	DTL	DTL	CCNH	CCNH	CCNH	BCL2L1	BCL2L1	BCL2L1	LYZ	LYZ	LYZ
8	PTPN11	PTPN11	PTPN11	CDKN2C	CDKN2C	CDKN2C	SH2D1A	SH2D1A	SH2D1A	HSPA4	HSPA4	HSPA4	LCP1	LCP1	LCP1	RRMΩ	RRMΩ	RRMΩ	HSPA9B	HSPA9B	HSPA9B	BCL2	BCL2	BCL2
9	GUSB	GUSB	GUSB	IFI16	IFI16	IFI16	HMBS	HMBS	HMBS	18S	18S	18S	ALDH1A1	ALDH1A1	ALDH1A1	MAD2L1	MAD2L1	MAD2L1	DNAJA2	DNAJA2	DNAJA2	IRF4	IRF4	IRF4
10	CENPF	CENPF	CENPF	DPP4	DPP4	DPP4	CDC2	CDC2	CDC2	NBN	NBN	NBN	FOXP3	FOXP3	FOXP3	MAPRE1	MAPRE1	MAPRE1	ITGA4	ITGA4	ITGA4	H1F0	H1F0	H1F0
11	CCNE2	CCNE2	CCNE2	ITMΩA	ITMΩA	ITMΩA	CCNA2	CCNA2	CCNA2	CTSL	CTSL	CTSL	BUB3	BUB3	BUB3	H2AFX	H2AFX	H2AFX	TYMS	TYMS	TYMS	HIST1H3D	HIST1H3D	HIST1H3D
12	CASP3	CASP3	CASP3	CSE1L	CSE1L	CSE1L	TBP	TBP	TBP	CDC6	CDC6	CDC6	BCCIP	BCCIP	BCCIP	CHEK1	CHEK1	CHEK1	HMMR	HMMR	HMMR	MAPK6	MAPK6	MAPK6
13	CDK7	CDK7	CDK7	NUMA1	NUMA1	NUMA1	MAPK14	MAPK14	MAPK14	EMR3	EMR3	EMR3	RSN	RSN	RSN	CYCS	CYCS	CYCS	BUB1B	BUB1B	BUB1B	AURKA	AURKA	AURKA
14	GRB2	GRB2	GRB2	TNFRSF8	TNFRSF8	TNFRSF8	MLH1	MLH1	MLH1	CENPE	CENPE	CENPE	MAPK9	MAPK9	MAPK9	STAT1	STAT1	STAT1	DCK	DCK	DCK	HSP90AA1	HSP90AA1	HSP90AA1
15	CD3D	CD3D	CD3D	MAP3K7	MAP3K7	MAP3K7	CD8B	CD8B	CD8B	TOP2A	TOP2A	TOP2A	DTL	DTL	DTL	CCNH	CCNH	CCNH	BCL2L1	BCL2L1	BCL2L1	LYZ	LYZ	LYZ
16	PTPN11	PTPN11	PTPN11	CDKN2C	CDKN2C	CDKN2C	SH2D1A	SH2D1A	SH2D1A	HSPA4	HSPA4	HSPA4	LCP1	LCP1	LCP1	RRMΩ	RRMΩ	RRMΩ	HSPA9B	HSPA9B	HSPA9B	BCL2	BCL2	BCL2

Figure 2.2 TLDA Format 64.

b. TLDA model _ Format 32_ 4 samples per card, 30 genes study. Second, the best candidate genes from the initial TLDA 64 assay were selected based on their amplification efficiency, biological significance and treatment response correlation to design and set up a novel assay of genes. This platform was applied to a larger series of 282 samples and used to develop the final predictive model (MRS_ Molecular Risk Score). 2 endogenous control genes (HMBS and GUSB) were present and chosen from the 4 initial ones.



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	BCL2	BCL2	BCL2	CLIP1	CLIP1	CLIP1	RRM2	RRM2	RRM2	GAPDH	GAPDH	GAPDH	SH2D1A	SH2D1A	SH2D1A	GRB2	GRB2	GRB2	LYZ	LYZ	LYZ	STAT1	STAT1	STAT1
2	GUSB	GUSB	GUSB	FOXP3	FOXP3	FOXP3	CCNA2	CCNA2	CCNA2	HMBS	HMBS	HMBS	DNAJA2	DNAJA2	DNAJA2	IRF4	IRF4	IRF4	CD3D	CD3D	CD3D	CD8B	CD8B	CD8B
3	MAPK6	MAPK6	MAPK6	HMMR	HMMR	HMMR	ALDH1A1	ALDH1A1	ALDH1A1	BUB3	BUB3	BUB3	IFI16	IFI16	IFI16	CASP3	CASP3	CASP3	MAD2L1	MAD2L1	MAD2L1	BCL2L1	BCL2L1	BCL2L1
4	CENPF	CENPF	CENPF	CCNE2	CCNE2	CCNE2	TYMS	TYMS	TYMS	TOP2A	TOP2A	TOP2A	HIST1H3D	HIST1H3D	HIST1H3D	H2AFX	H2AFX	H2AFX	CDC2	CDC2	CDC2	HSP90AA1	HSP90AA1	HSP90AA1
5	BCL2	BCL2	BCL2	CLIP1	CLIP1	CLIP1	RRM2	RRM2	RRM2	GAPDH	GAPDH	GAPDH	SH2D1A	SH2D1A	SH2D1A	GRB2	GRB2	GRB2	LYZ	LYZ	LYZ	STAT1	STAT1	STAT1
6	GUSB	GUSB	GUSB	FOXP3	FOXP3	FOXP3	CCNA2	CCNA2	CCNA2	HMBS	HMBS	HMBS	DNAJA2	DNAJA2	DNAJA2	IRF4	IRF4	IRF4	CD3D	CD3D	CD3D	CD8B	CD8B	CD8B
7	MAPK6	MAPK6	MAPK6	HMMR	HMMR	HMMR	ALDH1A1	ALDH1A1	ALDH1A1	BUB3	BUB3	BUB3	IFI16	IFI16	IFI16	CASP3	CASP3	CASP3	MAD2L1	MAD2L1	MAD2L1	BCL2L1	BCL2L1	BCL2L1
8	CENPF	CENPF	CENPF	CCNE2	CCNE2	CCNE2	TYMS	TYMS	TYMS	TOP2A	TOP2A	TOP2A	HIST1H3D	HIST1H3D	HIST1H3D	H2AFX	H2AFX	H2AFX	CDC2	CDC2	CDC2	HSP90AA1	HSP90AA1	HSP90AA1
9	BCL2	BCL2	BCL2	CLIP1	CLIP1	CLIP1	RRM2	RRM2	RRM2	GAPDH	GAPDH	GAPDH	SH2D1A	SH2D1A	SH2D1A	GRB2	GRB2	GRB2	LYZ	LYZ	LYZ	STAT1	STAT1	STAT1
10	GUSB	GUSB	GUSB	FOXP3	FOXP3	FOXP3	CCNA2	CCNA2	CCNA2	HMBS	HMBS	HMBS	DNAJA2	DNAJA2	DNAJA2	IRF4	IRF4	IRF4	CD3D	CD3D	CD3D	CD8B	CD8B	CD8B
11	MAPK6	MAPK6	MAPK6	HMMR	HMMR	HMMR	ALDH1A1	ALDH1A1	ALDH1A1	BUB3	BUB3	BUB3	IFI16	IFI16	IFI16	CASP3	CASP3	CASP3	MAD2L1	MAD2L1	MAD2L1	BCL2L1	BCL2L1	BCL2L1
12	CENPF	CENPF	CENPF	CCNE2	CCNE2	CCNE2	TYMS	TYMS	TYMS	TOP2A	TOP2A	TOP2A	HIST1H3D	HIST1H3D	HIST1H3D	H2AFX	H2AFX	H2AFX	CDC2	CDC2	CDC2	HSP90AA1	HSP90AA1	HSP90AA1
13	BCL2	BCL2	BCL2	CLIP1	CLIP1	CLIP1	RRM2	RRM2	RRM2	GAPDH	GAPDH	GAPDH	SH2D1A	SH2D1A	SH2D1A	GRB2	GRB2	GRB2	LYZ	LYZ	LYZ	STAT1	STAT1	STAT1
14	GUSB	GUSB	GUSB	FOXP3	FOXP3	FOXP3	CCNA2	CCNA2	CCNA2	HMBS	HMBS	HMBS	DNAJA2	DNAJA2	DNAJA2	IRF4	IRF4	IRF4	CD3D	CD3D	CD3D	CD8B	CD8B	CD8B
15	MAPK6	MAPK6	MAPK6	HMMR	HMMR	HMMR	ALDH1A1	ALDH1A1	ALDH1A1	BUB3	BUB3	BUB3	IFI16	IFI16	IFI16	CASP3	CASP3	CASP3	MAD2L1	MAD2L1	MAD2L1	BCL2L1	BCL2L1	BCL2L1
16	CENPF	CENPF	CENPF	CCNE2	CCNE2	CCNE2	TYMS	TYMS	TYMS	TOP2A	TOP2A	TOP2A	HIST1H3D	HIST1H3D	HIST1H3D	H2AFX	H2AFX	H2AFX	CDC2	CDC2	CDC2	HSP90AA1	HSP90AA1	HSP90AA1

Figure 2.3 TLDA Format 32

2.6.1.2 Reference gene selection

The TLDA format 64 included HMBS, GUSB, TBP and 18S as reference genes to normalize the data on the basis on their proven role as housekeeping genes and their uniform expression tested in a small set of FFPE tumor samples from the studied series. Expression stability of the selected endogenous genes was determined by the geNORM Visual Basic application from Real Time StatMiner™ software (Integromics, Madrid,Spain), also known as the pairwise approach(Vandesompele et al., 2002). This analysis relies on the principle that the expression ratio of two perfect endogenous genes should be identical in all samples. GeNorm program calculates a gene stability measure termed as M by determining the average pairwase variation between a particular reference gene and the other studied genes. A higher value of M means greater variation in RNA expression. By stepwise exclusion of the least stable gene and recalculation of the M values, most stable reference genes are identified and a final normalization factor (NF) is calculated (based on the geometric mean of the expression levels of the best reference genes).The best two candidates were HMBS and GUSB (NF = 0.0902), thus being chosen to be included in the other TLDA platforms (Format 32 and 16) and for normalization of the gene expression levels in all the analysis.

Table 2.7 GeNorm Analysis candidate endogenous genes TLDA assays

Gene	**18S**	**TBP**	**HMBS/GUSB**
NF	0.206508201	0.104911765	**0.090217764**

NF- Normalization factor

Table 2.8 Assays details included in the RT-qPCR TLDA platforms used in the study

Control genes	Description	Assay ID	bp		Format 64	Format 32
GUSB	glucuronidase, beta	Hs99999908_m1	81		x	x
18S	Eukaryotic 18S rRNA	Hs99999901_s1	187		**x**	x
TBP	TATA box binding protein	Hs00427620_m1	91		**x**	
HMBS	hydroxymethylbilane synthase	Hs00609297_m1	64		**x**	
Study Genes	**Description**	**Assay ID**	**bp**		**Format 64**	**Format 32**
BCL2	B-cell CLL/ lymphoma2	Hs00608023_m1	81	t	**x**	x
BCL2L1	BCL2-like 1	Hs00236329_m1	65	t	**x**	x
CASP3	caspase 3, apoptosis-related cysteine protease	Hs00234385_m1	66	t	**x**	x
CCNA2	cyclin A2	Hs00153138_m1	110	t	**x**	x
CCNE2	cyclin E2	Hs00180319_m1	92	t	**x**	x
CDC2	cell division cycle 2, G1 to S and G2 to M	Hs00176469_m1	101	t	**x**	x
CENPF	centromere protein F, 350/400ka (mitosin)	Hs00193201_m1	99	t	**x**	x
GADPH	glyceraldehyde-3-phosphate dehydrogenase	Hs99999905_m1	125		**x**	x
HMMR	hyaluronan-mediated motility receptor (RHAMM)	Hs00234864_m1	98	t	**x**	x
IRF4	Interferon regulatory factor 4	Hs00180031_m1	81	m	**x**	x
LYZ	lysozyme (renal amyloidosis)	Hs00426231_m1	84	m	**x**	x
STAT1	signal transducer and activator of transcription 1, 91kDa	Hs00234829_m1	79	m	**x**	x
TYMS	thymidylate synthetase	Hs00426591_m1	87	t	**x**	x
MAD2L1	MAD2 mitotic arrest deficient-like 1 (yeast)	Hs00829154_g1	161	t	**x**	x
CD3D	CD3D antigen, delta polypeptide (TiT3 complex)	Hs00174158_m1	92	m	**x**	x
SH2D1A	SH2 domain protein 1A, Duncan's disease (lymphoproliferative syndrome)	Hs00158978_m1	93	m	**x**	x
CD8B1	CD8 antigen, beta polypeptide 1 (p37)	Hs00174762_m1	95	m	**x**	x
FOXP3	forkhead box P3	Hs00203958_m1	64	t/m	**x**	x
ALDH1A1	aldehyde dehydrogenase 1 family, member A1	Hs00167445_m1	97	m	**x**	x
BCCIP	BRCA2 and CDKN1A interacting protein	Hs00255585_m1	104	t	**x**	x
BUB3	BUB3 budding uninhibited by benzimidazoles 3 homolog	Hs00190920_m1	*78*	t	**x**	x
DNAJA2	DnaJ (Hsp40) homolog, subfamily A, member 2	Hs00195365_m1	76	t	**x**	x
GRB2	growth factor receptor-bound protein 2	Hs00257910_s1	68	t	**x**	x
H2AFX	H2A histone family, member X	Hs00266783_s1	48	t	**x**	x
HIST1H3D	histone 1, H3d	Hs00371415_s1	62	t	**x**	x
HSPCA	heat shock 90kDa protein 1, alpha	Hs00743767_sH	133	t	**x**	x
IFI16	interferon, gamma-inducible protein 16	Hs00194261_m1	83	m	**x**	x
MAPK6	mitogen-activated protein kinase 6	Hs00833126_g1	100	t	**x**	x
TOP2A	topoisomerase (DNA) II alpha 170kDa	Hs00172214_m1	125	t	**x**	x
RRM2	ribonucleotide reductase M2 polypeptide	Hs00357247_g1	79	t	**x**	x
MAPK9	mitogen-activated protein kinase 9	Hs00177102_m1	102	t	**x**	
MLH1	mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli)	Hs00179866_m1	103	t	**x**	
NUMA1	nuclear mitotic apparatus protein 1	Hs00272062_m1	94	t	**x**	
RSN	restin (Reed-Steinberg cell-expressed intermediate filament-associated protein)	Hs00161477_m1	65	m	**x**	

Study Genes	Description	Assay ID	bp		Format 64	Format 32
AGC1	aggrecan 1 (chondroitin sulfate proteoglycan 1, large aggregating proteoglycan, antigen identified by monoclonal antibody A0122)	Hs00202971_m1	93	m	x	
BUB1B	BUB1 budding uninhibited by benzimidazoles 1 homolog beta (yeast)	Hs00176169_m1	90	t	x	
CCNH	cyclin H	Hs00236923_m1	127	t	x	
CD30 (TNFRSF8)	tumor necrosis factor receptor superfamily, member 8	Hs00174277_m1	94	t	x	
CDC6	CDC6 cell division cycle 6 homolog	Hs00154374_m1	77	t	x	
CDK7	cyclin-dependent kinase 7	Hs00361486_m1	68	t	x	
CDKN2C	cyclin-dependent kinase inhibitor 2C (p18, inhibits CDK4)	Hs00176227_m1	86	t	x	
CENPE	centromere protein E, 312kDa	Hs00156507_m1	88	t	x	
CHEK1	CHK1 checkpoint homolog (S. pombe)	Hs00176236_m1	112	t	x	
CSE1L	*CSE1 chromosome segregation 1-like (yeast)*	Hs00354853_m1	*123*	t	x	
CYCS	cytochrome c, somatic	Hs01588974_g1	85	t	x	
CTSL	cathepsin L	Hs00377632_m1	85	m	x	
DCK	deoxycytidine kinase	Hs00176127_m1	93	t	x	
DPP4	dipeptidylpeptidase 4 (CD26, adenosine deaminase complexing protein 2)	Hs00175210_m1	90	m	x	
EMR3	egf-like module containing, mucin-like, hormone receptor-like 3	Hs00261470_m1	72	m	x	
H1F0	H1 histone family, member 0	Hs00271174_s1	107	t	x	
HSPA4	heat shock 70kDa protein 4	Hs00382884_m1	62	t	x	
HSPA9B	heat shock 70kDa protein 9B (mortalin-2)	Hs00269818_m1	92	t	x	
ITGA4	integrin, alpha 4 (antigen CD49D, alpha 4 subunit of VLA-4 receptor)	Hs00168433_m1	61	m	x	
LCP1	lymphocyte cytosolic protein 1 (L-plastin)	Hs00158701_m1	72	m	x	
MAP3K7	mitogen-activated protein kinase kinase kinase 7	Hs00177373_m1	95	m	x	
MAPK14	mitogen-activated protein kinase 14	Hs00176247_m1	62	m	x	
MAPRE1	microtubule-associated protein, RP/EB family, member 1	Hs00606526_mH	*93*	t	x	
NBS1	Nibrin	Hs00159537_m1	75	t	x	
PTPN11	protein tyrosine phosphatase, non-receptor type 11 (Noonan syndrome 1)	Hs00275784_m1	83	t	x	
RAMP	null	Hs00212788_m1	91	t	x	
STK6	serine/threonine kinase 6 (AuroraA)	Hs00269212_m1	85	t	x	

bp_ base pairs ; m_ microenvironment; t_tumor cells

2.6.1.3 Experimental procedure and data analysis

Each cDNA sample (30 µl) was added to 20 µl of RNase-free water and 50 µl of 2 x TaqMan® Universal PCR Master Mix (No AmpErase UNG) (Applied Biosystems, Foster City, CA). The mixture was then transferred into a loading port on a TLDA card. The card was centrifuged twice, sealed, and PCR amplification was performed using Applied Biosystems Prism 7900HT Sequence Detection System under the following thermal cycler conditions:

Table 2.9 PCR conditions used for TLDA platforms		
Program	**Temperature**	**Time (min)**
HOLD	50ºC	2
CYCLE (x 40)	94,5ºC	10
	97ºC	0.5
	59,7ºC	1

As previously explained, a preamplification step (PreAmp, Applied Biosystems, CA) was used to improve the sensitivity of the assays for low-abundance target genes available from FFPE samples. Ct values were exported using SDS software (SDS 2.3), and the data were analyzed with Real-Time StatMiner software (INTEGROMICS TM; www.Integromics.com) for normalization. -ΔCT values (-(Ct value of gene – median Ct value for HMBS and GUSB) were considered for further statistical analysis if they meet quality criteria (Ct<35).

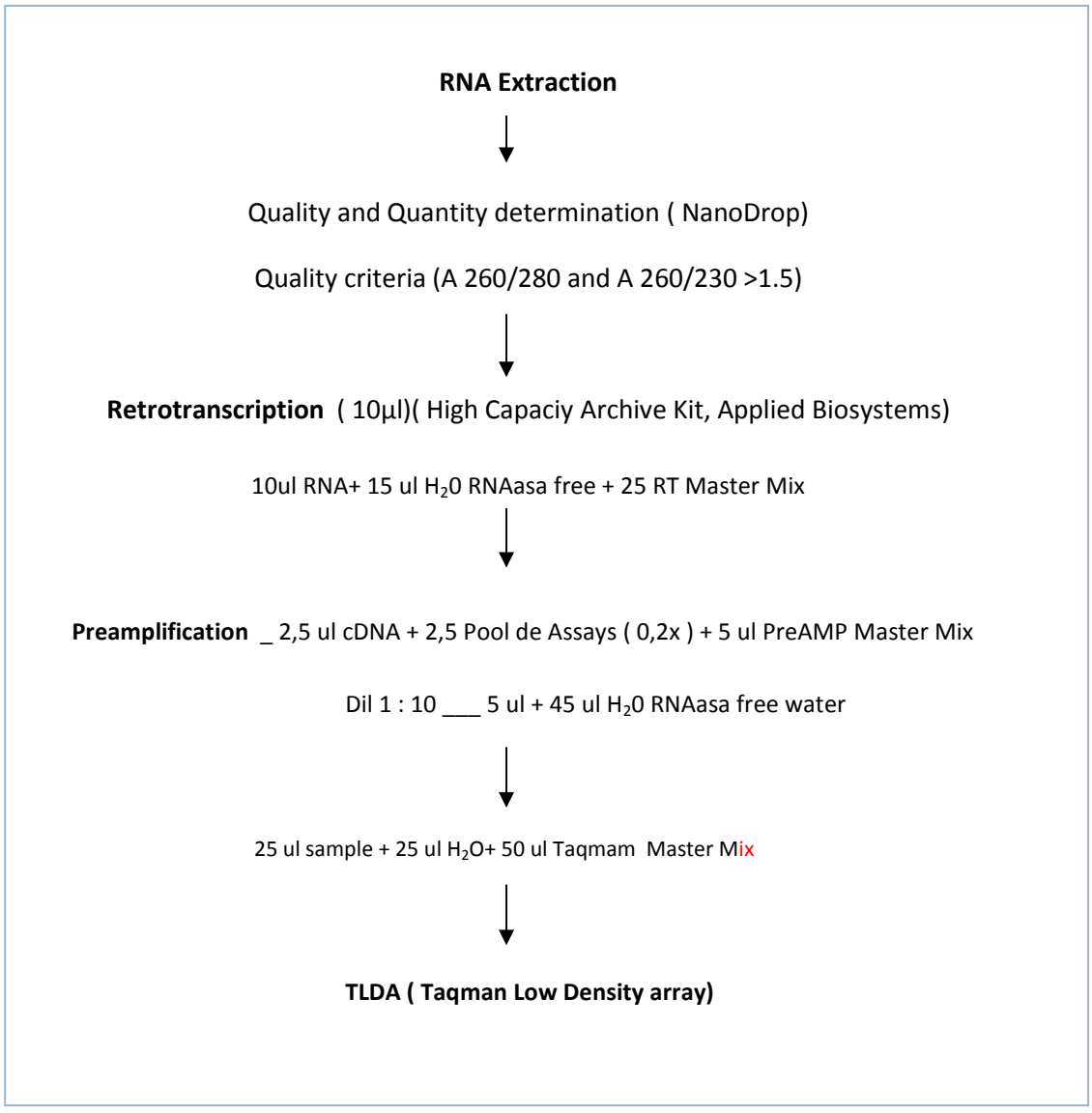


Figure 2.4. Protocol summary TLDAs

2.6.2 Real-time PCR for relative miRNA quantification using RNA from FFPE tissue.

For FFPE samples, miRNA RT-qPCR was performed for the 34 selected miRNAs (Table 2.10) following manufacturer protocol (Applied Biosystems, Foster City, CA, USA) by the Applied Biosystems 384-well multiplexed real-time PCR assay using 250 ng of total miRNA.

Table 2.10 Assay Details miRNAs analyzed in the study m_microenvironment; t_tumor cells

Name	Assay ID	Description	Signal	miRBase accession
RNU44	001094	Endogenous Control		
RNU48	001006	Endogenous Control		
hsa-miR-15b*	002173	Mature miRNA	t	MIMAT0004586
hsa-miR-17*	002421	Mature miRNA	t	MIMAT0000071
hsa-miR-216a	002220	Mature miRNA	t	MIMAT0000273
hsa-miR-23a	000399	Mature miRNA	t	MIMAT0000078
hsa-miR-25	000403	Mature miRNA	t	MIMAT0000081
hsa-miR-30d	000420	Mature miRNA	t	MIMAT0000245
Hsa-miR-30e	002223	Mature miRNA	t	MIMAT0000692
hsa-miR-320	002277	Mature miRNA	t	MIMAT0000510
hsa-miR-93	000432	Mature miRNA	t	MIMAT0002762
hsa-miR-21	000397	Mature miRNA	t	MIMAT0000076
hsa-miR-21*	002438	Mature miRNA	t	MIMAT0004494
hsa-miR-92a	000431	Mature miRNA	t	MIMAT0000092
hsa-miR-130b	000456	Mature miRNA	t	MIMAT0000691
hsa-miR-378	002243	Mature miRNA	t	MIMAT0000732
hsa-miR-551a	001519	Mature miRNA	t	MIMAT0003214
hsa-miR-708*	002342	Mature miRNA	t	MIMAT0004927
hsa-miR-92b*	002343	Mature miRNA	m	MIMAT0004792
hsa-miR-26a	000405	Mature miRNA	m	MIMAT0000082
hsa-miR-27b*	002174	Mature miRNA	m	MIMAT0004588
hsa-miR-132	000457	Mature miRNA	m	MIMAT0000426
hsa-miR-148b*	000471	Mature miRNA	m	MIMAT0000759
hsa-miR-204	000508	Mature miRNA	m	MIMAT0000265
hsa-miR-300	241035_mat	Mature miRNA	m	MIMAT0004903
hsa-miR-31*	002279	Mature miRNA	m	MIMAT0004504
hsa-miR-423-5p	002340	Mature miRNA	m	MIMAT0004748
hsa-miR-503	001048	Mature miRNA	m	MIMAT0002874
hsa-miR-539	001286	Mature miRNA	m	MIMAT0003163
hsa-miR-559	001527	Mature miRNA	m	MIMAT0003223
hsa-miR-609	001573	Mature miRNA	m	MIMAT0003277
hsa-miR-621	001598	Mature miRNA	m	MIMAT0003290
hsa-miR-628-3p	002434	Mature miRNA	m	MIMAT0003297
hsa-miR-637	001581	Mature miRNA	m	MIMAT0003307
hsa-miR-888	002212	Mature miRNA	m	MIMAT0004916
hsa-miR-937	002180	Mature miRNA	m	MIMAT0004980

Briefly, miRNA from each case was reversed-transcribed as previously explained in triplicate using a multiplex looped primer pool with the selected miRNA probes for the study.. Each completed reaction was loaded onto the 384-well plate and real-time PCR was done on the ABI 7900HT Prism using RNU44 and RNU48 as the endogenous normalization controls. Ct values were exported using SDS software (SDS 2.3), and data were analyzed with Real-Time StatMiner software (INTEGROMICS TM; www.Integromics.com) using the ΔCT method. All assays were performed in triplicate and results which did not meet methodological quality control criteria were omitted considering a miRNA to be present if the Ct was less than 36 in all three biological replicates; -ΔCT values (-(Ct value of miRNA of interest – median Ct value for RNU44 and RNU48) were considered for further statistical analysis.

2.7 Gene expression microarrays

In this work the Whole Human Genome Microarray kit v2 (1x44K, Agilent Technologies) was used.

2.7.1 cDNA synthesis from total RNA

2µg of total RNA were mixed with 2µl of a 5,000-fold dilution of Agilent´s Two-Color Spike-in RNA control and amplified using Agilent Low RNA Input Fluorescent Amplification Kit (Agilent Technologies, Inc., Santa Clara, CA). A final mixture volume of 6,5 µl (total concentration at least 5ng/µl) was mixed with 5µl of T7 promoter primer. The primer and the template were denatured by incubating the reaction at 65ºC for 10min and placing on ice for 5min. Following, 8.5µl of cDNA Master Mix was added and the samples were incubated first at 400C in a circulating water bath for 2h and then at 65ºC in a heating block for 15 min to inactivate MMLV-RT. After this time, samples were incubated on ice for 5 min.

cDNA Master Mix composition: 4µl of 5X First strand buffer, 0.1M DTT 2µl, 10mM dNTP mix 1µl of MMLV-RT and 0.5µl of RNase OUT.

2.7.2 Fluorescent cRNA synthesis: In Vitro Transcription and incorporation of fluorochromes.

To each sample tube, either 2.4µl of 10mM cyanine 3-CTP (sample) or 2.4µl of 10mM cyanine 5-CTP (Stratagene Universal Human Reference RNA) were added and mixed. Following, to each sample, 57.6µl of Transcription Master Mix was added and incubated in a circulated water bath at 40ºC for 2h. Following amplification and labeling, each sample was assessed on the Nanodrop ND-1000 to measure yield and specific activity.

Transcription Master Mix composition: 15.3µl of Nuclease-free water, 20µl of 4X Transcription buffer, 0.1 M DTT 6µl, 8µl of NTP mix, 50% PEG 6.4µl, 0.5 µl of RNase OUT, 0.6µl of Inorganic pyrophosphatase and 0.8µl T7 RNA Polymerase.

2.7.3 Hybridization

cRNA target was prepared as follows: 0.75µg cyanine 3-labeled, linearly amplified sample cRNA was mixed with 0.75µg cyanine 5-labeled, linearly amplified reference pool cRNA, 50µl of 10X control targets and Nuclease-free water to final volume of 240µl. The hybridization solution was prepared by adding 240µl of 2X target solution to 10µl of 25X fragmentation buffer. The mixture was incubated at 60ºC in the heating block for 30 min. Following, 250µl of 2X hybridization buffer (from In situ Hybridization kit) to the final volume of 500µl, mixed, spinned and 490µl of the hybridization solution was applied to 60-mer Agilent 44 K (or 4X 44K) Human Whole Genome oligonucleotide microarrays and assembled in microarray hybridization chamber (G2534A). Once fully assembled, the chambers were loaded into the hybridization rotator rack and set to rotate at 4rpm. The hybridization was performed in a rotating oven at 60ºC for 17h. All the washing steps were performed at room temperature. First the sandwiched slides were submerged in Wash Solution 1 to remove oligo microarray slide. The slides were washed for 1min in the Wash Solution 1 with the magnetic stir and transferred to the staining dish containing Wash Solution 2 and washed for 1min. Following, the slides were transferred to the staining dish containing the Wash Solution 3 and washed for 30 seconds. All steps were performed in darkness. The dried slides were scanned with a G2565BA Microarray Scanner System (Agilent Technologies, Palo Alto, CA).

Wash solution 1 composition: 6X SSPE, 0.005% N- Lauroylsarcosine, deionized nuclease free water.

Wash solution 2 composition: 0.06X SSPE, 0.005% N- Lauroylsarcosine, deionized nuclease free water.

The buffers 1 and 2 are passed through a 0.2µm sterile filtration unit before use.

Wash solution 3 composition: Agilent Stabilization and Drying Solution containing an ozone scavenging compound dissolved in acetonitrile.

2.8 MicroRNA expression microarrays

2.8.1 Labeling

100 ng of total RNA were hybridized onto Agilent 8x15 Human microRNA Microarray kit version2 containing probes from 723 human and 76 viral microRNAs based on Sanger

miRbase, following the manufacturer´s instructions (Agilent Technologies). Briefly, this procedure consisted of a RNA dephosphorylation step where the mixture of 100 ng RNA, 10X CIP Buffer and CIP (16 U/µl) was incubated at 37ºC for 30 min. Following, a denaturation step was performed through addition of DMSO to samples and carried out at 100ºC for 5min. Samples were then treated with T4 ligase (15 U/µl)and labeled with Cy3 at 16ºC for 2 hours. The reaction was performed in a total volume of 20µl. Subsequently, RNA was purified with Bio-spin 6 columns (BioRad, Hercules, CA) and dried in a speed vacuum (45ºC, 45 minutes).

2.8.2 Hybridization

Labeled RNA was resuspended in the appropriate volume of hybridization mix (H_2O 18µl, blocking agent 4.5µl and hybridization buffer 22.5µl) to a final volume of 45µl and incubated at 100ºC for 5 minutes, quickly cooled and pippeted into the 8X gasket, closed with the Agilent 8x15K Human miRNA v2 platform, and locked with the appropriate support. Hybridization was carried out at 55ºC at 20rpm. After that time the slides were washed first with WB1 buffer and then with WB2 buffer. Both washing steps were performed at 37ºC for 7 minutes each, with shaking. Finally, the array was submerged in acetonitrile for 1 minute and dried. The arrays were scanned for Cy3 fluorescence with a G2565BA Microarray Scanner System (Agilent Technologies, Palo Alto, CA).

2.9 Tissue Microarrays

Tissue Microarrays (TMA) technology allows for the simultaneous analysis of multiple individual tissue samples on a single slide through immunohistochemistry, i*n situ* hybridization and fluorescence *in situ* hybridization techniques (Kallioniemi et al., 2001);(Kononen et al., 1998) and consist on the arrayed disposition of a variable number of circular biopsies (0.6-2mm diameter) originated from different FFPE blocks into a unique block. This technology takes advantage of the possibility of analyzing a big number of samples in homogeneous experimental conditions with a minimum destruction of the original tissue sample and a low time and money consumption. The possibility of bias due to non-representative tissue sections is overcome by the number of tumors analyzed as demonstrated by several studies.

The expression of a selected number of markers was assessed in a set of 142 FFPE samples Primary chosen antibodies were as follows: anti-BCL2 and anti-LYZ (DAKO), anti-BCLXL (anti-BCL2L1) (Zymed Laboratories); anti-CASP3 and anti-CCNA2 (Novocastra Laboratories); anti-CTSL (Alexis Biochemicals); anti-STAT1, anti-SH2D1A, anti-CDK7, anti-HSP70 and anti-

MUM1 (IRF4; Santa Cruz Biotechnology) ; anti-CCNH (Cell Signaling Technology); anti-HIST1H and anti-HISTH2A (Upstate Biotechnology).

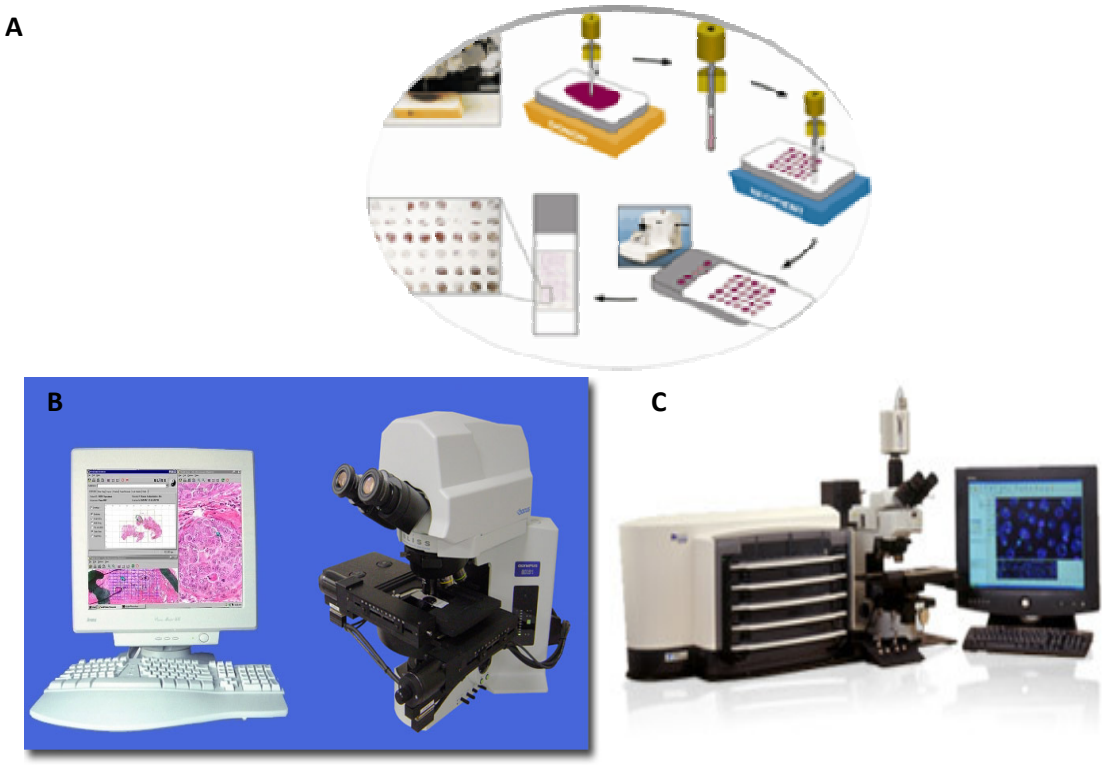


Figure 2.5 TMA´s construction scheme procedure (A), scanning method (B) (Olympus digital microscope and quantification method (C) (Ariol SL.50, Genetix).

The TMAs were scanned using a computerized microscope (Olimpus BX61) and stored as digital images. Protein levels were subsequently quantified by adjusting the appropriate thresholds for each marker using the TMA score software package (Ariol SL-50, Genetix).

2.10 Analysis of the presence of the Epstein- Bar virus (EBV).

The presence of EBV in cHL lymph nodes was examined by in situ hybridization for EBV RNA (EBER 1 and 2, Dako) and LMP1 immunohistochemistry (CS1-4, Dako) using routine procedures on TMAs constructed with paraffin-embedded tissue samples. This EBV detection was carried out in FFPE tissue by the Histology and Immunohistochemistry Core Unit at the Spanish National Cancer Research Center (CNIO).

2.11 Laser capture microdisecction

Laser Capture microdissection (LCM) was performed for $CD30^{+}$ HRS cells in a small set of 4 frozen tissue samples to validate the expression of specific miRNAs in the HRS cells.

Briefly, five-micrometer serial sections from the frozen tissue blocks were cut, fixed with acetone at 4 °C, and incubated at 1:250 with anti-human CD30 monoclonal antibody (clone Ber-H2, DAKO) at 1:1,000 to identify the HRS cells, using a standard indirect avidin-biotin horseradish peroxidase method and diaminobenzidine color development.

After immunostainig, 600 to 800 CD30-positive HRS cells were microdissected using the PALM UV Laser Microbeam system (Carl Zeiss Inc., Germany). In each experiment, equal number of CD30-negative cells was randomly picked from the reactive microenvironment.

After microdissection, miRNA expression was analyzed by RT-qPCR as previously explained. Quantitative RT-PCR for miR-21 and miR30d was performed with microdissected samples testing also the levels of the same miRNAs in microdissected fragments obtained from the reactive surrounding infiltrate of each case. Ct values using RNU44 and RNU48 as endogenous controls were used for analysis.

2.12 Microarray data analysis

2.12.1 Gene expression and miRNA microarrays preprocessing

Data were processed and extracted by Feature Extraction software (v.9.5.3.1) (Agilent Technologies) being afterwards combined and filtered for saturated and non-uniformly stained flags using the combine program (http://combine.bioinfo.cnio.es/). Between-array median normalization was carried out.

2.12.2 Bioinformatic methods

2.12.2.1 Supervised analysis. Pomelo II Tool.

The analyses of the initial gene expression datasets (either Gene Expression or miRNA expression dataset) were performed using Pomelo II tool(Morrissey and Diaz-Uriarte, 2009), available free at http://pomelo2.bioinfo.cnio.es . To identify genes and miRNAs differentially expressed between tumoral samples and cHL cell lines we used a supervised method based on Student´s t test with a correction of multiple testing. Unadjusted P values were obtained from 100,000 permutations of the dataset and false discovery rates (FDRs) were calculated by the method of Benjamini and Hochberg(Green and Diggle, 2007) considering genes with an FDR<0.15 as differentially expressed. Genes and miRNAs with FDR<0.15 were considered to be differentially expressed and used to construct the Tumoral and Microenvironmental signatures from both gene (Tumoral Database (TDB; Microenvironmental Database (MDB)) and

microRNA studies. This first analysis was done without consideration of clinical outcomes of the patients.

2.12 .2.2 Gene Set Enrichment Analysis

Gene Set Enrichment Analysis (GSEA) (http://www.broad.mit.edu/gsea) (Subramanian et al., 2005) is a computational method that determines whether an a priori predefined set of genes (pathways) shows statistically significant differences among two biological states. In this work GSEA tool was used in the Gene Expression study (Results I) to identify sets of related genes that might correlate to therapy resistance (Favorable versus Unfavorable Outcome) in the previously identified TDB and MDB databases testing each one separately . GSEA´s method uses t statistics to search for predefined list of gene sets that were in this work selected taking into consideration the particular heterogeneity of cHL tumors including all public pathways associated with immune response and cell-cell interactions.

Biocarta (www.biocarta.com) and other public sources available through the Molecular Signature Database (MSigDB: http://www.broad.mit.edu/gsea/msigdb/index.jsp) were used to generate the gene-set databases. Final gene sets used were manually curated and enriched in pathways also known to be involved in lymphoma pathogenesis and immune response. The rank of all genes in the sets were determined and an enrichment score (ES) was calculated for each pathway as a measure of its relevance. This enrichment score (ES) value was afterwards normalized in order to calculate the Normalized enrichment score (NES) used for pathway evaluation (See Results I).

The analysis was performed with 1,000 random class permutations. Genes with more than 30% of missing values were excluded , and only gene sets meeting the gene set size criteria (min=10, max=500 genes) were analyzed. Thus, final analysis of TDB included 46 pathways whereas 21 pathways were taken into account for MDB analysis.

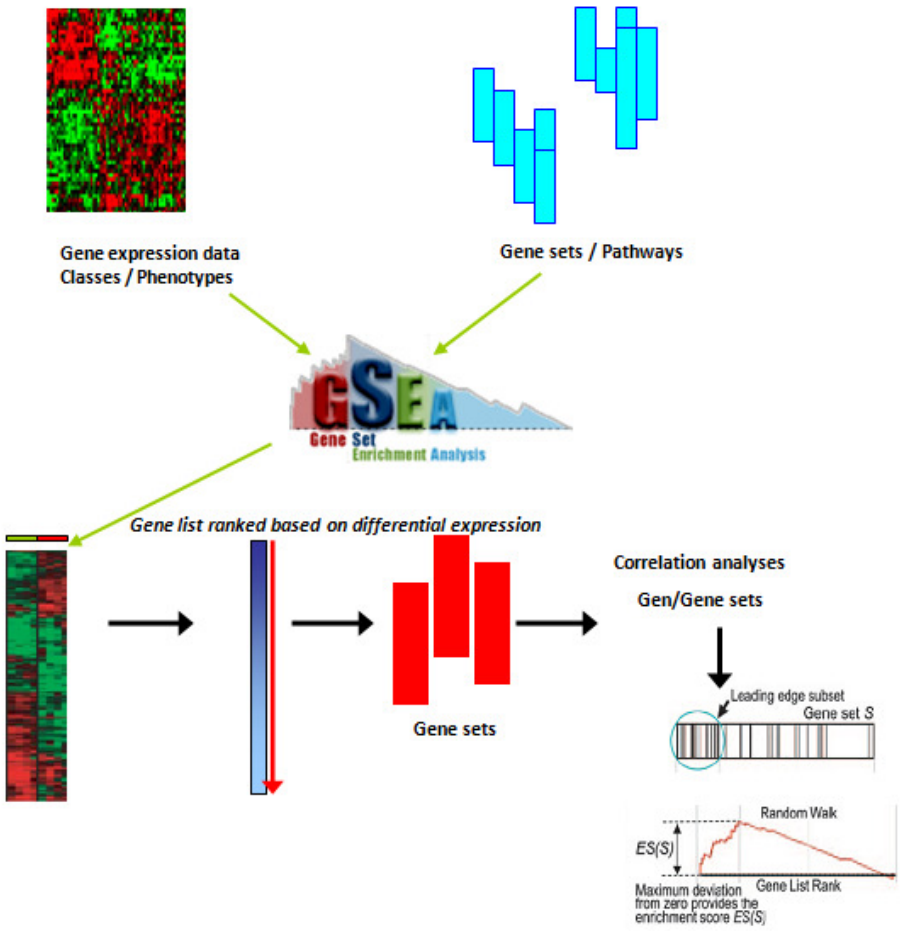


Figure 2.6 GSEA workflow scheme analyses.

2.13 Bioinformatics Target miRNA prediction.

miRBase (miRBase v11.0, MICROCOSM)(Griffiths-Jones, 2004; Griffiths-Jones et al., 2006; Griffiths-Jones et al., 2008)and TargetScan (TargetScan release 5.1, Whitehead Institute for Biomedical Research)(Friedman et al., 2009; Grimson et al., 2007; Lewis et al., 2005) databases were interrogated to identify the genes that might be regulated by the 4- miRNAs identified with outcome in this series of cHL cases. Details of algorithms used by this bioinformatic tools can be found on their respective web sites (http://www.mirbase.org/ and http://www.targetscan.org/). Additionally, TarBase (Sethupathy et al., 2006) (http://diana.cslab.ece.ntua.gr/tarbase/), a database which houses a manually curated collection of experimentally supported microRNA targets was included in the analysis.

Finally, with the aim of incorporating microRNAs in pathways DIANA-mirPath (http://diana.cslab.ece.ntua.gr/pathways/index_multiple.php) (Papadopoulos et al., 2009) was used. It consists on a web-based computational tool developed to identify molecular pathways potentially altered by the expression of single or multiple microRNAs. The software performs an enrichment analysis of multiple microRNA target genes comparing each set of microRNA targets to all known KEGG pathways (http://www.genome.jp/kegg/pathway.html). The combinatorial effect of co-expressed microRNAs in the modulation of a given pathway is taken into account by the simultaneous analysis of multiple microRNAs. The graphical output of the program provides an overview of the parts of the pathway modulated by microRNAs, facilitating the interpretation and presentation of the analysis results.

2.14 Statistical data analysis

All statistical analyses were two-sided, taking values of $p < 0.05$ to be significant. These were performed with SPSS 15.0 (SPSS Inc., Chicago, IL). Survival curves were assessed by the Kaplan-Meier method and risk groups were compared by the log-rank test. Plots were generated using GraphPadPrism v.5 (GraphPad Software, Inc.).

2.14.1 Clinical endpoints

Differences in the distributions of standard clinical parameters (age, gender, stage, IPS and outcome) in all patient series used in the different analyses were tested by the Pearson chi-square test When patient series were split into estimation and validation sets, it was randomly done in order to avoid any bias into the event distribution, thus with a balanced proportion in favorable versus unfavorable outcomes.

The major aim of this study was the analysis of biological parameters associated with response to the initial therapy and treatment failure, thus data from second-line and salvage therapies and/or bone-marrow transplantation were not considered.

2.14.1.1 Treatment outcome (Favorable versus Unfavorable)

The first endpoint of all studies used for initial gene and miRNA selection was maintained response to therapy (treatment outcome).

- Favorable outcome (F) . Good response was considered in patients with sustained complete response (18 months).

- Unfavorable outcome (U). Bad response was considered in cases without complete response or patients with early relapses to therapy following previously published criteria(Carde et al., 2002).

2.14.1 Survival analysis (FFS and OS)

-Failure-free survival (FFS) was used as the fundamental endpoint for survival analysis, defined as the time interval between treatment initiation and treatment failure or last follow-up. Failure was defined as either the failure to achieve complete remission (CR), or the occurrence of progressive disease, irrespective of whether there had been an initial CR.

-Overall survival (OS) was included as a secondary end point in the survival analyses, defined as the time interval between diagnosis and death due to the lymphoma or any other causes. However, its significance in the study was imperfect since it is conditioned by the effect of subsequent eventual treatments and complication of treatment.

2.14.2 Logistic Regression analysis

Associations between gene expression levels and the probability of treatment response and outcome were studied by logistic regression analyses. Normalized expression levels were used as independent continuous variables and either treatment outcome (F vs. U) or FFS event (relapse versus non relapse) as the dependent variable. Univariate and multivariate backward stepwise logistic regression analyses were used to elucidate and develop the different predictive signatures.

2.14.3 ROC curves

Receiver operating characteristic (ROC), or simply **ROC curve**, is a graphical plot of the sensitivity, or true positive rate, versus false positive rate (1 – specificity or 1 − true negative rate), for a binary classifier system as its discrimination threshold is varied. ROC analysis provides a tool to select possibly optimal models and to discard suboptimal ones independently from (and prior to specifying) the cost context or the class distribution.

Thus, ROC curves were used in the present work to select the best markers (genes and miRNAs) and to evaluate the predictive sensitivity and specificity of the different elucidated signatures and derived scores (Integrated Risk, Molecular Risk Score and miRNA score) using outcome (F versus U) as binary classifier variable. Additionally, the MRS and miRNA score cut-off points were prespecified by using area ROC analysis to define different risk groups (low risk versus high risk) choosing those cut-off points with highest compromise between sensitivity and specificity. In clinical practice, it is exceedingly rare that a chosen cut point will achieve perfect discrimination between cases. Thus, the best compromise between sensitivity and specificity has to be selected as done in this thesis work (Zweig and Campbell, 1993).

2.14.4 Survival analysis (Cox regression analysis and Kaplan-Meier Test)

Survival was estimated using the Kaplan-Meier method and compared by log-rank tests either in the Gene expression or miRNA project. For graphical representation, survival analyses were done by the Kaplan-Meier method and long-rank test separately in the training and validation series for MRS development, and in the entire series for both the MRS and miRNA signatures respectively. In addition, multivariate Cox proportional hazards models were used to determine and test the joint associations of factors initially found to have potential association with outcome and derive final models to be applied for the entire series: Integrative Model (N=262) (Results I) and miRNA score (N=220) (Results II).

2.15 The derived scores.

2.15.1Gene Expression project

2.15.1.1 Integrated Risk Score

To illustrate the relationship among gene expression data from TLDA Format 64 and treatment response an integrated risk score was derived, defined as the logarithmic mean of the expression levels of all the genes included in the analysis. This approach combined with ROC analysis of the individual genes from TLDA Format 64 was also used to select the genes with the better prognostic ability in order to design the TLDA Format 32 that was applied to a larger series of patients (N=282).

2.15.1.2 Molecular Risk Score

The Molecular Risk Score (MRS) algorithm was defined as a continuous function obtained by logistic regression analysis and comprised expression levels of 11 genes

significantly found to predict failure that were grouped into their corresponding biological functional pathways.

The MRS was developed using gene expression data from the estimation set (N=183) and further validated in the validation set (N=79). Briefly, univariate regression analysis was first performed with treatment response (F versus U) as the dependent variable to identify genes associated with outcome. Genes were ranked on basis on their p value and those ones found to be significant were further studied by cross-validation.

A cross-validation step was included to test the classification ability of the initial set of significant genes to choose the strongest predictor ones. For this purpose, Tnasas, a web tool for building class prediction models available at (http://tnasas.bioinfo.cnio.es/) was used and three prediction algorithms available were applied: Diagonal linear discriminant analysis(DLDA)(Solberg, 1978), Support vector machines (SVM)(Chu and Wang, 2005) a and K-nearest neighbor (KNN).

Using this bioinformatic tool several predictive models including the previously identified significant genes were derived. To construct each predictive model the program used values of the F statistic (the popular ANOVA F-ratio) to rank genes and each one of the above described algorithms were used. The frequency of genes in the different developed models were used for ranking, and counted after 150 cross-validation runs of each model. Final selected genes to be included in the MRS were those ones with frequencies higher than 40 percent in all-cross validated runs (CV).

Selected genes were classified into functional groups on the basis of their known biological relationship and their coregulated expression as estimated by Pearson correlation coefficient. Three biological functional groups were defined: Cycle, Apoptosis and macrophage with the inclusion of a gene as independent variable not ascribed to any functional group. Individual genes from each functional group were weighted using diagonal linear discriminant analysis (DLDA). Finally, these functional gene clusters associated with cHL outcome were analyzed in a multivariate logistic regression model with response to therapy (F versus U) as a dependent variable. In this way, an algorithm was derived that combined these measurements into a quantitative “molecular risk score” (MRS), used as a continuous variable to estimate the probability of treatment response. Further details of MRS development are explained in Results I section.

2.15.1.3 Integrative Model.

Initially, a multivariate Cox's proportional hazards model, including the data at diagnosis, the IPS stratified (low <3 versus high ≥3), and its seven individual variables (hemoglobin < 10.5 g/dl; albumin < 4 g/dl; leucocytosis ≥15,000/mm^3; lymphopenia < 600/mm^3; age ≥ 45 yr; male gender and stage IV), was applied to the entire series in order to test the independence of the MRS with respect to clinical variables. The integrative model was defined as a continuous survival function derived by Cox proportional hazard model using backward stepwise selection algorithm that was applied to the entire series (N=262). This model comprised two variables: the previously described MRS and the clinical variable Stage IV found to be the only clinical variable that remained significant after multivariate Cox regression analysis (See Results I).

2.15.2 MiRNA expression project

2.15.2.1 MiRNA score

For miRNA signature score, Cox regression analysis was performed to determine the association of the miRNAs initially found to have potential association with outcome in the logistic regression analyses. A final Cox proportional hazard model based on the miRNA expression was finally derived for the entire series of FFPE samples analyzed (N=220) using backward selection algorithm. The miRNA score was defined as the continuous survival function derived from the model, being afterwards validated by leave one out crossvalidation. Its predictive performance was tested by ROC analyses and patients were stratified into high and low risk Cox groups by the median of the miRNA score in the whole series (using also ROC analysis to check suitability of this cut-off point).

RESULTS

RESULTS I- Gene Expression Signatures with Prognostic Significance in Advanced Classical Hodgkin Lymphoma.

The purpose of this work was the identification of biological processes underlying treatment failure in advanced Hodgkin lymphoma (HL) patients, and the subsequent development of a quantitative RT-PCR based assay to be applied to routine formalin-fixed paraffin-embedded (FFPE) samples.

Taking into account the heterogeneity and peculiar composition of cHL tumors, this study identified gene subsets expressed either by the tumoral cells and the Hodgkin microenvironment, and showed that robust methodologies based on quantitative RT-PCR are suitable for expression profiling of tumors and can be easily applied to paraffin embedded samples allowing to analyze a limited number of selected genes in a single sample.

In addition, it identified functional signatures associated with treatment response and demonstrated the potential prognostic capacity of the developed assay finding a positive correlation between the expression of the proposed genes and treatment response (Molecular Risk Score_ MRS)

3.1 Identification of tumor and microenvironment signatures. TDB and MDB.

As mentioned, clinical and histological heterogeneity of classical Hodgkin lymphoma and the characteristic scarcity of HRS cells have prevented molecular studies. Despite the lack of a perfect model of HRS cells, established cHL cell lines are commonly used as a model system to characterize HRS cells biology while tumor samples represent a complex mixture of tumoral cells and the reactive immune infiltrate.

Comparison of gene expression profiles from 29 frozen tumoral advanced cHL samples and gene profiles from 5 cHL-derived cell lines (L428, L540, L1236, HDLM2 and HDMYZ) by supervised methods using T-test revealed 3.463 genes differentially expressed (FDR≤0.15). This analysis allowed the recognition of gene signatures attributable to both the neoplastic components:

a) ***Genes overexpressed in cHL cell lines (1.705 genes)*** were considered as putatively expressed by HRS cells and grouped in a database named Tumor Database ***(TDB).***
b) ***Genes overexpressed in the tumor samples (1.758 genes)*** were considered as characteristic from the microenvironment and grouped in another database named Microenvironment Database ***(MDB).***
c) The remaining not differentially expressed genes (FDR>0.15, 1,671 genes) were considered to be potentially expressed by both populations and included either in TDB and MDB in order to avoid loss of information.

Composition of TDB and MDB was very complex with a high number of genes that informed from functions and processes already known to be implicated in cHL pathogenesis. TDB contained genes known to be expressed by HRS cells, including cell cycle regulators, signaling, surface receptors and transcription factors. According to previous published data(Kuppers et al., 2003), genes such as TNFRSF8 (CD30) , GATA3, the tumor necrosis factor receptor family member RANK , and the metalloproteinase TIMP1(Fiumara et al., 2001) were present in TDB whereas MDB was mainly composed of genes involved in the immune response (STAT1, LYZ, SH2D1A and CTSL) (Sanchez-Aguilera et al., 2006).

As described above, the identified gene signatures were used to generate two different datasets containing the gene expression data from the 29 advanced cHL patients to be further analyzed.

3.2 TDB and MDB validation by immunohistochemistry.

Since the elucidation of the tumoral and microenvironmental signatures (TDB and MDB) was based on a bioinformatic approach (virtual or In silico microdissection) (Alizadeh et al., 2001) ,we wanted to confirm that obtained results were really informative and consistent. Thus, a limited number of genes were selected for validation at protein level using tissue microarray-based immunohistochemical techniques (TMAs) on an independent series of 142 cHL patient samples.

Validation markers were chosen based on their representative value for either tumoral or microenvironment signature and the availability of reliable antibodies for paraffin-embedded tissue. According to TDB composition, BCL2, BCL2L1, CASP3, MUM1/IRF4, CDK7 and CCNA2 were mainly expressed by the HRS tumoral cells, whereas genes present in MDB (STAT1, LYZ, SH2D1Aand CTSL) were expressed by reactive T cells, macrophages, reactive T cells and fibroblasts (Figure 3.1).

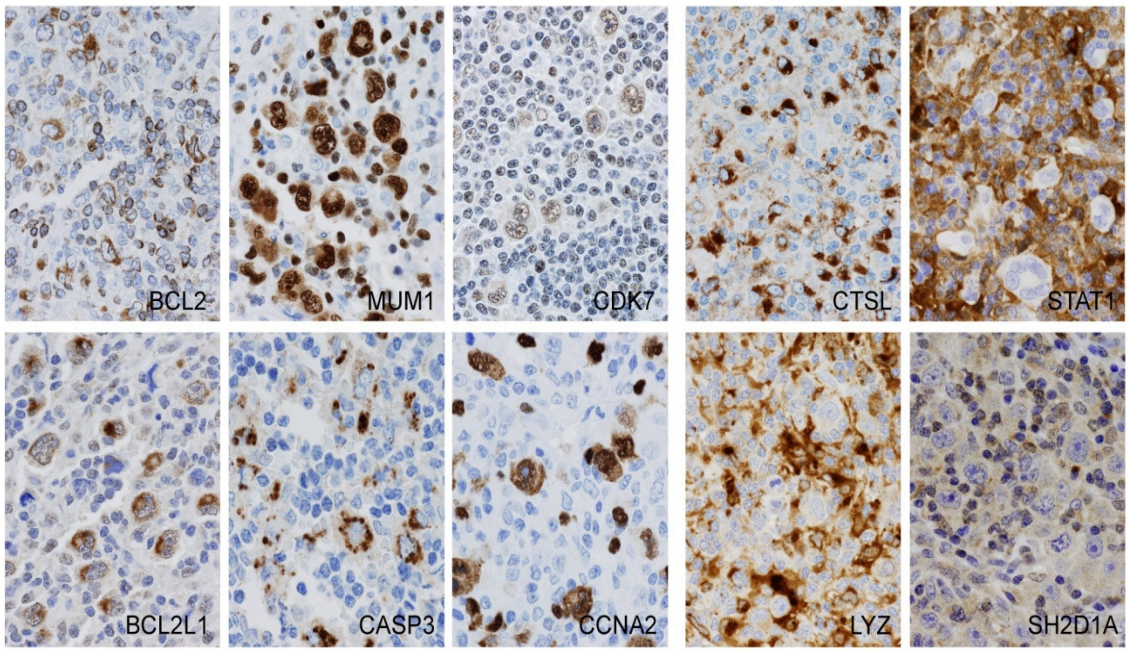


Figure 3.1 Immunohistochemical analyses of selected markers. Proteins in the six panels on the *left* correspond to genes included in the TDB, and their expression was mainly restricted to tumoral HRS cells (BCL2, MUM1, CDK7, BCL2L1, CASP3 and CCNA2). The four panels on the *right* correspond to genes included in the MDB, and the expression of the respective proteins was mainly restricted to fibroblasts, macrophages, and reactive T cells (CTSL, STAT1, LYZ, and SH2D1A) whereas tumoral cells are negative.

When the number of positive cHL samples for each marker was evaluated considering expression levels showed either by HRS cells or the microenvironment the same trend was observed (Figure 3.2). The absence in HRS cells from those markers considered to be expressed by microenvironment (CTSL; SH2D1A, STAT1 and LYZ) was clearly seen being negative for almost all cases. A more complex expression pattern was found in the reactive immune infiltrate component in agreement with the intrinsic complexity of cHL microenvironment.

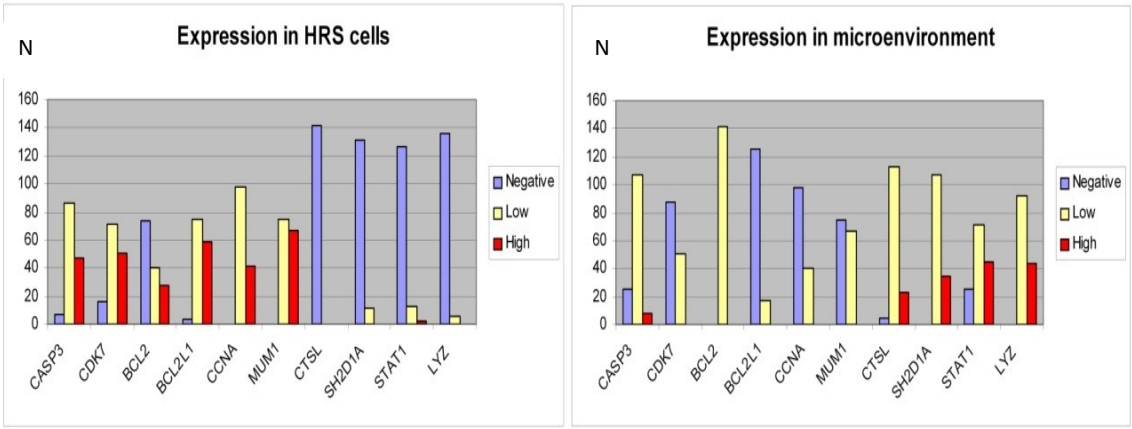


Figure 3.2 The histograms represent the number of positive cHL samples for each marker expressed by the HRS cells or the microenvironment. Expression levels codified as negative, low or high expression depending on signal intensity observed in the microscope. Total number of evaluated cases (N=142).

3.3 Pathway analysis (GSEA) of TDB and MDB

In order to clarify the complex composition of TDB and MDB, we reanalyzed both databases using a different bioinformatic approach able to look for pathways or corregulated genes, taking also into consideration the clinical outcome of patients (Favorable or good outcome versus Unfavorable or bad outcome) not considered in the first analysis. As explained in material and methods, favorable outcome (F) was considered in patients with sustained complete response (18 months) whereas unfavorable outcome (U) was considered in cases without complete response or patients with early relapses to therapy following previously published criteria(Carde et al., 2002)

The hypothesis that treatment response in cHL could be determined by a combination of factors related to both the neoplastic cells and their tisular microenvironment and might be associated with alterations in biological pathways rather than with randomly identified individual genes was tested by using Gene Set Enrichment Analysis (GSEA). Noteworthy,

comparisons just made on the level of individual gene lists clearly diverged when different statistical methods were used, thus prompting us to avoid this approach. Two independent GSEA analyses to either TDB or MDB with independent gene sets for each database were performed, looking for functional pathways overrepresented in the unfavorable outcome patient´s group (U). In this way, specific gene patterns by comparing expression data from two completely opposed biological situations (treatment responders (F) versus non-responders (U)) were identified.

As the number of analyzed samples was pretty low (29 patients) and taking into consideration their uneven phenotypic distribution, GSEA was performed looking for pathways associated with lack of maintained response in order to select an initial panel of potentially predictive genes to be included in an RT_PCR assay (TLDA assay). Thus, we focused our attention to the most relevant pathways within all the studied ones with independence of strict statistical significance (based on a biological descriptive interest).

3.3.1 Tumor TDB

TDB analysis showed that 15 out of the total 46 pathways evaluated were associated with unfavorable outcome (Table 3.1 and Table 3.3). Interestingly, two gene sets (G2/M and GS) were significantly associated ($p<0.005$) to treatment failure containing genes associated with the regulation of the spindle checkpoint such us Aurora Kinase A (AURKA), MAD2L1, BUB1B, BUB3, CHEK1, and CDK1. These results are in agreement with already previous published data regarding the identification of genes related with genome integrity checkpoints as involved in cHL progression and the observation of disrupted transition through mitosis in cHL Hodgkin cells (Sanchez-Aguilera et al., 2006).Additionally, G1 pathway, histones, chaperones , MAP Kinase signaling and mitochondrial signaling pathways (including genes implicated in apoptosis regulation like CASP3, BCL2 and BCL2L1) were enriched in the unfavorable group of patients highlighting the complexity and wide variety of processes that could be associated to therapy resistance.

Interestingly, the drug resistance pathway was also enriched in the unfavorable group of patients with the important presence of genes such TOP2A and MLH1, genes known to be involved in the transport and metabolism of doxorubicin (basis of ABVD standard chemotherapy regimen used for cHL patients).

Table 3.1. GSEA pathways related to unfavorable in Tumoral Database (TDB)

Pathway	Genes
G2M	CDC27, MDM2,CHES1,RPS6KA1,BRCA1,CCNF,WEE1,GADD45B,EP300,CDKN2D,CDC25A,MAD1L1,CDC14A,CDC20,YWHAH, YWHAQ,PRKDC,KIF2C,BIRC5,STK6,CCNH,BUB1,MAPRE1,BUB3,CENPE,CHEK1,CCNA2,CENPA,NBS1,BUB1B,CDC2,CENPF, NMUA1, HMMR, NEK2, MAD2L1
GS	PRKACA,GAS2,TFDP2,DDX11,E2F2,GNAS,MYC,POLR3D,CDC20,CDC45L,CDC7,MYBL2,CCNE2,CCNH,CDC6,CDK7,CCNA2,PNA, GSPT1,CDC2,RAMP
G1	CDK6,MDM2,RBL1,TGFB2,CDKN2B,ABL1,CDK2,HDAC1,CDC25A,SKP2,CCND2,DHFR,SKP1A,CKS2,CDK5,GSK3B,CDK4,NPM1, TFDP1,BCCIP,CCNH,CDN2C,GSPT1,CDC2
Histone	HDAC4,HDAC5,HDAC11,HIST2H2BE,HIRA,HUS1,HIST1H1C,HIST2H2AA,MYST2,HDAC1,H2AFV,HIST1H2AC,HDAC2,HIST1H4H, NASP,H2AFZ,HDAC8,HDAC7A,HIST1H3D,H2AFX,H1F0
Chaperone	CDC37,HSPB3,DNAJB2,HSF1,DNAJB12,DNAJC11,DNAJB1,DNAJC3,DNAJC13,HSPA4L,BAG5,DNAJB6,HSPA2,DNAJA3,DNAJB5, HSPCB,HSPA1L,HSBP1,HSPH1,DNAJC8,HSPA8,DNAJA1,DNAJA2,HSPCA,HSPA9B,HSPA4
Drug resistance	CYP3A4,ABCC5,SMPD1,ABCC1,UHRF1,ERCC2,RALBP1,ABCB1,PSME4,ACO1,RPL11,ABCC2,APEX1,GLRX2,MSH2,SOD1,GSTP1, ABCC3,TYMS,H2-ALPHA,PSMC1,MLH1,RRM2,TOP2A,DCK
Metabolism	MAP4K1,HRAS,PTK2,SOS1,MAPK3,PAK1,HGF,ACTA1,MET,STAT3,RAP1A,CRKL,ELK1,RAF1,PIK3CA,MAP2K1,GRB2,RAP1B,PTPN11
Epidermal Growth Factor	HRAS,SOS1,MAP2K4,EGFR,JAK1,MAPK3,SHC1,STAT5A,STAT3,ELK1,PRKCA,RAF1,PIK3CA,STAT1,CSNK2A1,MAP2K1,GRB2
ERK pathway	RPS6KA1,HRAS,SOS1,NGFR,DPM2,EGFR,GNAS,MAPK3,IGF1R,MYC,SHC1,STAT3,ELK1,RAF1,MAP2K1,PPP2CA,GRB2
Mitogen-activated protein kinase	MAPK11,MAP3K6,MAP4K1,MAP2K3,MAP2K5,MAP3K4,RPS6KA1,RIPK1,MAPK12,MAPK10,MAP3K14,HRAS,MAP3K8,TGFB2, MAP2K4,NFKB1,DAXX,RELA,MAPK3,PAK1,MYC,MAX,SHC1,MAPKAPK3,MAP4K4,ELK1,RAF1,MAPK14,MAP3K7,STAT1,BRAF,MAPKAPK5, MAP2K1,MAPK9,GRB2,MAPK6
Tumor Necrosis factor 1 (TNFR1)	DFFA,DFFB,TNFRSF8,TNFRSF5,TNF,RIPK1,MAP2K4,CRADD,PAK1,LMNA,TNFRSF10B,MAP3K7,TTRAP,PRKDC,FADD,LMNB2, CASP3
Integrines	HRAS,VCL,PTK2,SOS1,BCAR1,ZYX,CSK,MAPK3,RCK1,SHC1,ACTA1,PPP1R12B,RAP1A,CRKL,RAF1,MAP2K1,GRB2
Mitochondrial	DFFA,DFFB,PDCD8,BAX,BCL2L1,BCL2L2,BCL2,BIRC3,BAG5,BIRC4,CASP9,BAG1,CASP7,BAK1,BIRC2,CYCS,CASP3

3.3.2 Microenvironment MDB

MDB analysis

GSEA analysis of MDB revealed 13 pathways associated to unfavorable outcomes with T-cells, monocyte/macrophage and dendritic cells pathways as the most enriched gene sets. These pathways included genes reporting from specific T-cell populations, thus highlighting the important role of the microenvironment (Table 3.2).

Table 3.2 GSEA pathways related to unfavorable outcome in Microenvironment Database (MDB)

Pathway	Genes
T-Cells	CR1,ADAM17,FURIN,CD4,TRA,CD1C,MAL,MLL,PRKCG,CD3G,PRKCH,FYN,CD3E,LEF1,DPP4,CD69,CD3D ,SH2D1A,TXN,CD8B1,RAB27A
Monocyte/ macrophage	VDR,PTGS1,ITGAM,PECAM1,SIRPB1,ANPEP,CD4,SELL,MMP9,ITGB2,SELPLG,SDC2,CD44,CTSS,ITGA4,LCP1, ALDH1A1,STAT1,LYZ,TXN,ATP6AP2,RAB27A
Dendritic cells	TNFRSF11A,CR1,LILRB1,CD80,LILRB4,SIRPB1,S100A1,LAMP3,SOCS1,LTB,S100A10,CD83,CD86,S100A8,CYP27B1, TNFSF11,S100A11,S100B,CYP24A1,S100A3,RSN
T cytotoxic cell	THY1,CD2,TRA,ITGB2,CD8A,PRF1,GZMB,HLA-A,CD3G.CD3E,LEF1,CD3D,CD8B1
T cell receptor	JUN,NFKBIA,NFATC4,FOS,VAV1,PTPN7,TRA,PIK3R1,ZAP70,MAP2K4,NFKB1,NFATC3,LCK,RELA,CD3G,FYN, MAPK3K1,CD3E,PRKCB1,PPP3CB,CD3D,PIK3CA,GRB2
Neutrophils	ITGAM,PECAM1,ANPEP,NR3C1,FCGR2A,MMP9,FPR1,SELPLG,BST1,FCGR1A,PAK1,LYN,BIRC3,SERPINA1,LCP1 ,CD69,EMR3
FAS ligand	JUN,DFFB,CASP10,MAP2K4,CFLAR,PTPN13,DAXX,PAK1,MAP3K1,ARHGDIB,CASP7,MAP3K7,CASP3
Fibroblast	FGFR1,FGF3,TACR1,SHOX2,FRS2,THY1,ST5,DDR2,MICB,FGFR4,FGFR3,GTF2IRD1,OPTN,FGF7,FGF11,SDC2,ECHS1 ,ACVR1,FGF19,FGF13,BLNK,CTSL,SDCBP,GRB2,AGC1
IL2RB	JAK3,IL2RA,FOS,SYK,BCL2L1,SOCS3,PIK3R1,SOCS1,IL2RB,CFLAR,IL2RG,BCL2,PIK3CA,GRB2
TH1	IL2RA,CXCL12,MMP2,IL18,MMP9,IL18R1,IFNGR2,CD86,IFNGR1,IFI30,IFI16
B cell receptor	JUN,RAFTLIN,BTK,NFATC4,FOS,SYK,VAV1,CD79A,CD79B,NFATC3,LYN,MAP3K1,PRKCB1,MAPK14,PPP3CB,BLNK,GRB2
Mastocytes	GAB2,ADAM17,PLD2,FCGR2B,APC,PRKCD,FCER2,VEGF,PTK2,LCP2,PLSCR3,FCGR2A,MMP9,USF2,SELPLG,BST1, ATP2A3,ICAM3,AREG,LYN,PLSCR2,PRKCB1,PLSCR1,GRB2
Natural Killer cells	TGFB3,SYK,IL18,VAV1,CCR7,CD4,CCL3,PIK3R1,IL18R1,IFNGR2,PTK2B,PAK1,CXCR4,IFNGR1,HLA-A,CD69,PIK3CA

3.4 Selection of genes with potential prognostic capacity and TLDA design format 64.

As shown in chapter 3.3, GSEA analysis revealed specific gene patterns identified after comparison of gene expression data from two completely opposed biological situations (*treatment responders (F) versus non-responders (U)).* These results showed that supervised methods are useful when asking for specific biological questions between phenotypes and verified the initial hypothesis that treatment response in cHL was determined by a combination of factors related to both the neoplastic HRS cells and their tisular microenvironment.

- Thus, Genes overexpressed by the neoplastic HRS cells (indeed confirmed by TMA analysis) comprised:

a) Genes related with signaling/ apoptosis

b) Cell-cycle regulatory genes, mainly related with G2/M transition, and the regulation of spindle checkpoint.

- Additionally, genes that reflected specific cell populations of the immune response were identified, such as specific T-cells and monocyte/macrophage cells.

These genes could reveal as useful prognostic markers; and prompted the design of a TLDA assay platform based on quantitative RT-PCR technology that was optimized to be applied to FFPE tissues (see material and methods for more details). A general overview of the experimental approach followed is summarized in Figure 3.3

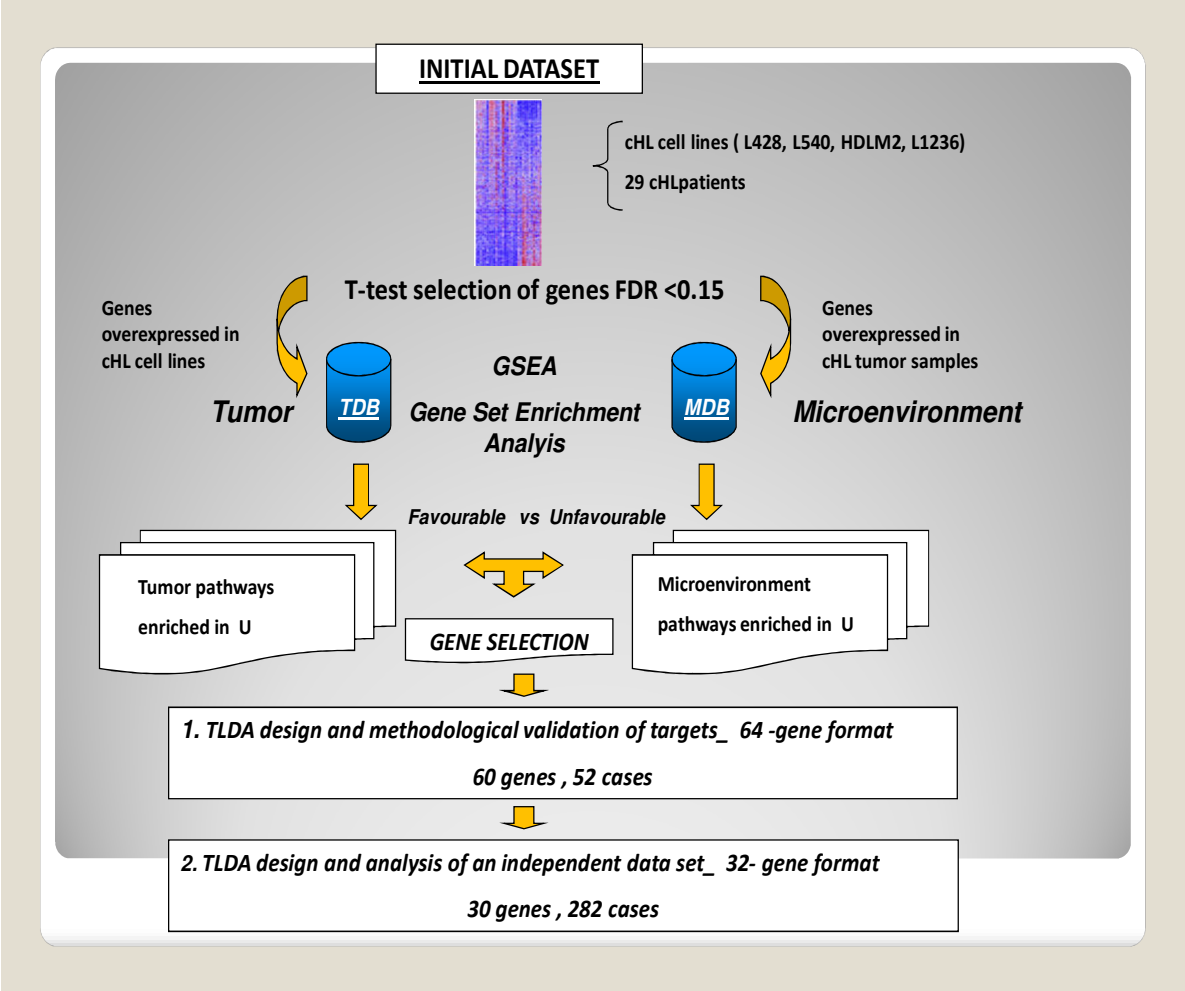


Figure 3.3 Experimental design. After characterization of the tumor (TDB) and microenvironment (MDB) databases, GSEA allowed the selection of genes to be included in TLDA assays. F, favorable; U, unfavorable.

Initially, an assay that allowed the study of 64 genes was designed. From the most highly enriched pathways in the TDB and MDB, 56 representative genes (top score) were chosen (Table 3.3) for inclusion in TLDA assay. They were selected based on the strength of their ability to discriminate patients with a good or bad treatment response (enrichment score of the genes in each pathway) and the biological relevance of the pathway. In addition, 4 other

genes were also included in the platform: *BCL2, IRF4, and FOXP3* based on their previously prognostic significance (Alvaro et al., 2005; Rassidakis et al., 2002a) and *TNFSRF8 (CD30*) due to its expression in HRS cells

In this first model, 4 endogenous genes (HMBS, GUSB, TBP and GAPDH) were included in order to select the two best ones for data normalization. GeNorm algorithm (Vandesompele et al., 2002)(available inside Integromics package) detected HMBS and GUSB as the most stable ones, thus being selected for further analysis (see material and methods for details).

Table 3.3 Top GSEA enriched pathways associated with unfavorable outcome and highest scoring genes

GENE SETS ENRICHED IN TDB	NES	Nom p-val	GENES TLDA assay 64		
G2M PATHWAY	-1.67	0.016	AURKA BUB1B BUB3 CCNH CDC2	CENPE CENPF CHEK1 CSE1L HMMR	MAD2L1 MAPRE1 NBS1 NUMA1
GS PATHWAY	-1.50	0.052	RAMP. CCNA2 CCNE2	CCNH CDC6 CDC2	CDK7
G1 PATHWAY	-1.35	0.145	BCCIP	CCNH	CDKN2C
HISTONE PATHWAY	-1.26	0.151	H1F0	H2AFX	HIST1H3D
CHAPERON PATHWAY	-1.24	0.269	DNAJA2 HSP90AA1	HSPA4	HSPA9B
DRUGS	-1.23	0.204	DCK MLH1	RRM2 TOP2A	TYMS
MET PATHWAY	-1.20	0.203	GRB2	PTPN11	
MAPK PATHWAY	-0.83	0.736	GRB2 MAP3K7	MAPK14 MAPK6	MAPK9
MITOCHONDRIAL PATHWAY	-0.77	0.883	BCL2 CYCS	BCL2L1	CASP3
GENE SETS ENRICHED IN MDB	**NES**	**Nom p-val**	**GENES TLDA assay 64 MODEL**		
T-CELLS	-1.50	0.063	CD3D SH2D1A	CD8B1	DPP4
MONOCYTE/MACROPHAGE	-1.39	0.161	ALDH1A1 LYZ	ITGA4 STAT1	LCP1
TCYTOTOXIC PATHWAY	-1.13	0.367	CD3D	CD8B1	
TH1	-0.80	0.656	IFI16	IRF4	
DENDRITIC PATHWAY	-0.78	0.745	RSN		

NES_ Normalized Enrichment Score; Nom p val_ Nominal probability

The final set of genes included a balanced representation of those associated with the tumor cells and microenvironment being intrinsically associated with the pathogenetic characteristics of cHL. As commented, G2-M pathway included several important genes associated with the regulation of the spindle checkpoint, such as AURKA, MAD2L1, BUB1B,

BUB3, CHEK1, and CDK1 (CDC2). Most of these genes are known to be involved in chemoresistance and some of them are considered as potential therapeutic targets. RSN, an intermediate filament-associated protein highly expressed in the HRS cells that may contribute to neoplastic transformation in Hodgkin´s disease (Hilliker et al., 1994),was also found in this study.

Interestingly, there was also a representation of cell cycle regulators specifically related with drug metabolism including TOP2A (cellular target of Adriamycin), TYMS, and RRM2, which have been associated with drug resistance in different tumor models (Farrugia et al., 2003; Wang et al., 2004). In addition, the presence of the chaperones (frequently overexpressed in cancer cells) in the selected genes was taken into account with HSP90AA1, HSPA4 and HSP90A.

On the other hand, microenvironment signature was overrepresented by genes involved in the Th2 immune response, such as cytotoxic and regulatory T cells (FOXP3), already described to independently predict patient survival (Kelley et al., 2007).

3.5 TLDA analysis and relationship with treatment response. The Integrated Risk Score.

This first TLDA assay model was applied to an initial set of patient samples obtaining adequate RT-PCR profiles (Ct value <35) for 43 out of the 52 FFPE cases analyzed (82.7%). Differences in the distribution of classic clinical variables (age, gender, stage) between the two groups were not statistically significant. Although the main aim was the development and technical optimization of an exploratory tool for cHL tumor model, normalized expression levels (ΔCt) (using HMBS and GUSB) of individual selected genes were on average overexpressed in the unfavorable group. Additionally, T-test comparison among outcome groups of all genes grouped into the Integrated Risk Score was significant (p=0.025) with higher expression levels in the Unfavorable outcome group (Figure 3.4).

Due to the limited number of cases analyzed (43 cases); some overlap of the Integrated Risk Score values was observed. However, it confirmed the potential prognostic value of the genes included in this preliminary assay and prompted to study a larger series of patients for validation and development of a predictor model. In fact, ROC analyses showed a set of genes with significant p values ($p<0.05$): BCCIP; CASP3, CCNE2, CTSEL1, CTSL, CYCS, DCK, DNAJA2, HSP90AA1, HSPA4, ITGA4, LYZ, RSN and TYMS. A logistic regression model integrating these genes had an overall 100% predictive accuracy for the series (χ^2 =51.049; $p< 0.001$).

When leave-one out cross validation (CV) was applied, classification accuracy decreased to 69.5 % thus highlighting the need of further validation studies.

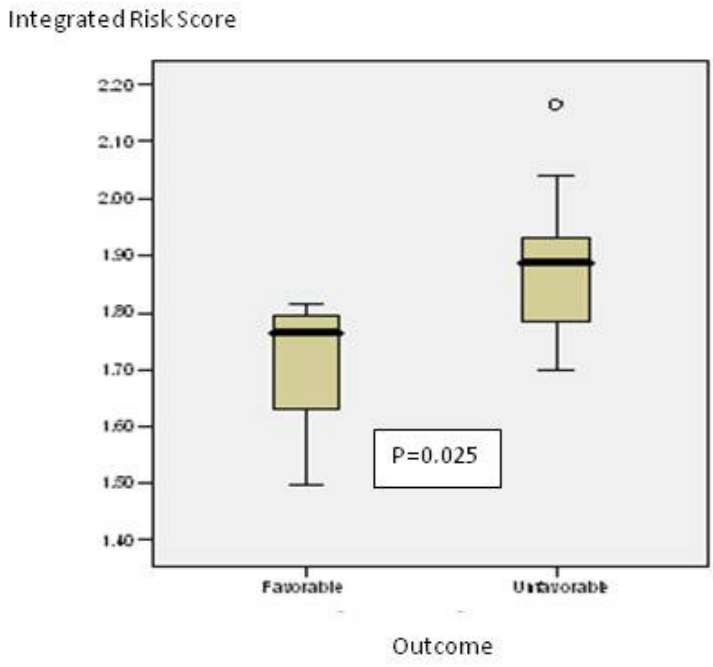


Figure 3.4 Box-plots of the Integrated Risk Score. It was defined as the logarithmic mean of expression levels of all genes included in the TLDA assay format 64 (metagen), in each patient outcome group. *Black line within the box* represent the median of each. Box group values are within the first and third quartiles. *Whiskers* (error bars) extend from each end of the box to the adjacent values in the data and represent the most extreme values within 1.5 times the interquartile range from the ends of the box. *Symbols*, outliers.

Despite the fact that RT_PCR is a sensitive and precise tool for determining gene expression in tissues, some cases yielded low quality Ct results (Ct>35) due to the inherent degradation and small amount of RNA extracted from FFPE tissues. The inclusion of a pre-amplification step in the analysis was tested in 13 cases, including 2 previously considered to be low quality. Adding a pre-amplification step (PreAmp Master Mix , Applied Biosystems) all cases yielded good results (100% Ct<35) showing a mean reduction in Ct values of 4,49 cycles (mean Ct value= 25,62± 4,76 using PreAmp versus 30,1± 3,40 without preamp) without introducing any amplification bias. Thus, pre-amplification was included in all analyses did in the second part of this study to ensure results' reliability.

3.6 Development of the predictor model: Molecular Risk Score

The genes included in the assay (TLDA format 32) used to develop the final predictive model (TLDA assay 32) were selected from the initial ones, chosen on the basis of their prognostic capability (evaluated by logistic regression and ROC analysis as previously commented) and their capacity to represent biological functions identified as relevant in cHL pathogenesis (Sanchez-Espiridion et al., 2009). Additionally, the strength and consistency of primer and probe performance was taken into account (Sanchez-Espiridion et al., 2009). This selection consisted of 30 genes that were tested in 282 samples and adequate RT-PCR profiles were obtained in 262 of 282 cases (93.9%). The samples were afterwards randomly split and assigned to either training (n= 183) or validation series (n=79) (See material and methods).Normalized expression levels (ΔCt) of individual genes varied considerably among samples. However, when univariate regression analysis was applied to the normalized expression data, a set of 20 genes were found to be significantly associated with treatment response ($p < 0.05$) (Table 3.4).

These 20 genes identified by univariate logistic regression analysis were further studied by cross-validation to improve the initial selection. When cross validation methods were applied, the genes most frequently found in prognostic models consisted on a panel of 11 genes included in the final model: BCL2, BCL2L1, CASP3, HMMR, CENPF, CCNA2, CCNE2, CDC2, LYZ, STAT1 and IRF4.

Table 3.4 Univariate logistic regression analysis of significant genes and crossvalidation analysis

			Cross-validation algorithms						
			SVM		DLDA		KNN		Total CV count (%)
Gene name	p value	Hazard ratio (95% CI)	CV	%	CV	%	CV	%	
HMMR	0.000	1.160 (1.082- 1.243)	136	90.67	148	98.67	149	99.33	96.22
LYZ	0.000	0.800 (0.738- 0.866)	131	87.33	149	99.33	147	98.00	94.89
CENPF	0.005	1.148 (1.041- 1.265)	22	14.67	111	74.00	142	94.67	61.11
CASP3	0.014	1.104 (1.020- 1.194)	26	17.33	115	76.67	134	89.33	61.11
BCL2L1	0.001	0.836 (0.751- 0.931)	26	17.33	98	65.33	129	86.00	56.22
STAT1	0.000	0.810 (0.748- 0.878)	21	14.00	103	68.67	125	83.33	55.33
CCNE2	0.001	1.152 (1.057- 1.255)	21	14.00	98	65.33	127	84.67	54.67
CCNA2	0.000	1.243 (1.137- 1.358)	18	12.00	102	68.00	120	80.00	53.33
BCL2	0.000	1.191 (1.089- 1.303)	20	13.33	88	58.67	123	82.00	51.33
IRF4	0.020	0.872 (0.777- 0.978)	13	8.67	63	42.00	108	72.00	40.89
CDC2	0.000	1.123 (1.052- 1.199)	20	13.33	60	40.00	102	68.00	40.44
MAD2L1	0.001	0.853 (0.779- 0.933)	15	10.00	49	32.67	98	65.33	36.00
GRB2	0.000	0.868 (0.818- 0.921)	9	6.00	52	34.67	96	64.00	34.88
TYMS	0.002	1.167 (1.058- 1.287)	12	8.00	46	30.67	92	61.33	33.33
DNAJA2	0.004	0.849 (0.760- 0.947)	9	6.00	37	24.67	37	24.67	30.89
H2AFX	0.000	0.884 (0.831- 0.941)	9	6.00	39	26.00	90	60.00	30.66
IFI16	0.000	0.823 (0.740- 0.916)	9	6.00	37	24.67	90	60.00	30.22
BUB3	0.001	0.838 (0.752- 0.934)	9	6.00	36	24.00	90	60.00	30.00
ALDH1A1	0.007	0.878 (0.799- 0.965)	9	6.00	37	24.67	37	24.67	30.00
HISTH3D	0.000	0.876 (0.827- 0.928)	9	6.00	35	23.33	89	59.33	29.56

Note. Genes with total crossvalidation (CV) count >40 % were selected. SVM_supprt vector machine; DLDA_diagonal linear discriminant analysis. NN_ K-nearest neighbor

Ideally, a model based on relative contributions of genes clustered on their respective functional groups would be more consistent than one derived for individual genes. Then, these final 11 selected genes were classified into biological functional groups on the basis of their known biological relationship and their correlated expression estimated by the Pearson correlation coefficient (Table 3.5). Pearson correlation coefficient was significant between the genes included in each of the signatures ($p<0.001$) apart from IRF4 that was included as independent predictive gene since there were neither distinct functional relationship nor statistical correlation with other genes or pathways.

Individual genes were weighted using diagonal linear discriminant analysis (DLDA) into their corresponding functional pathways, and then subsequently correlated with the clinical patient outcome using multivariate logistic regression. The multivariate logistic regression analysis correlating the mentioned functional pathways with the clinical patient outcome showed that Cell Cycle and Apoptosis terms were associated with an unfavorable outcome whereas IRF4 and Macrophage activation seemed to have protective effects. The optimized final model was based on relative contributions of each of the four gene functional terms, as described by the following equation:

Molecular Risk Score = exp (fx)/ [1+exp (fx)], where f(x) =

Constant (-0-913) + (0.401 x Apoptosis) + (0.284 x Cell Cycle) + (-0.301x Macrophage Activation) + (-0.143 x IRF4).

This equation generated a continuous probability function defined as ***Molecular Risk Score (MRS)*** to treatment failure that was calculated for each case and ranged from 0.06 to 0.813 values. In order to stratify patients according to their MRS value we used ROC analysis to define a threshold dividing the series into high-risk group (MRS≥0.3) and low-risk group (MRS≤0.3) (Figure 3.5).

Table3.5 Gene expression correlation analysis of genes grouped in pathways

Macrophage activation		**LYZ**	**STAT1**			
LYZ	Pearson correlation	1				
	p value					
	N	263				
STAT1	Pearson correlation	.559(**)				
	p value	**0.000**				
	N	263	263			
Apoptosis		**BCL2**	**BCL2L1**	**CASP3**		
BCL2	Pearson correlation	1				
	p value					
	N	263				
BCL2L1	Pearson correlation	.293(**)	1			
	p value	**0.000**				
	N	263	263	1		
CASP3	Pearson correlation	.387(**)	.132(*)			
	p value	**0.000**	**0.032**			
	N	263	263	263		
Cell Cycle		**CENPF**	**CCNA2**	**CCNE2**	**CDC2**	**HMMR**
CENPF	Pearson correlation	1				
	p value					
	N	263				
CCNA2	Pearson correlation	.474(**)	1			
	p value	**0.000**				
	N	263	263			
CCNE2	Pearson correlation	.688(**)	.433(**)	1		
	p value	**0.000**	**0.000**			
	N	263	263	263		
CDC2	Pearson correlation	.595(**)	.525(**)	.535(**)	1	
	p value	**0.000**	**0.000**	**0.000**		
	N	263	263	263	263	
HMMR	Pearson correlation	.615(**)	.497(**)	.609(**)	.596(**)	1
	p value	**0.000**	**0.000**	**0.000**	**0.000**	
	N	263	263	263	263	263
IRF4	-					

***** significant correlation**

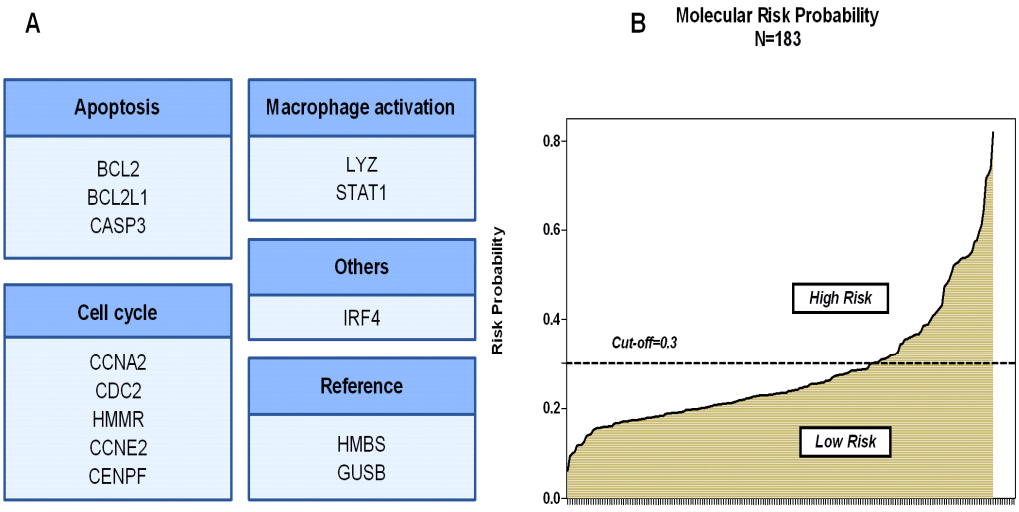


Figure 3.5.Panel of 11 genes and the molecular risk algorithm. A) The molecular risk algorithm was based on the relative contributions of each of the four gene functional groups from the tumoral HRS and their reactive microenvironment as follows: Molecular Risk Score = exp (fx)/ [1+exp (fx)], where fx = (-0.319) + (0.401 x Apoptosis) + (0.284 x Cell Cycle) + (-0.301 x Macrophage activation) + (-0.143 x IRF4). Coefficients were derived from a multivariate analysis, in which positive values indicate that a higher level of expression is correlated with a worse outcome, and negative coefficients indicate that a higher level of expression of the pathways is associated with better outcomes. B) Molecular Risk Score as a continuous function was used to set a threshold for patient´s stratification by ROC analysis. Patients were stratified according to the levels of the molecular risk score into low-risk (< 0.3) and high-risk (≥0.3) groups.

3.7 Validation of the Molecular Risk Score.

3.7.1 T test comparison: Favorable versus Unfavorable patients.

The MRS showed significant differences between outcome groups either in training or validation datasets and predicted treatment response with an accuracy of 68.9 percent in the estimation dataset and 70 percent in the validation dataset. As expected, patients with unfavorable responses showed, on average, higher expression levels than those with favorable responses (Figure 3.6). Pathways of MRS score were indeed differentially expressed among the Outcome groups at significant values (with exception of IRF4 p=0.074). In agreement with estimated coefficients Monocyte/ Macrophage activation and IRF4 were overexpressed in the favorable outcome groups (thus, with a potential protective effect) whereas Apoptosis and Cell Cycle pathways showed higher expression in the unfavorable group of patients(Figure 3.7).

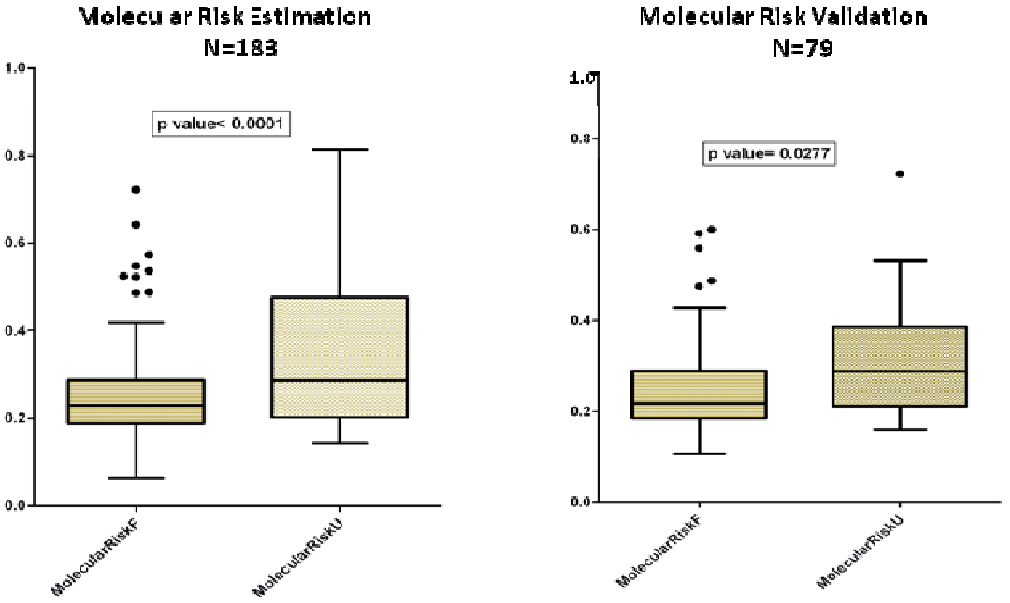


Figure 3.6 MRS differences in validation and estimation sets of samples between favorable (F) and unfavorable (U) outcomes. Significant p values were obtained from comparisons using t-tests statistics.

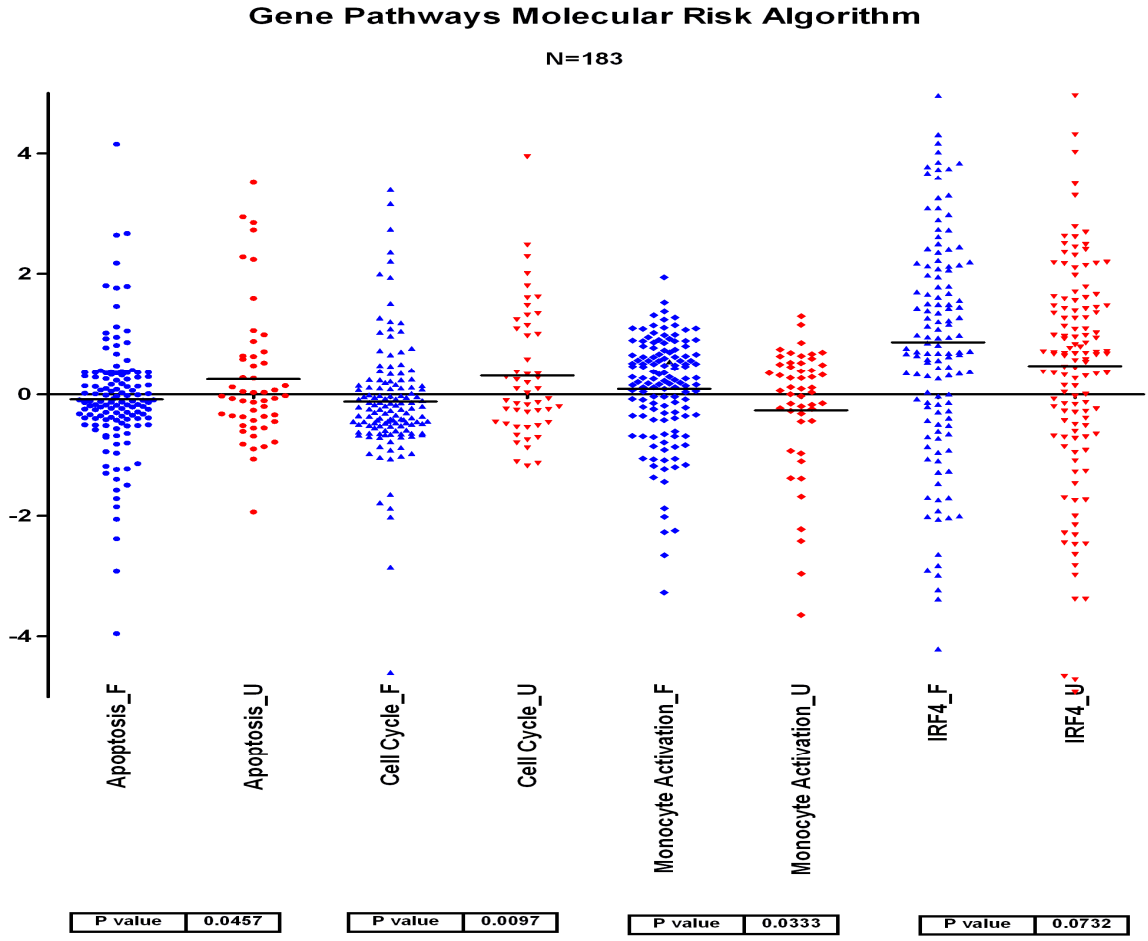


Figure 3.7 T-test comparisons of biological functional groups included in the MRS between patients with favorable (F) and unfavorable outcome (U).

3.7.2 Survival Analysis. Is our model valid?

The survival estimates of failure-free survival (FFS) in patients from estimation (N = 183) and validation (N = 79) sets after classification into risk groups according to their MRS values (high risk versus low risk) gave significant results in both datasets: 5-year FFS probabilities 75.6% versus 45.9 %,(log rank statistic $p \approx 0.000$) in estimation and 5-year FFS probabilities 71.4% versus. 43.5%, (log rank statistic $p < 0.004$) in validation set respectively. However, when overall survival (OS) was used as clinical endpoint not significant results were obtained for the validation dataset (log rank test p= 0.38). This was probably due to the low number of events (deaths) presented in the validation dataset, thus indicating MRS score was not valid for this clinical endpoint. In addition, when all the entire series was analyzed significant results were obtained either for FFS or OS (Figure 3.8).

It should be mentioned that main´s clinical endpoint for this study was first treatment outcome (Favorable versus Unfavorable) used for GSEA analysis and MRS development; and secondly, FFS was considered for survival analyses. As commented above, the risk score had a prognostic effect on FFS in the small independent validation dataset, but not significant effect on OS in the validation set was observed. Results obtained for all entire series were not taken into account because of a biased statistical analysis and over fitting, thus being just used for graphical representation and description of the series.

These clinical endpoints (outcome and FFS) were selected because overall survival (OS) involves three other factors that have to be considered separately and were not considered in this work (Hasenclever and Diehl, 1998) :

1. The ability of the initial treatment to control the disease.
2. An appreciable second chance of a cure with salvage treatment.
3. Deaths due to the late toxicity or disorders unrelated to Hodgkin´s disease in patients with continuous complete remissions,

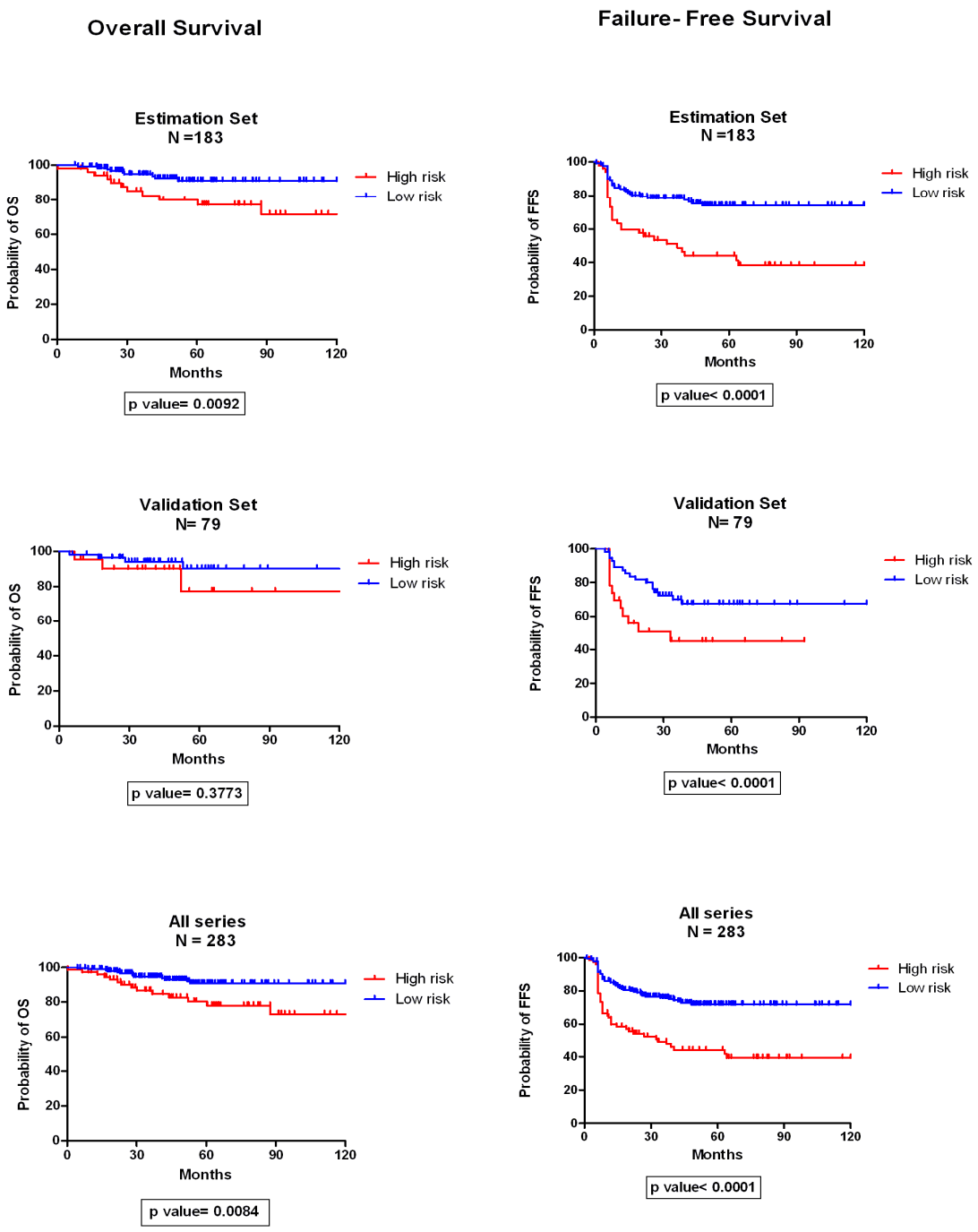


Figure 3.8 Overall and failure-free survival (OS and FFS) analyses using Kaplan-Meier and log-rank test in the estimation and validation sets, and in the whole series of patients. Patients are stratified by the MRS values (high risk ≥0.3 versus low risk w<0.3).

3.8 Integrative Model. Molecular Risk Score and the traditional clinical variables.

Next step was to evaluate whether the MRS was adding any prognostic value to the already established and commonly used IPS. No relation was observed between the individual IPS factors and the molecular risk score (low- and high-risk groups), with neither interaction nor confusion factors between variables observed ($\chi 2$ test non significant).

In a multivariate Cox proportional hazards model using FFS as dependent variable; just MRS was significant (Table 3.6 A). This result was striking, and probably reflecting the limitations of this traditional clinical score (Hasenclever and Diehl, 1998).Thus, we analyzed each one of the different variables comprising IPS (Hemoglobin, Albumin, Leucocytosis, Lymphopenia, Age, Stage IV and Gender) finding Stage IV also significantly associated with FFS in the entire series when a backward stepwise selection Cox regression model was applied (Table 3.6.B).

Table 3.6 The Molecular Risk Score (MRS) and the IPS variables

A. International Prognostic Score (IPS) and Molecular Risk Score (N = 262)

	p value	Hazard ratio (95% CI)
Molecular Risk Score	**0.000**	24.362 (6.268-94.691)
IPS	0.205	1.111 (0.944-1.309)

B. IPS variables and Molecular Risk Algorithm (N = 262)

	p value	Hazard ratio (95% CI)
Molecular Risk Score	**0.000**	24.715 (6.804-89.768)
Hemoglobin (< 10.5 g/dl)	0.352	1.243 (0.787-1.963)
Albumin (< 4 g/dl)	0.504	1.157 (0.755 - 1.772)
Leucocytosis (≥ 15,000/mm^3)	0.256	1.312 (0.821 - 2.095)
Lymphopenia (< 600/mm^3)	0.555	0.820 (0.425 - 1.583)
Age (≥ 45 yr)	0.369	1.227 (0.785 - 1.916)
Stage IV	**0.041**	1.552 (1.018 - 2.360)
Male gender	0.753	0.936 (0.618 - 1.416)

Finally, and integrative model (Stage IV and MRS) was developed in a Cox regression model able to identify a subgroup of patients with a very bad outcome when patients series was divided into quartiles. (Fourth quartile: 5-year FFS of 24.3% p<0.0001) (Table 3.7 and Figure 3.9).

Table 3.7 Cox regression analysis of variables from Integrative model

Integrative Cox model (N = 262)		
	p value	Hazard ratio
Molecular Risk Score	**0.000**	23.782 (6.041 - 94.340)
Stage IV	**0.025**	1.409 (1.044 – 1.900)

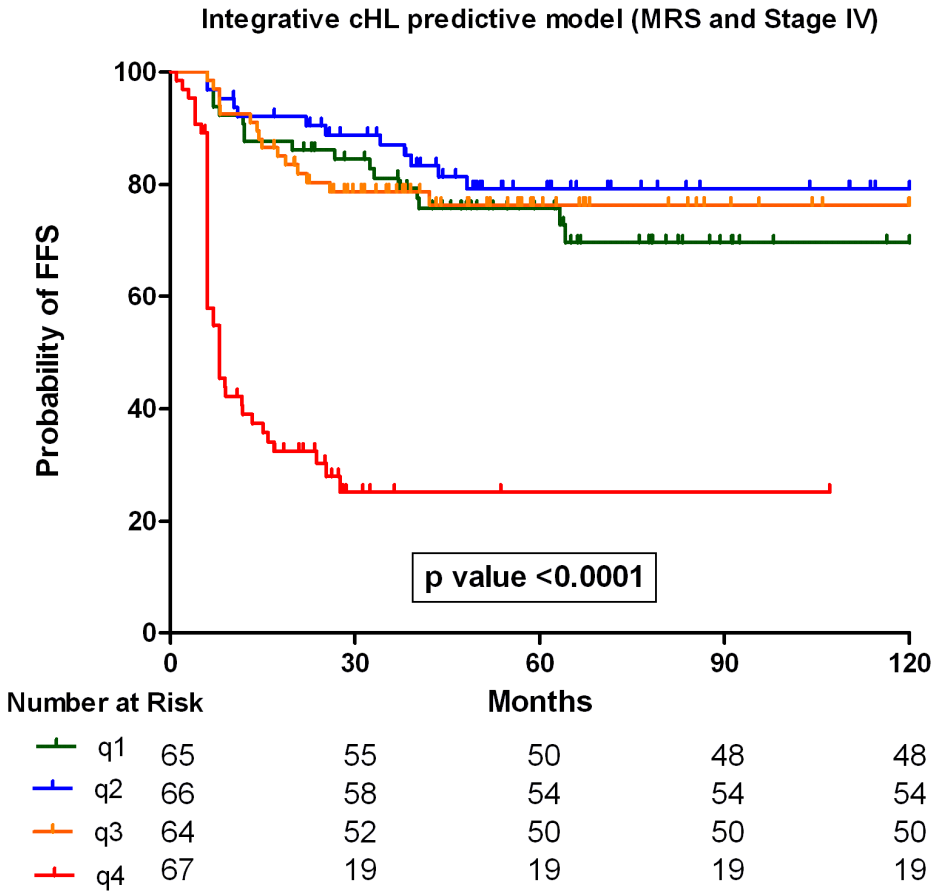


Figure 3.9 Integrative Model Risk of cHL. The final Cox model integrated the molecular risk score (MRS) and Stage IV. Patients in the quartiles 1, 2 and 3 had comparable FFS rates at 5 years (5-year FFS: 76.4 %, 79,3% and 69,7% respectively) while patients in the quartile 4 showed a 5-year FFS of 24,3 percent (p<0,001)

Moreover, these results showed that MRS biological model was independent of and complementary to the conventional International Prognostic Score using multivariate Cox proportional hazards analysis.

3.9 Conclusion and final remarks.

As a summary, in this first part of the work:

- Using two series of training and validation sets of advanced classical Hodgkin lymphoma patients an IPS-independent Molecular Risk Score (MRS) has been developed.

- The MRS described consists on a quantitative RT-PCR assay for patients with advanced cHL, incorporating a limited number of genes and pathways from tumor and microenvironment cell components, designed for its application to formalin-fixed paraffin-embedded routine samples.

- The MRS integration with clinical stage IV makes possible to derive an Integrative model able to identify a subgroup of patients with very bad outcome (a quartile with 24, 3 % FFS at 5 years).

- The roles that both tumor components, HRS cells and their reactive microenvironment, are playing in patient outcome need to be validated in further studies, since other recent studies are showing a relationship between the presence of tumor associated macrophages and a poorer patient outcome.

Some comments and their respective answers than can be concerned are listed below.

1. It could be thought that MRS results were biased_ However; different datasets were used to develop and test the model and its genes.
2. It could be thought that results are overoptimistic_ To avoid this, crossvalidation algorithms were applied
3. What are they adding to the classical variables? _The model is adding biological information (reporting from specific genes and pathways) underlying treatment resistance.

RESULTS II _MicroRNA signatures with Prognostic Significance in Advanced Classical Hodgkin Lymphoma.

MicroRNAs (miRNAs) are small noncoding RNA molecules that regulate gene expression at the post-transcriptional level through mRNA interference. They are currently emerging as important regulators of biological processes such as proliferation and differentiation, and their role in cancer pathogenesis has been demonstrated in several malignancies.

Some attempts have been made to elucidate specific miRNA signatures for the characteristic HRS tumor cells and their microenvironment. However, despite these studies, the potential prognostic role of miRNA signatures in cHL remains unclear and it has not been yet well investigated the relevance of microRNA expression in cHL and their putative value for patient outcome prediction.

In the study reported here, the aim was to investigate the relevance of microRNA expression in cHL within a prognostic context. Thus, miRNA expression data from the analyses of advanced cHL samples and cHL derived cell lines were used to identify specific miRNA profiles from tumoral cells and their non-tumoral microenvironment associated with lack of maintained response (in a similar approach followed in the first part of this thesis work) including also reactive lymph nodes and tonsils as controls to elucidate specific miRNAs for the malignancy. .Additionally, the influence of EBV infection on the expression pattern of miRNAs in the patient's samples was examined.

Finally, a panel of miRNAs was selected to be further studied in a large set of patients to derive a miRNA signature able to classify patients and bioinformatic analyses were used for identification of molecular pathways and target genes potentially altered by the expression of these miRNAs.

A general overview of the work is summarized in Figure 4.1

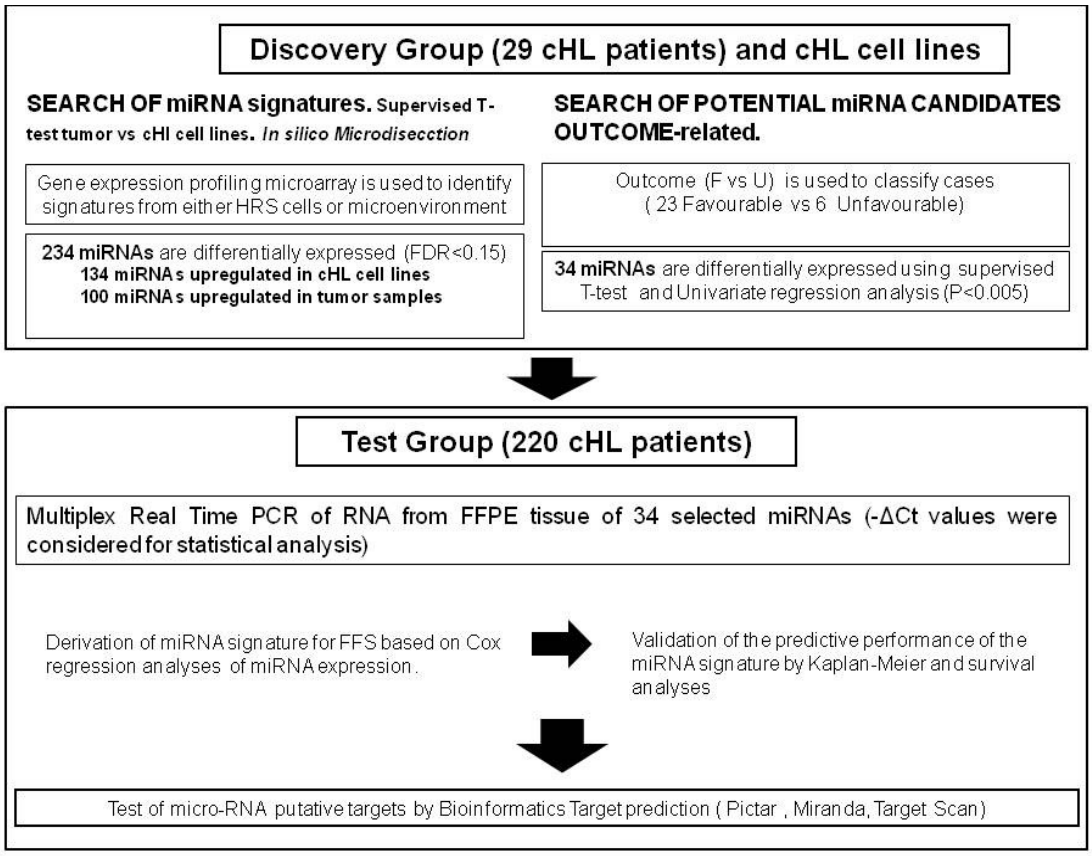


Figure 4.1: Experimental design. The whole series of patients was divided in two major groups, a discovery group (29 patients with available frozen tissue for profiling using array technologies) and a test group (220 patients with available FFPE tissue from the diagnostic pathological sample). In contrast to the experimental approach followed in the first part of this thesis work, all patient samples from test group (N=220) were used in the subsequent statistical analyses. This was used initially looking for miRNAs with the most prognostic ability (logistic regression analyses), to afterwards derive a final Cox model.

4.1 Identification of miRNA signatures

4.1.1 Correlation between EBV- encoded miRNA profiles and EBV status

To date, hundreds of miRNAs have been identified in humans and other species, including viruses. Furthermore, it has been reported that viruses have their own miRNA set and that there is an interaction between the host miRNAs and virus miRNAs (Chang et al., 2004; Kewalramani et al., 2004). EBV is present in the malignant cells of 40% to 60% of cHL patients. However, the precise role of the EBV in the pathogenesis of cHL is unknown. The interaction between the virus and the malignant cells in cHL might be mediated in part by miRNAs. Thus, the presence of EBV in tumors was assessed by conventional immunohistochemistry for LMP1 protein and in situ hybridization for EBERs (Figure 4.1) to further explore whether a specific signature of miRNAs was associated with EBV status. For the 30 cases analyzed, 6 out of the 30 cases were EBV+ whereas the 24 remaining ones didn´t shown neither LMP1 nor EBER1 and EBER2 protein expression.

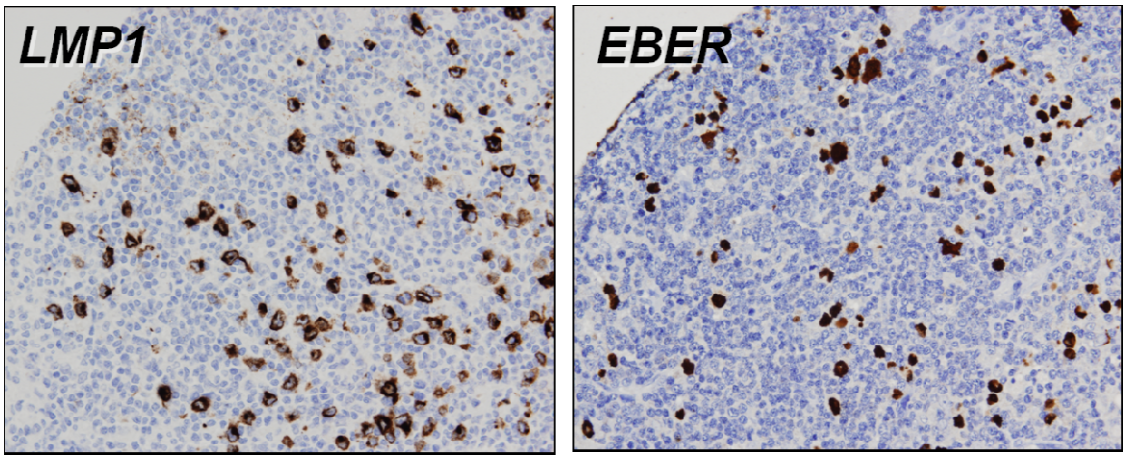


Figure 4.1 EBV immunohistochemistry pattern. Immunohistochemical staining of a EBV positive case showing positive expression for LMP1 and positive EBERs in situ hybridization.

First, non supervised analyses of EBV encoded miRNAs showed a random distribution of EBV+ and EBV- samples, with only some EBV+ tumors clustering together associated with overexpression of some EBV miRNAs. Interestingly, many EBV+ cases showed no overexpression of these miRNAs whereas some EBV- cases did show high levels of some EBV-encoded miRNAs.

Prior studies have shown miRNAs whose expression is influenced by the presence of EBV (Wong and Altekruse, 2009). In EBV+ cHL cases compared to EBV- cHL cases, miR-96 , miR-128a, miR-128b, miR-129 and miR-205 have been shown to be underexpressed whereas miR-28, miR-132, miR-140 and miR-330 are overexpressed. However, in this set of cHL cases, when supervised analysis (Pomelo T-test) was applied, the obtained results showed no statistically significant differences between EBV+ and EBV- cases in the expression of neither EBV-encoded miRNAs (Figure 4.2) nor human miRNAs. This could be due to the low number of EBV positive cases among the analyzed ones (just 4 out of the 30 cases analyzed).

However, although no miRNA signature associated with EBV infection was found,; when EVB status was not taken into account, high expression levels of some EBV encoded were observed in cHL tumor samples. Indeed, the signature identified in cHL (when compared tumor samples to lymph nodes and tonsil controls) included some EBV-encoded miRNAs such as ebv-miR-BART2-5p, ebv-miR-BART5, ebv-miR-BART12, ebv-BART13, ebv-BART16 and ebv-miR-BART-19-3p. These miRNAs are potentially associated with the viral latency status and may serve as an immunomodulatory mechanism in the tumors.

Thus, although EBV infection seems to be a common finding in cHL, this study was not able to identify any specific miRNA signature associated with EBV status. In contrast, some EBV negative cases were observed as able to overexpress specific EBV miRNAs thus pointing at the limitations that current inmunohistochemical methods based on viral proteins (LMP1, and EBERs) that might not be enough robust to accurate identify all EBV positive cases.

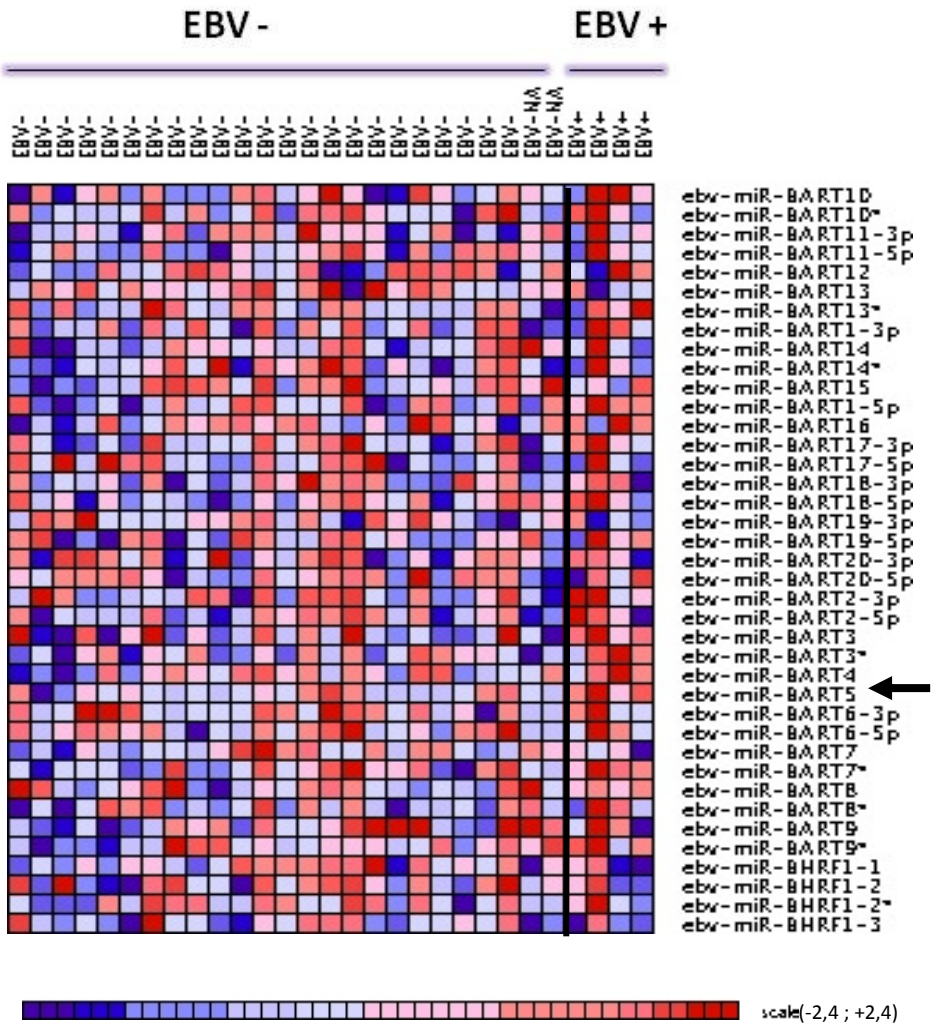


Figure 4.2 Supervised t-test (Pomelo tool) of samples EBV-positive and EBV-negative cases. As commented, no significant ebv-associated miRNAS were differentially expressed between EBV- and EBV+ cases. NA _ ebv- status not determined. The arrow is highlighting Ebv-miR-Bart-5p remarkable expression observed in ebv+ cases.

The presence of ebv-miR-BART5 in the cHL signature must be remarked and mentioned due to the previous demonstration of its antiapoptotic function of this viral miRNA that provided the first example for viral modulation of apoptosis through a miRNA. Plausibly, different viral miRNAs might target different cellular proapoptotic proteins to promote host cell survival. (Choy et al., 2008). Moreover, PUMA has been identified as a cellular target of this EBV encoded miR-BART5. This might reveal a new molecular mechanism through which EBV confers resistance to apoptosis in HRS cells by counteracting the proapoptotic function of p53. Interestingly and also in agreement with prior studies (Navarro et al., 2007) the small number of either cellular miRNAs or EBV- encoded mirRNAS among EBV+ cases and EBV – cases , could lead to hypothesize that EBV is not a primary transforming event in cHL.

4.1.2 Identification of tumor and microenvironment signatures

Initial studies are showing that miRNA losses and gains may explain some of the biological and clinical features of different lymphoma types, and recent analyses of miRNA expression in cHL cells and tumors have also demonstrated deregulated expression of some miRNAs in this malignancy.

To investigate whether specific miRNAs signatures are associated with cHL pathogenesis, we assessed the expression levels of 723 human and 76 viral miRNAs using a commercially available microarray platform (Agilent technologies). We identified the miRNA profile of cHL tumor samples after comparison with normal tissue controls, and characterize the signatures attributable to both the neoplastic cells and the microenvironment using miRNA expression profiles of derived cHL cell lines.

"In silico" microdissection approach considering cHL derived cell lines as a model system of HRS cells was used in the elucidation of miRNA signatures attributed either for tumoral cHL cells or microenvironment. Thus, supervised T-test (Pomelo II tool, http://pomelo2.bioinfo.cnio.es) of miRNA expression data from 29 frozen tumor samples (without overlap with the ones used in gene expression project (Results I) and 5 cHL derived cell lines (L428, L1236, L540, HDLM2, KHM2) allowed the identification of 234 miRNAs differentially expressed (FDR<0.15) between them.

Those miRNAs overexpressed in cHL cell lines (134 out of 234) represented HRS miRNA signature whereas those overexpressed in tumor samples (the remaining 100 out of 234) represented microenvironment signature and were evaluated afterwards for outcome association (F vs U).This initial analysis described a miRNA signature in cHL tumors, that contained miRNAs putatively expressed by the neoplastic HRS cells or the microenvironment. Some of the miRNA identified as part of the cHL signature have been previously described in this disease such as miR21, miR31 and miR-17-92 cluster members (miR-17, miR-19a) (Navarro et al., 2008; Wong and Altekruse, 2009). As expected, miR-155 known to be overexpressed in cHL cell lines showed a strong expression pattern in the cHL derived profiled in this study (Figure 4.3). Other members included in the signatures like miR-590-5p, miR-143, miR-145, miR-126 have been also described in different tumor types and could have a role in cHL.

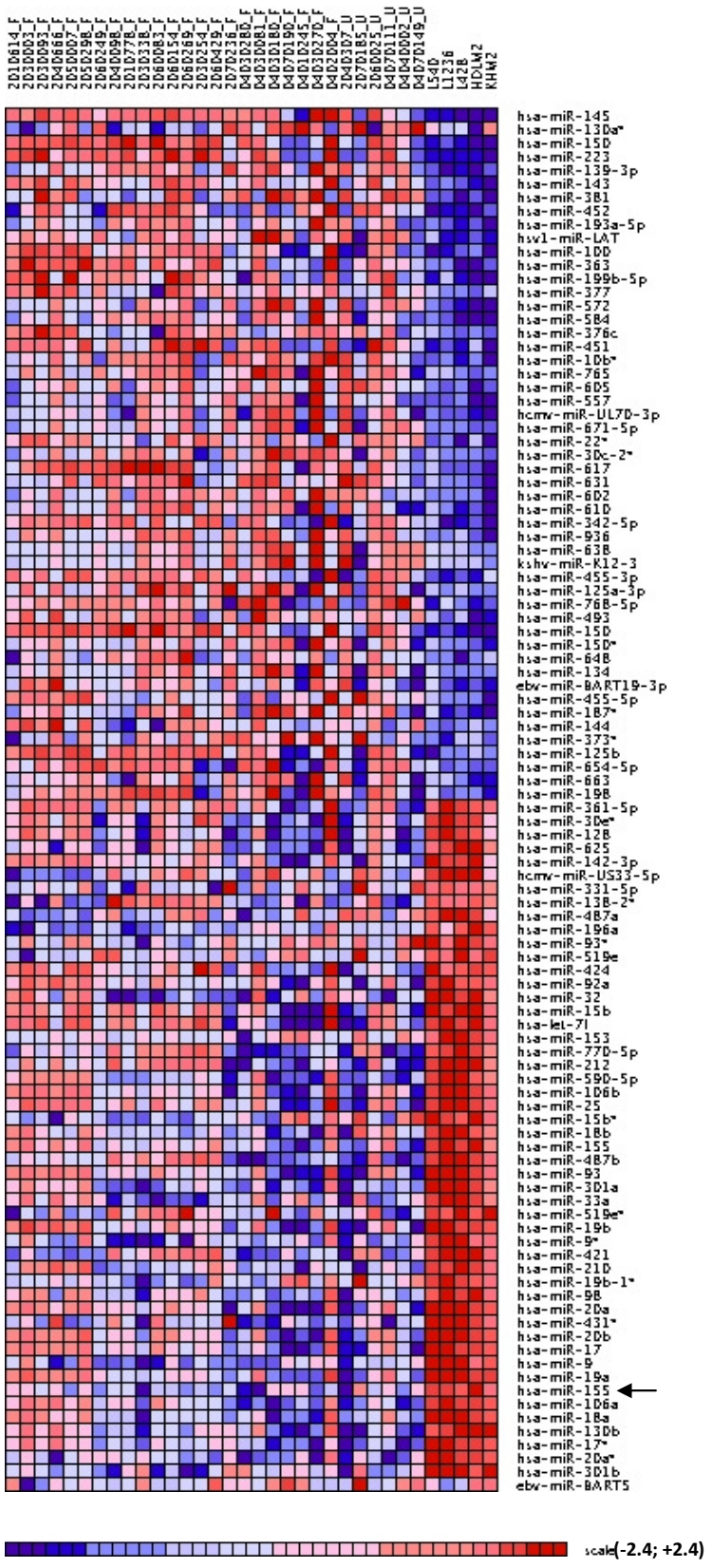


Figure 4.3 Supervised T-Test between tumoral samples (1) and cHL derived cell lines (2). Heatmap represents the 100 top-score differentially expressed genes FDR (<0.15) (either overexpressed (RED) or underexpressed (GREEN). As commented, miR-155 showed strong expression levels in cHL cell lines in comparison to tumoral samples (arrow).

When these identified miRNAs were compared with the previously published miRNA signature of cHL (Navarro et al., 2007), 7 out of the 25 miRNAs published by Navarro, showed higher expression levels in tumor samples compared to the controls. These consisted on miR-21, miR-23b, miR-26b, miR-30b, miR-31, miR-126 and miR-135a. However, other known miRNAs to be specific for cHL such as miR-16 and miR-24 (Wong and Altekruse, 2009) were not observed in the cHL´s signature identified by this work.

4.1.3 Clinical correlations. Identification of miRNAs outcome-associated in cHL

Clinical data of 29 patients were available, thus allowing the study of the relationship between miRNA signature and clinical outcome. Using univariate logistic regression analysis, response to ABVD therapy of cases was compared: favorable outcome (F), (sustained complete response) versus unfavorable outcome (U) (refractary disease).

When patient outcome was taken into account and based on univariate logistic regression analysis, a set of 34 miRNAs were differentially expressed ($p<0.05$) among outcomes and thus being considered for further evaluation by qRT-PCR using FFPE samples (Table 4.1). From this set, some miRNAS showed remarkable expression in the cHL cell lines in agreement with the previously elucidated tumoral and microenvironmental signatures, thus probably reflecting functional properties of the HRS cells whereas the others informed from properties of the reactive infiltrate without expression in cHL cell lines in agreement with the signatures previously identified (Figure 4.4). Interestingly, these initial selection included members of the miR-17-92 cluster such as miR-17, miR-92a, miR-92b and other miRNAs known to be involved in cancerogenesis (like miR-21, miR-21* and mirR-30d).

The observed differential expression pattern between HRS cell lines (representative of HRS cells) and tumor samples (representative of microenvironment) of miRNAS was clearly seen with exception of miR-92b*. As seen in Figure 4.4, expression of miR-92b* was observed for either from both HRS cells and the microenvironment. Thus, it should be remarked that miRNA adscription to microenvironmental or HRS cells signatures was done taking into account statistical T-test results and median general expression of miRNAS between the two compared groups.

Table 4.1. MirRNA prognostic signatures identified by "in silico" microdisecction. (N=29)

Composition of the HRS cells miRNA signature			*Composition of the microenvironment miRNA signature*		
miRNA	**Univariate Regression Analyses**	**cHL cell lines expression**	**miRNA**	**Univariate Regression Analyses**	**cHL cell lines expression**
hsa-miR-17*	0.0055339	low	hsa-miR-92b*	0.004834	low
hsa-miR-25	0.0243501	low	hsa-miR-26a	0.0260448	low
hsa-miR-30d	0.0289642	low	hsa-miR-27b*	0.0446511	low
hsa-miR-320	0.0359278	low	hsa-miR-148b*	0.0006499	no
hsa-miR-15b*	0.0359478	low	hsa-miR-204	0.0011748	no
hsa-miR-30e	0.0367626	low	hsa-miR-539	0.0037592	no
hsa-miR-23a	0.0476105	low	hsa-miR-609	0.0045141	no
hsa-miR-216a	0.0486403	low	hsa-miR-132	0.0060738	no
hsa-miR-93	0.0489202	low	hsa-miR-31*	0.0118676	no
hsa-miR-378	0.0087782	yes	hsa-miR-628-3p	0.0157169	no
hsa-miR-21*	0.0131724	yes	hsa-miR-503	0.0162567	no
hsa-miR-708*	0.0174565	yes	hsa-miR-621	0.0197161	no
hsa-miR-551a	0.0253349	yes	hsa-miR-637	0.0272346	no
hsa-miR-92a	0.0302539	yes	hsa-miR-300	0.029759	no
hsa-miR-130b	0.0330384	yes	hsa-miR-423-5p	0.0324985	no
hsa-miR-21	0.0371476	yes	hsa-miR-937	0.0346731	no
			hsa-miR-559	0.0371576	no
			hsa-miR-888	0.0489402	no

Note. Univariate regression analyses between patient outcomes (p<0.05 as cut-off value)

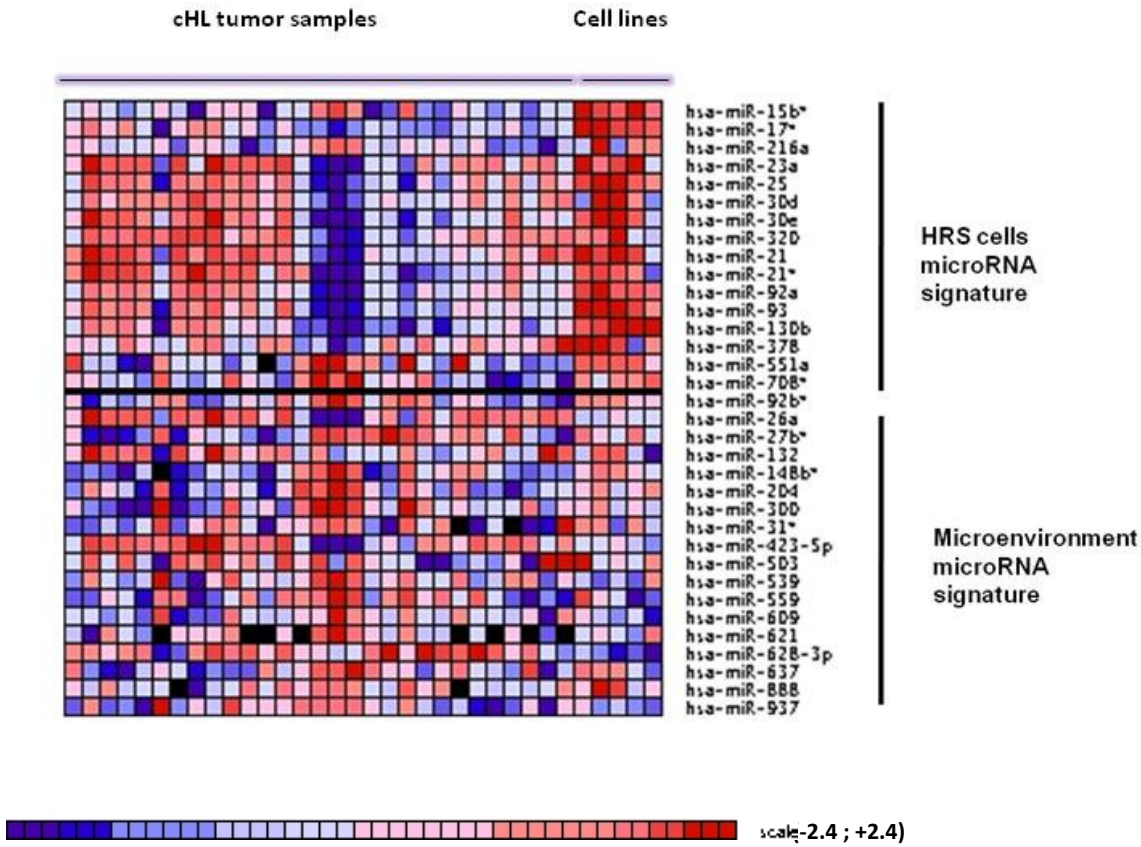


Figure 4.4 Heatmap miRNAs prognostic signatures identified by "in silico" microdisecction and Univariate Regression Analyses using outcome (Favorable versus Unfavorable) as dependent variable.

4.2 Laser capture microdissection of HRS cells and microenvironment

To further investigate whether miRNA composition of both signatures identified by "in silico" microdissection was a consistent approach, two of the most expressed miRNAS in cHL cell lines (miR-21 and miR-30d) were additionally validated by laser capture microdissection (LCM). First, CD30 marker was analyzed as a positive control for HRS cells, thus selecting 800-1000 HRS cells to be microdiseccted including also fragments of their surrounding microenvironment in separate tubes and measuring miRNA expression by RT-qPCR. Comparison of both normalized miRNA expression levels (ΔCt) from HRS cells and the reactive infiltrate showed higher expression of miR-21 and miR-30d in tumoral HRS cells, thus supporting the previously identified miRNA signatures (Figure 4.5). Higher expression in HRS cells was more remarkably for miR-30 whereas subtle differences were seen for miR-21.In agreement with this results, preferential expression of miR-21 in the cytoplasm of HRS cells has been already demonstrated by other studies by chromogenic in situ hybridization(Navarro et al., 2007).

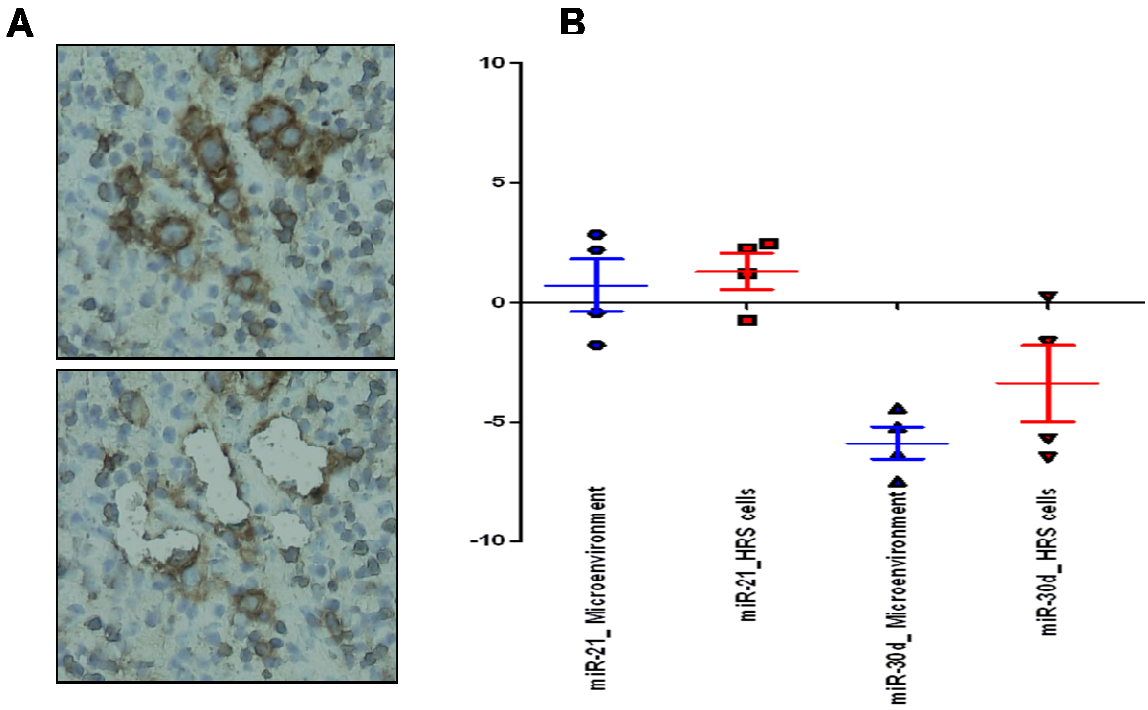


Figure 4.5 Laser capture microdissection of HRS cells (≈ 800-1000) and quantitative RT-PCR for miR-21 and miR-30d, normalized levels using RNU44 and RNU48. A) CD3O positive staining for HRS cells identification before and after microdissection process. B) Normalized (ΔCT) values (RNU44 and RNU48 used as endogenous miRNAs, from both HRS cells and microenvironment fragments.

4.3 MicroRNA signatures related to treatment response and patient outcome

The first endpoint of this study used for initial miRNA selection was the response to standard first-line treatment considering favorable response (F) and unfavorable response (U), as mentioned above without considering data from second-line and salvage therapies and/or bone-marrow transplantation. As previously commented and shown in Table 4.1 a set of 34 miRNAs were differentially expressed among patients with favorable and unfavorable outcomes, thus being selected and further studied in an independent set of 220 FFPE samples recruited from different institutions of the Spanish Hodgkin lymphoma group and MD Anderson Cancer Centre (Houston, TX). Differences in the distributions of standard clinical parameters (age, gender, stage, IPS, the individual variables contained in IPS and outcome –favorable versus unfavorable-) in the initial discovery dataset (N=29) and the test sample series (N=220) were not statistically significant (Pearson chi-square test). A summary of the clinical characteristics of the whole sets of FFPE patients used in the study can be found in Material and Methods Section.

4.3.1 Logistic regression analyses

Initially, miRNA final selection was done by backward stepwise logistic regression analysis process and revealed a 4-miRNA signature (miR-21, miR-30d, miR-30e and miR-92b*) related to outcome (Table 4.2) that was able to predict treatment response with an overall accuracy of 69.5 percent.

Table 4.2. Logistic regression model Test dataset (N= 220)

	B	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
mir21	.193	.030	1.212	1.019	1.442
mir92b*	.280	.003	1.323	1.099	1.594
miR30d	-.209	.018	.811	.683	.964
mir30e	-.279	.000	.756	.665	.860
Constant	-1.722	.000	.179		

Table 4.2 Logistic regression analysis and classification table of the miRNA model. 220 FFPE tumor samples were analyzed and adequate miRNA RT-PCR profiles were obtained in all cases. Normalized expression levels (ΔCt), using RNU44 and RNU48 as endogenous of the individual miRNAs varied considerably among samples. However, a multivariate logistic regression analysis identified a 4-miRNA signature able to predict treatment response with an overall accuracy of 69, 5 percent.

In addition, a miRNA expression-based model using the expression of these 4 microRNAs as a continuous variable was able to predict two risk groups with differences in survival probability with FFS at 5 years of 70,7% versus 42,3% (p=0.03). Its predictive performance in the test group of patients was confirmed using the integrated area under the ROC curve (ROC value = 0.694).

4.3.2 Survival analyses. The miRNA score

In a similar approach as followed in the first part of this thesis work, a Cox proportional hazards model analysis was performed to determine the association of the miRNAs initially found to have potential association with outcome in the logistic regression analyses. Finally, Cox proportional hazard and survival models based on the miRNA expression were derived for the entire series (N=220).

Using a backward stepwise selection algorithm, a Cox model for the entire series (N=220) was constructed containing miRNAs found to be associated with FFS. This analysis

highlighted the same 4-mIRNA signature previously found by regression analysis with miR-21, miR-92b*, miR-30d and miR-30e significantly associated to FFS (Table 4.8).

In this multivariate Cox analysis, miR-21 and miR-30d were shown to be positively associated with poorer outcomes and FFS (Exp (B) = 1.239 and 1.275) whereas overexpression of miR-92b* and miR-30e seemed to have a protective effect. The 4-miRNA signature ranged from 0.291 to 0.895 and was associated to FFS. Its classification accuracy performance was confirmed using the integrated area under the ROC curve (ROC value = 0.727) and after leave-one-out cross-validation, a 70 % correct classification rate was observed. Additionally, this signature was able to stratify patients according to the median expression of the miRNA signature into high and low risk groups respectively (cut off point =0.669). Thus, Kaplan-Meier estimates showed significant differences in FFS probability between the low risk and the high risk groups with 5-year FFS of 77.8 % versus 40.2% (Figure 4.7).

Table 4.8 Backward stepwise Cox regression analyses in the whole entire series N=220

	B	Sig.	Exp(B)	IC 95.O% para EXP(B)	
mir21	0.214	0.040	1.239	1.009	1.520
mir92b*	-0.294	0.000	0.745	0.640	0.868
miR30d	0.243	0.018	1.275	1.043	1.559
mir30e	-0.225	0.023	0.799	0.658	0.970

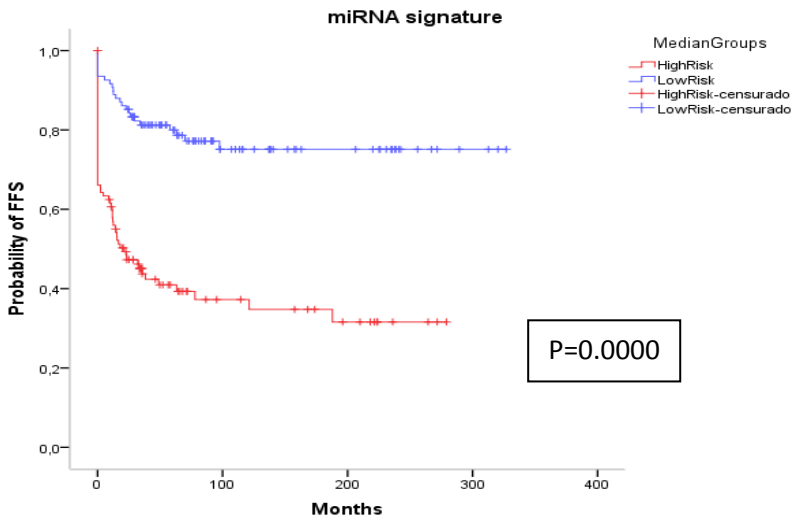


Figure 4.7 Kaplan Meier analysis failure free survival model miRNA signature. The final Cox model integrated the 4 miRNAs was able to identify two risk groups according to the median miRNA signature expression: FFS- 5 years survival rates 77.8 % versus 40.2 % into low and high risk groups respectively.

4.4 MicroRNA target analysis. The miRNA regulatory network

4.4.1 Bioinformatics target prediction. Diana-MirPath

In order to gain a deeper insight into the network of genes that might be regulated by the set of outcome- related miRNAs we interrogated the current miRNA target prediction databases available, MiRBase, TargetScan and Tarbase (this last one containing target genes experimentally validated). A high amount of targets for each miRNA were identified: as an example 198 genes for miR-21, 994 genes either for miR-30d or miR-30e and 640 genes for miR-92b* using TargetScanv5. Thus, to summarize results or extract the most relevant ones we focused on those target genes commonly found in the different databases.

Interestingly, the predicted miRNA- gene pairs found for both miR-21 , miR-30d and miR-30e included BCL2, BCL2L1, BCL11, BAX, PDC4, RAS, TRAIL3 and p53 (PTEN) among others, thus suggesting the potential crucial role of the miRNA regulatory network in the deregulated apoptosis mechanism observed in cHL malignancy. Again, genes implicated in apoptosis regulation (BCL2, BCL2L1, BCL11 and BAX) were highlighted, in agreement with the already identified signatures as outcome related by microarray gene expression and RT-PCR assays (Results I).

Table 4.9 Target genes mirRNAs include in the signature

miRNA ID	**Chr. Pos.**	**Putative Targets**
hsa-miR-21	17q23.1	PTEN*,PDCD4*,TMP1*,SERPINB5*, BCL2, FASLG, IL12A, CDK6, CDCD25A, CCL1, CCL20, CCL22, CCR7, PIK3R1, PPP3CA, ACV1R1C,ACVR2A,BMPR2,PITX2,SAMAD7,TGFBR2, BTK, NFAT5, MAP2K3, NTF3, MAP3K1, FASLG, FGF1, DUSP8, PPP3CA,TGFBR2, RPS6KA3,CNTFR, LIFR,SPRY1, SPRY2, LRP6,
hsa-miR-30d	8q24.22	BCL2, CASP3, CCNE2, CDC7, IL1A, IL2RA, IL28RA, MAP3K12, MAP3K5,MAP4K4, ACVR1,SMAD1,SMAD2,TFDP1, THBS2, KRAS, LYN, VAV3, VAV2, PPP3CA, PPP3CB, PIK3R2, NFAT5, PIK3CD,RPS6KA2, CACNB2, BDNF, CRKL, TAOK, SOS1, RAP1B, NLK, NF1, RASA1, PDGFRB, NCK2, CBLB, PLCG1, PRKAR1A, SOCS1, LIFR, LEPR, IFNAR, SOCS3, CBLB, JAK1, CLCF1, DBF4, E2F3, YWHAZ, ORC2L, TFDP1, ABL1, NFYB, PSME3,HSPA5, ITGA6,ITGA5,ITGB3,CAMK2D,LRP6,TBL1X,VANGL1,PRICKLE1,NLK,IFNAR2,SERPINE1,IGF1
hsa-miR-30e	1p34.2	BCL2, CASP3, IL1A, IL2RA, IL28RA, CCNE2, CDC7,ACVR1, SMAD1, SMAD2, TFDP1, THBS2, KRAS, LYN, VAV3, VAV2, PPP3CA, PIK3R2, NFAT5, PIK3CD, PPP3CB, MAP4K4, RPS6KA2, CACNB2, BDNF, CRKL, TAOK1, SOS1, RAP1B,NLK, NF1, RASA1, MAP3K12, MAP3K5, PDGFRB, NCK2, PIK3R2, CBLB, PLCG1, PRKAR1A, SOCS1, LIFR, LEPR, IFNAR2, SOCS3, CBLB, JAK1, CLCF1, DBF4, E2F3, YWHAZ, ORC2L, TFDP1, ABL1, NFYB, PSME3, HSPA5, ITGA6, ITGA5, ITGB3, CAMK2D, LRP6, TBL1X, VANGL1, PRICKLE1, SERPINE1, IGF1
hsa-miR-92b*	1q22	CDK6, CDKN1C , CDC27, CDC42,MAP2K4, MAP3K13, BMPR2, PPP2R2A, RBL2, SMAD6, SMAD7, SMURF1, SP1, NFAT5, PIK3R3, , DUSP5, RPS6KA4, RAP1B, NLK, DUSP10, NF1, SPHK2, , PRKAR2B, E2F3, , NFYC, NFYB, ITGA6, ITGA5, CXCL5, CTNNBIP1, DOCK5,FZD10,TRAF3,SESN3

The putative target genes shown were selected based on functional aspects (Diana-mirPath) and statistical significance ($p<0.005$). Chr. Pos. _ Chromosomal position.

Putative targets identified using mirBase and Target Scan

*** Target genes experimentally validated were identified from TarBase (experimentally validated).**

A summary of all putative target genes found by the different bioinformatic target prediction softwares used (TargetScan, mirBase and TarBase) is shown in Table 4.9.

The different target prediction algorithms used to look for potential miRNA targets , followed similar rules such us the near- perfect sequence match between the first 8 nucleotides of the 5´end of the miRNA (the so called "seed-region"), conservation of the binding site between related species and thermodynamic parameters. The consequences of varying so many parameters are that many times there is little overlap between different predictions making it difficult to distinguish true targets from false positives. Nonetheless, it´s widely accepted that targets predicted by more than one algorithm have more chances to be true targets, thus being the experimental procedure followed in this work (choosing those common targets among the different algorithms used) a way to minimize the putative false positive results.

Additionally, to further understand the biological significance of the different identified targets, we included Diana-mirPath software (Papadopoulos et al., 2009), a web-based computational tool developed to identify molecular pathways potentially altered by the expression of single or multiple microRNAs (http://diana.cslab.ece.ntua.gr/pathways). This approach looked for pathways associated to these target genes, paying special attention to those ones related to lymphoma pathogenesis. These combined analysis of miRNAs, with bioinformatic target prediction from Diana- mirPath revealed a series of genes and pathways potentially targeted by the 4 miRNAs of the signature (Table 4.10).

Table 4.10 Diana-mirPath. Pathway analyses miRNAs included in the signature

		miRNA ID			
KEGG Pathway	**(all miRNAS)**	**hsa -mir21**	**hsa-mir30d**	**hsa-mir-30e**	**hsa-mir-92b**
	- ln(pvalue)	**- ln(pvalue)**	**- ln(pvalue)**	**- ln(pvalue)**	**- ln(pvalue)**
TGF-beta signaling pathway	5,39	9,50	0,17	0,17	3,37
B cell receptor signaling pathway	5,03	5,98	5,52	5,52	0,32
MAPK signaling pathway	4,52	7,73	1,69	1,69	0,24
T cell receptor signaling pathway	3,98	1,44	5,86	5,86	0,12
Apoptosis	**2,09**	**3,82**	**1,77**	**1,77**	**0,11**
Jak-STAT signaling pathway	1,83	4,39	1,94	1,94	2,22
Cell cycle	1,29	0,01	1,0	1,0	0,40
P53 signaling pathway	0,09	0,42	0,15	0,15	0,14

Note _ -ln (p-value): negative natural logarithm of the enrichment p-value calculated for the specific pathway. The program considers as biologically relevant all those pathways with -ln (p value) >0

A total number of 156 Pathways were highlighted including essential pathways for lymphoma survival like TGF-signaling pathways, MAPK signaling pathway, Cell cycle, Apoptosis, Jak-STAT signaling pathway, T-cell and B cell receptor signaling pathways. This Diana-mirPath bioinformatic analysis also highlighted genes from p53 signaling pathways as putatively targeted by the 4 miRNA signature mainly associated to mir30d in agreement with other works ((Kumar et al.).

4.4.2 Correlation analyses.

Finally, correlation analysis was performed using gene expression and miRNA expression data from samples that overlapped among both projects, gene expresion project (RT-PCR from TLDAS) and miRNA project (RT-PCR data for miRNA Taqman Assays) including the 11 genes for the MRS model and the 4 miRNAs in the signature.

A set of 96 patient samples had both data (GEP from the 11 genes included in the model and miRNA expression obtained by RT-PCR), corresponding all samples to MD Anderson Houston cancer centre series. Interestingly, significant correlation pairs were identified, including miR21-BCL2L1 (Spearman's rho correlation coefficient=0.222; p=0.033) and miR21-CASP3 (Spearman's rho correlation coefficient =-0.207; p=0.043). The pair miR21-BCL2 showed a negative correlation coefficient (Spearman's rho correlation coefficient=-0.136) although this comparison was not significant.

When apoptosis pathway (BCL2, BCL2L1 and CASP3) was correlated with mir21 expression a significant correlation was found (Spearman's rho correlation coefficient=0.158); p=0.041). These results were again suggesting a crucial role of the identified miRNAS in the signature (especially for miR21) into the observed deregulated apoptosis in cHL. Additionally, significant correlations were found between both mir21 and mir30d and genes from Cell cycle pathway (CCNA2, CCNE2, CENPF and CDC2). However, no other significant correlations were found between the other miRNAs included in the signature and the genes included in the MRS model neither for mir30e nor mir92b*.

DISCUSSION

5.1 RT-PCR based models. Technological development

5.1.1 RNA from FFPE samples and qPCR. TLDAs

As demonstrated by all the studies done until now, transcriptional analysis of cancer is proving to be a powerful and increasily useful tool in biomedical research. A goal of these studies is the use of gene expression patterns revealed by transcriptional profiling to understand the pathogenesis of the disease and then predict prognosis and responsiveness to therapy. However, the practical applications of these advances in the routine clinical care of patients are rare, mainly due to the limitations arising from the relatively few cases that can be studied when using frozen specimens.

In this study it has been shown that robust methodologies, based on quantitative RT-PCR, are suitable for expression profiling of tumors and can be applied to routine FFPE tissue samples and that new platforms such as the TLDAs allow an accurate analysis of a limited number of genes. Hundreds of reactions can be done simultaneously, enabling large numbers of cases to be rapidly assessed. These technologies would be especially useful for identifying gene signatures of prognostic significance as it has been already done in other tumor types (Glenn et al., 2007).

From a methodological point of view, quantitative RT-PCR is a robust method to quantify RNA abundance being the procedure highly sensitive and reproducible as long as the initial RNA is intact. However, breaks in the RNA due to chemical or enzymatic cleavage reduce the number of molecules that contain intact amplicons and the number of molecules available for amplification decreases. When the described TLDA platforms were initially tested to evaluate the feasibility of this tool and the suitability of FFPE samples some limitations were detected. Those limitations were overcome with the optimization of extraction protocols and the correct design of the TLDAS.

As a summary , in Table 5.1 are listed the problems found during initial setting up of the technique, their main cause of origin , the solutions and additional factors that can be taken into account to improve RNA extraction. These included the efficiency of RNA recovery from archival samples, fragmentation of the RNA during fixation, quality of PCR primers and the size of the amplified amplicon. Some of these factors were quite difficult to control. While many problems related to the isolation of RNA from FFPE can be solved with high concentration of proteinase K and long incubation times (overnight) at elevated temperature

followed by phenol-chloroform extraction, the fragmentation of the RNA remained the most critical factor inherent to the sample and there was not solution. However, in our hands 93, 90 % of FFPE analyzed samples yielded significant results using phenol /chloroform extraction protocol as described in Material and Methods and in agreement with previous works ((Koch et al., 2006).

Table 5.1. Limitations concerning RNA from FFPE tissues.

Problem	**Origin**	**Solution**	**Additional factors**
1.Low quality RNA	It depends on the sample	There is not solution 93.90% of the samples yielded analyzable data based on our results	Sample processing Storage time (less than 10 years)
2.Low amount RNA	It depends on the sample	Number and thickness of paraffin sections / (3 sections of 30um) Extraction protocol Preamplification	Phenol / chloroform extraction Commercial kits (column based)
3.Genes selected	Study´s design	Amplicon length(<100bp)	Choose correct assays Load enough RNA

Additional recommendations included also the preferential use of samples no older than 10 years in storage time and to avoid column based extraction protocols (von Ahlfen et al., 2007). Interestingly, more recent works suggest that it is also possible to extract mRNA of sufficient standard good quality for further transcriptomic analysis from minute FFPE samples up to 15 years old (Stewart et al., 2011). Thus, exploitation of retrospective tissue collections with robust associated clinical data is possible, although in our hands cases very old didn't work as well as the others, thus suggesting the convenience of quite new samples as published.

We analyzed in more detail the consequences of RNA fragmentation and amplicon size on RT-PCR and Ct values. From a theoretical point of view, qRT-PCR is affected by RNA fragmentation, as every break in the RNA that occurs between two primers inevitably separates the two ends of the amplicon into two different cDNA molecules and is therefore lost as template for subsequent PCR reactions. In agreement with this hypothesis, small amplicons resulted in more stable delta Ct values (< 32 cycles) than relatively large amplicons (median Ct values 32-36 cycles). This phenomenon was previously shown for large amplicons

which failed to generate measurable expression values (Godfrey et al., 2000; Lehmann and Kreipe, 2001; Specht et al., 2001). As a conclusion, short amplicons should be used for quantitation. Therefore, to ensure high quality and fidelity of the TLDA assays amplicon length was also considered as selection criterion for genes included in the study, discarding genes with longer amplicon length and choosing those ones with length that were preferentially lower than 100 base pairs (bp).

Additionally, we investigated the optimal system of normalization to ensure reliability of the qPCR data and 4 putative control genes (HMBS, GUSB, TBP and GADPH) were included in the Low Density Arrays (TLDAS) , afterwards selecting two of them (HMBS and GUSB) as controls according to GeNorm procedure developed by Vandesompele et al.(Vandesompele et al., 2002). As published, degradation related shifts of Ct values can be compensated by calculating DeltaCt values between test genes and the mean values of several control genes. Thus, these two endogenous selected genes were used to compensate for the changes in expression levels resulting from the increased fragmentation of RNA. GeNorm normalization´s procedure based on several genes has been already demonstrated by previous studies (Antonov et al., 2005) and allowed us to obtained reliable and robust data.

This study shows how the capabilities of RT-qPCR have been expanded with the development of the TLDA platform. Low amount of RNA is required, small amplicons can be included and the results obtained from RNA isolated from FFPE tissues correlates well with those ones obtained from fresh tissue (Steg et al., 2006). The application of this RT-PCR based platform on two subsequent series of paraffin embedded samples (52 and 282 samples) demonstrated the feasibility of applying the platform to paraffin embedded specimens, although only short amplicons gave rise to estimable results in the initial TLDA 64 format platform as previously commented. However, when the genes included in the TLDA platform were subsequently selected taking into account their biological interest, short amplicon lengths and several endogenous control genes were used for data normalization, good quality data were obtained.

The approach proposed in this work has been done in breast cancer and other tumor types. Based on the preliminary results, we moved a step forward and examined whether RNA preamplification could be used in combination with TLDA in order to recover those samples that didn´t amplify. Preamplification is a method whereby nanogram amounts of total cDNA undergo a multistep process for linear amplification of the initial RNA fraction once it has been retrotranscribed to cDNA. Reproducibility and reliability of Preamplification have been proved

and applied also in other works (Ciotti et al., 2009). In this thesis work was an optimal way to ensure precision of Taqman assays, yielding Ct values inferior to 30 cycles, something that is especially Important to note taking into account that precision of Taqman arrays started dropping of at Ct 30-32 as the number of copies approached to less than 10 copies /µl. (As noted by the manufacter´s recommendations, Applied Biosystems).

As a final conclusion, TLDA methodology presented in this study represents a robust and reproducible technique for quantifying gene expression in tens to hundreds of independent genes concurrently in RNA samples isolated from FFPE tissues. It constitutes a significant advance in multivariate gene analysis that is less time and labor intensive than individually analyzing single genes.

5.1.2 MiRNA from FFPE samples and qPCR

Despite the successful results obtained by TLDA (qRT-PCR) from total RNA extracted from FFPE whole tissue sections, we additionally evaluated microRNA as an alternative analyte for gene expression studies of FFPE samples. Since recovery of quality intact mRNA compatible with molecular techniques is often difficult due to the high degradation as previously commented, we analyzed the performance of the total miRNA and RNA yield, miRNA recovery and robustness of real-time PCR in cHL FFPE samples. Hypothetically the small size of these miRNAs (21 nucleotides medium length) would make them ideal for this kind of molecular analyses. In agreement, miRNAs generated quantitative reverse transcription-polymerase reaction signals that were more robust and reliable (in terms of lower Ct values, median Ct values 25-30 cycles) in comparison with longer mRNAs independently of FFPE block´s age and RNA quality. Based in these results, we concluded that miRNA recovery and expression analysis when using FFPE samples are more robust and accurate than total mRNAs. This observation has been also previously commented by other studies (Doleshal et al., 2008) highlighting the suitability and emerging interest of miRNA analyses as a potential diagnostic and prognostic tool (Chen et al., 2008).

In general, mRNA – based assays using total RNA seem to be less reliable, primarily due to the suboptimal RNA qualities in these total RNA specimens (92.90 % of adequate RT-PCR profiles in total RNA vs. 100 % accurate profiles in miRNA). In contrast to the high degradation observed in total RNA samples, miRNAs were uniformly well preserved in the formalin paraffin embedded tissues, presumably due to their small size. This advantage indicates that miRNA based assay is most well suited for the molecular analysis of routinely processed formalin-fixed

tissues. Despite differences in the efficiency of the amplifications, this work has proven that archival tissue samples can be used for expression profiling either for both genes and miRNAs. RNA derived from FFPE samples will be always substantially degraded and this inevitably will affect accuracy and sensitivity of expression measurements based on real-time PCR. Nevertheless, reliable measurement can still be made for many genes and miRNAs. A combination of both miRNA and genes would be and ideal approach to help in the identification of prognostic markers and biological processes underlying the different malignancies.

5.2 Gene expression profiling and classical Hodgkin lymphoma

5.2.1 Microarray-based gene expression and microdisecction

Gene expression signatures have been identified for the most common types of non-Hodgkin lymphoma. These studies have proved the ability and feasibility of these technologies to identify pathogenic mechanisms, new molecular targets and biological processes involved in lymphomagenesis. In addition, the relationship between clinical outcomes and expression profiles of different lymphoma types has been proved(Alizadeh et al., 2001). Regarding Hodgkin lymphoma, gene expression analysis could provide vital clues and new insights to the pathogenesis of the disease and potentially identify specific patterns related with tumor aggressivity and/or sensitivity to therapy. However, as mentioned in the introduction of this thesis work, the clinical and histological heterogeneity of classical Hodgkin lymphoma and the characteristic scarcity of HRS cells have prevented molecular studies.

Fortunately, during the past ten years, a number of approaches including microdissection have tried to resolve the variability in measurements derived from tumor heterogeneity. Another approach used in this work and designated as virtual or *in silico* microdissection (Alizadeh et al., 2001), avoids the laborious and time-consuming step of anatomic microdisecction and consists of confronting the gene expression profiles of whole tissue samples to those of cell lines representative of different cell lineages, differentiation stages, or different signaling pathways. This strategy has been used in recent studies (Navarro et al), aiming the analysis of whole tissue sections as a good alternative.

Despite the lack of a perfect model of HRS cells, we hypothesized that gene expression analysis on cHL derived cell lines would help to know more about the biology and putative pathogenic mechanisms underlying cHL as previously done by other groups (Kuppers et al.,

2003; Schwering et al., 2003a) allowing also the identification of genes that could be putatively expressed by HRS cells.

The first part of this thesis work wanted to perform gene expression analysis of whole cHL tissues focusing on advanced cHL cases with a double aim: first, to identify factors related either with the neoplastic HRS cells and the microenvironment (By supervised analysis comparing gene expression patterns from HL-derived cell lines) ; and second, to identify gene expression signatures that could be implicated in therapy resistance (By supervised analysis comparing gene expression patterns from patients with good and bad outcome) .Due to the complex composition of the elucidated gene signatures and thinking of a potentially clinical application, the attention was finally focused on the development of a model able to classify patients according to therapy response that could be applied to a routine diagnostic setting. In a multistep approach we first reduced the number of potential predictive genes by using the GSEA bioinformatic approach and second we move forward from Microarray techniques (frozen samples required, economically expensive) to RT-qPCR methods (suitable for FFPE samples easily available, economically affordable) in order to develop a predictive model of potential clinical application. Analyzing a larger and independent patient series, finally a 4-cluster/ 11-gene model was derived from an initial selection of 30 potentially predictive markers.

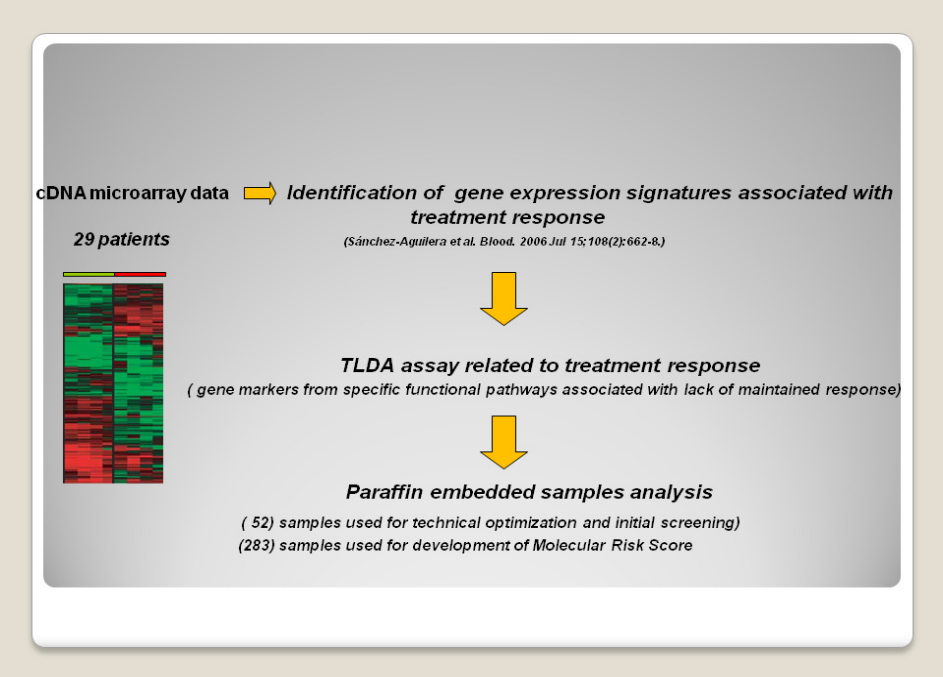


Figure 5.1 General overview of the different and subsequent steps followed for MRS development. TLDA- Taqman low density arrays.

5.2.2 GSEA

To develop a predictive model clinically useful and simplify the identified signatures, the reduction in the number of genes was needed, always preserving the biological significance. Thus, Gene set Enrichment Analysis (GSEA) was a very useful tool in this issue and allowed the identification of pathways and genes implicated in treatment failure.

Using GSEA, we found a number of gene sets relevant to biological treatment resistance processes with high complexity; these comprised members of different signaling pathways (such us JAK-Stat and apoptosis) **,**the interleukin- and chemokine-family, Cell-cycle regulatory genes, including cyclin-dependent kinases, and genes mainly related with G2/M transition, G1/S phase and the regulation of spindle checkpoint. Additionally, genes that reflected specific cell populations of the immune response were identified, such as specific T-cells and monocyte/macrophage cells in agreement with previous results. This suggests that, in cHL, there may be a limited number of gene sets responsible for the treatment failure observed in some advanced stages. Moreover, our identified signatures highlight the same biological processes found to be associated with outcome by other studies.

The comparison performed at a gene set level enabled the identification of common biological processes associated to a specific biological question (In this particular case treatment failure, favorable versus unfavorable outcomes), and demonstrated that GSEA embedded in the statistical workflow procedure is a suitable approach to study gene expression data, improving the analysis avoiding the loss of biological information derived from single gene studies. A systematic approach such as GSEA is needed to analyze raw data from microarray analyses. However, it must be said that when the mass of raw data is derived from a relatively small number of tissue samples (like in this case, 29 samples) results are descriptive and validation studies are afterwards needed.

5.2.3 The Molecular Risk Score model

Important functional pathways were highlighted by GSEA analysis to be related to treatment response but the final model presented in this work included just the expression of 11 genes grouped into 4 pathways.

In the MRS model, expression of BCL2, BCL2L1, CASP3, HMMR, CENPF, CCNA2, CCNE2 and CDC2 included in Apoptosis and Cell Cycle pathways respectively correlated with short FFS, in agreement with the repeated observed abnormal expression of antiapoptotic BCL2 family

proteins in HRS cells (Bai et al., 2007; Garcia et al., 2003), thereby contributing to their survival. Both, BCL2 and BCL2L1 are frequently expressed by HRS cells and their levels have been associated with short FFS patients treated with ABD or equivalent regimens (Rassidakis et al., 2002a). This confirms the importance of Apoptotic regulators for cHL outcome prediction and the complex regulatory network of this process.

The second signature, Cell Cycle (HMMR, CENPF, CCNA2, and CCNE2) was mainly composed by genes encoding for regulatory proteins of the S and G2/M phases of the cell cycle (as identified by GSEA analyses), pointing out directly to the implication of cell proliferation in treatment resistance, which is again in agreement with already identified markers. Interestingly, a significant prognostic value for CDC2 (CDK1) and CCNA2 (Cyclin A2) at protein level has been described for both cHL (Bai et al., 2004) and non- Hodgkin lymphomas (Wolowiec et al., 1999) (related to both FFS and OS). These results provide more evidences that the cell cycle and apoptosis regulation are profoundly disturbed and closely related in cHL disease.

In addition, *IRF4* (*MUM1*) expression was associated with longer FFS. This gene is an interferon regulatory factor, lymphocyte specific , induced after nuclear factor- B (NF-κB) activation , which controls B-cell proliferation and differentiation and has recently been shown to be up-modulated by CD40 engagement in HL cells (Aldinucci et al.). The lack of IRF4 protein has been previously associated with outcome in cHL, representing a potential adverse prognostic factor.

Interestingly, both *IRF4* and *BCL2L1* represent well-known NF-κB target genes whose expression is induced after NF-κB pathway activation. Thus, the MRS included important subrogates from the NF-κB activation, thought to be an essential pathogenetic mechanism in this disease. In fact, it has been recently described how the inhibition of the canonical NF-κB pathway enhances the proapoptotic effects of Adriamycin, thus also identifying NF-κB inhibition as an interesting therapeutic approach. (Bednarski et al., 2009).

Finally, LYZ and *STAT1* genes, observed to be expressed at high levels in a subset of tissue monocytes and activated macrophages, were also included in this model, and correlated with prolonged FFS and better outcome. These results are in contrast with other published data (Sanchez-Aguilera et al., 2006; Steidl et al.) in which a signature overexpressed in patients with unfavorable outcomes was found, mainly including genes expressed by specific subpopulations of macrophages. The discrepancy in the results concerning the role of

macrophages may have arisen from technical differences (RT-PCR vs. microarray gene-expression and immunohistochemistry) or the selection of markers such as *LYZ* and *STAT1* in this study, reflecting a specific functional status of the monocyte-macrophages. Different series of patients have been investigated in the different studies using three main techniques: GEP, IHC and QT-PCR. Distinct macrophages markers (LYZ, STAT1, CD68 and CD163) illustrating different functional subpopulations have been found associated with favorable or unfavorable outcomes in these series. The differences in the different studies are too important as for reaching a simple conclusion; this is why in the MRS model LYZ and STAT1 were denominated as markers of activated macrophages, avoiding the use of the term macrophage-markers. It is not clear whether the expression of these markers (STAT1 and LYZ)) is related or not with the number of macrophages, but both are markers either of specific macrophage subpopulation or specific stages of macrophage activation.

Importantly, RT-PCR studies do not intend to restrict the analysis to specific cell types, and are measuring RNA extracted from the whole tissue. Thus, we can´t conclude whether there is any inconsistency between the results, or if there is a contradiction between results obtained by others and our data. Different things are compared using diverse techniques in series of patients that have been collected using different criteria (advanced HL or just all classic HL stages). This fact indicates that prognostic parameters are strictly dependent of the technique used and the sample selection. Nevertheless, with the adequate case selection and using multicenter studies, general conclusions can be obtained when prospective validation studies are performed.

5.3 Gene expression profiling studies in cHL and survival.

5.3.1 Identifying genes associated with disease outcome in CHL

To date (year 2011), only four studies based on gene expression profiling have been published and reported by different teams, including the one from this thesis work. All of them using nonmicrodiseccted frozen samples (whole tissue samples) obtained from patients with cHL during diagnostic lymp-node biopsies and leading to the discovery of new prognostic factors. In general, the number of frozen profiles samples studied was relatively small with exception of the study reported from Steidl et al. (Table 5.2).

Table 5.2 Gene expression profiling of *in silico* microdissection

Reference	cHL tumor samples	Microarray platform Nº of genes	Control samples	Gene signature associated with	Validation set	Multivariate analysis
(Devilard et al., 2002)	21 samples limited and advanced stages	~1000 genes	Cell lines Tissues (adenitis, lymphomas)	Sustained CR	No	No
(Sanchez-Aguilera et al., 2006)	29 samples advanced stages	~9.348 genes	Cell lines	Sustained CR	250 samples IHC TMA	No
(Chetaille et al., 2009)	63 samples limited and advanced stages	~16.000 genes	Cell lines Tissues (adenitis, lymphomas)	Sustained CR	146 samples IHC TMA	EFS OS
(Steidl et al.)	130 samples limited and advanced stages	~25.000 genes	No	Absence of P	166 samples IHC TMA	PFS OS

cHL_ classical Hodgkin lymphoma; RS: Reed Sternberg, IHC: immunohistochemistry; TMA: Tissue microarray; EFS; event-free survival;CR: complete remission; P: progression; OS: Overall survival; PFS: progression-free survival, DSS: disease-specific survival

Despite some limitations inherent to all the studies, such as the use of different control samples, there are some common conclusions to most of them, such as the relevance of B cells and macrophages on patient outcome, thus pointing out to the important role of both the tumoral component and the microenvironment, as initially hypothesized in this work. Additionally, signatures found in this work by the GSEA analysis and the TLDA RT-PCR arrays highlighted the same functions identified by other studies (Table 5.3), confirming their results.

In addition to the already performed studies, the work done in this thesis adds something new, being an interesting (although preliminary attempt) to translate gene profiling results (always dependent from frozen sample availability) to the routine diagnostic clinical setting. Here, the design of a real time PCR –based low density array including relevant genes identified by reanalyzing gene expression data and finally leading to the development of an score (Molecular Risk Score _ MRS) able to identify patients with different risks to treatment failure, is presented .

Table 5.3. Gene expression signatures outcome related in cHL.

Technique	N	HRS cells signatures	Microenvironmental signatures	Ref
GEP	29	Cell cycle, apoptosis, signals transduction.	Immune response, extracellular matrix, adhesion and cell- cell signaling	**(Sanchez-Aguilera et al., 2006)**
RT-PCR array	52	Cell cycle (G2-M and G1 pathway s) , histones , chaperones and MAP kinase pathway,	T cell pathway , Monocyte Macrophage	**(Sanchez-Espiridion et al., 2009)**
RT-PCR array	282	Cell cycle, apoptosis	Macrophage activation, IRF4	**(Sanchez-Espiridion et al., 2010)**
GEP	63	BCR signaling, apoptosis, cell metabolism	Stroma remodeling genes	**(Chetaille et al., 2009)**
GEP	130	HRS cell related genes	Macrophages, Monocytes Metaloproteinases, angiogenic cells	**(Steidl et al.)**

N_number of samples; ref_ reference;GEP_ Gene expression profiling;RT-PCR_ reverse-transcription polymerase chain reaction. Adapted from (Derenzini and Younes, 2011)

5.3.2 Macrophages and outcome in Hodgkin lymphoma

The main conclusion that arises from the different works done till now is that the presence and features tumor associated macrophages is related with the outcome of classical Hodgkin lymphoma. Multiple reports now support the value of tumor associated macrophages in pretreatment biopsies for outcome prediction in classical cHL, although there is still some debate about the precise markers to be analyzed. Most important evidences have been suggested from gene expression profiling studies (Table 5.2) of large patient´s series, as also done in this work. Interestingly, from these studies two markers (CD68, CD163) in particular have emerged as putatively useful immunohistochemical markers. The different studies used gene expression profiling to identify a gene signature of tumor-associated macrophages associated with treatment failure, in an approach methodologically similar to the one here followed.

The exact role that macrophages play is still under discussion, their activation stage and immune polarization pattern (thus, balance between Macrophage type 2 (M2) or type1 (M1) populations, each one having different biological effects, seems to be critical in the final

outcome of patients. Conclusions are out of the scope this thesis work, but our GSEA analyses also identified a set of genes (ALDH1A1, LYZ, and STAT1) related to monocyte /macrophage as overexpresed in patient with unfavorable outcomes. Moreover, Molecular Risk Score Model included also genes related to macrophages as associated to treatment outcome, thus again highlighting their role in cHL outcome prediction.

5.3.3 Influence of other populations (GSEA pathways).

Data obtained either by GSEA analyses and the subsequent development of MRS reinforced previous observations done by previous works concerning the influence of B cells, different T-cell subsets (T regulatory cells, T helper, cytotoxic T cells), Natural killer (NK) cells, and stromal components in the cHL outcome. Besides further detailed molecular characterization of HRS cells, it appears that cellular enrichment, and a better characterization of macrophages, B cells and T cells from primary clinical material will be critical to gain further insight into the unique phenotype and profile of cHL. Therefore, as observed by the signatures here identified, in classical Hodgkin lymphoma many processes (related to either both HRS cells or microenvironment) are implicated in patient´s outcome. A better understanding of this complex scenario will help in the identification of patients at high risk of relapse.

5.4. MicroRNAs and classical Hodgkin lymphoma. New insights into outcome prediction and pathogenesis.

In the second part of this thesis work, we decided to investigate the relevance of microRNA expression in cHL and used miRNA expression data from the analysis of advanced cHL samples and cHL-derived cell lines to identify identify specific miRNA profiles from the tumoral cells and their non-tumoral microenvironment, that were associated with lack of maintained response, following the same schema as for the one followed for MRS development. Functional and practical relevance of microRNA expression signatures in advanced cHL was investigated. Thus, using microarray analyses of freshly frozen tissue we identified a specific 34 miRNA signature that included miRNA deregulated in the microenvironment and in the tumor cells. Further analyses using RT-qPCR in an independent series of 220 advanced cHL patients derived a four miRNA signature able to classify cHL patients according to treatment response.

This signature included already described oncomirs such as miR-21 and miR-30d (Kumar et al.; Medina et al.). In agreement with these results, overexpression of miR-21 has

been reported in cHL and other lymphoma types (Gibcus et al., 2009; Navarro et al., 2008; Rossi et al., 2010). Additionally, its overexpression has been associated with poorer prognosis in many tumor types: squamous cell carcinoma (Gao et al., 2011), non-small cell carcinoma of the lung(Gao et al., 2011), esophageal cancer (Hummel et al., 2011), colorectal adenocarcinoma (Kulda et al., 2010), pancreatic cancer (Hwang et al., 2010), and astrocytoma (Zhi et al., 2010). Also, anti-miR-21 treatment inhibits cell growth in vitro and tumor growth in a breast cancer mouse model (Si et al., 2007). Moreover, a mouse model conditionally expressing miR-21 lead to a pre-B-cell lymphoma (Medina et al.), and when miR-21 was inactivated, the tumors regressed completely.

The other interesting member of the identified signature related to poorer outcome is miR-30d that has been reported to promote tumor invasion and metastasis, in hepatocellular carcinoma, and a prognostic signature for this tumor also included this miRNA (Yao et al., 2010). MiR-30d is co amplified with miR-30b (8q24.22-23) in medulloblastoma, independent of MYC amplification (Lu et al., 2009). Also, inhibition of miR-30d expression increases endogenous p53 expression and induces cellular apoptosis in several cell lines, including one from multiple myeloma without TP53 mutations (Kumar et al.).

Finally, the other two members of the identified signature (miR-30e and miR-92b*) have been also related to several cancer types. Thus, miR-30e has been related to metastasis-related hepatocellular carcinoma (Budhu et al., 2008) whereas miR-92b* has been found to be overexpressed in brain primary tumors.

Due to the oncogenic features of these miRNAs found in other types of cancer, in Hodgkin lymphoma the overexpression of miR-21,miR-30d, miR-30 and miR-92b* could have a functional relevance in the prognostication of the disease course and could represent interesting therapeutic targets. Moreover, miR-21 and miR-30d, are thought to target apoptosis regulation (BCL2, BCL-XL, BCL11, BAX, PCD4, RAS) known to be altered in cHL, thus suggesting an interesting link among miRNA abnormal expression and the characteristic deregulated apoptosis of HRS cells. In agreement with this hypothesis, correlation analyses between our gene and miRNA expression data supported this relationship. Nevertheless, despite our promising results, this has to be further validated by functional analyses in cHL cell lines looking for targets and putative mechanisms underlying the observed effects (see 5.6 Perspectives).

5.5 A combined approach_ miRNAS and genes.

As commented, transcriptional analysis has identified gene, miRNAs and pathways associated with clinical failure in HL, but both the biological relevance and clinical applicability of these data are still pending of further development. Thus, a combined analysis of samples including both expression levels (genes and miRNAs) could help in a better understanding in the cHL malignancy.

The main goal of this work was to search for a gene and miRNA signature associated with clinical outcome and treatment response in cHL patients treated with ABVD or variants using FFPE tissue samples. Interestingly, when combined both signatures significant interactions where found between genes related to apoptosis (BCL2, BCL2L1 and CASP3) and Cell cycle (CCNA2, CCNE2, CENPF and CDC2) with miR-21 and miR-30d expression, thus probably indicating direct modulation of the intrinsic apoptosis pathway and the cell cycle, respectively.

Taken together, these results suggest that overexpression of these miRNAs could inhibit drug-induced apoptosis, and could be linked with the clinical resistance to chemotherapeutic therapy of cHL.

5.6 Perspectives. miRNAs, translating Basic Science into Therapy.

The emergence of miRNAs as important cancer prevention genes is likely to have a large effect on gene therapies that are designed to block tumor progression. Large-scale expression screens, similar to those described in this thesis work, that compare miRNA levels in tumors with different molecular phenotypes; will be useful for identifying novel miRNAs that are involved in chemoresistance and biological processes underlying cHL pathogenesis. Functional screens designed to select miRNA genes that specially control processes known to be altered in cHL, such as apoptosis will also help in this search. Interestingly, this study highlighted the abnormal expression of a small set of miRNAS, including miR-21 (expressed in different tumor types as mentioned). In the future, the administration of synthetic anti-sense oligonucleotides that encode sequences complementary to mature oncogenic miRNAs might effectively inactivate miRNAs in tumors and slow their growth. Preliminary data using cHL cell lines (L428 and L1236) suggest that inhibition of two of miRNAs from the identified signature (miR-21 and miR-30d) sensitizes cHL cells to doxorubicin treatment and thus might constitutive an alternative therapeutic approach in cHL.

CONCLUSIONS / CONCLUSIONES

CONCLUSIONS/CONCLUSIONES

1. Specific gene and miRNA signatures for the HRS cells and their microenvironment can be identified in cHL tissue samples, and improve the understanding of cHL pathogenesis.

2. Specific gene and miRNA patterns can be identified by comparing expression data from two completely opposed biological situations (treatment responders (Favorable outcome) versus non-responders (Unfavorable outcome) being GSEA analysis a convenient tool for elucidation of gene pathways and biological processes associated.

3. Treatment response in advanced cHL patients is partially determined by a combination of factors related to both either the HRS cells and their reactive microenvironment. Classical HL microenvironment plays a role in the disease pathogenesis, promoting tumor growth, survival and immune escape. Thus, treatment response in cHL was observed to be related with:

    a. Gene sets / functional pathways enriched in the HRS cells: G2/M checkpoint of the cell cycle, Regulation of the S phase and G1/S transition, Histone modification, Drug metabolism, Chaperons and the MAP kinase signaling pathways.

    b. Gene sets / functional pathways enriched in the microenvironment: T-cell populations: T-cytotoxic cells, Natural Killer cells and Macrophage activation.

4. Real time quantitative PCR (RT-qPCR) provides an effective technique for analyzing gene and miRNA expression in RNA and miRNA isolated from FFPE tissues. Additionally, it can be used for clinical prediction in HL paraffin-embedded diagnostic samples, using a selection of genes and miRNAs identified after analysis of the initial molecular signatures.

5. An 11-gene model including 4 functional pathways (cell cycle, apoptosis, macrophage activation and interferon regulatory factor 4) can identify groups of patients with different risks of treatment response and failure. This model defined as Molecular Risk Score (MRS) can be combined with Stage IV to improve patient´s stratification and derive an Integrative model able to identify a subgroup of patients with very bad outcome.

6. A 4-miRNA signature is related to outcome in advanced cHL patients and can stratify patients (high- versus low-risk patients) according to their median expression levels. This signature includes previously described oncomirs such as miR-21, miR-92b* and other relevant regulators of cell functions such as miR-30d and miR-30e.

CONCLUSIONES

1. En linfoma de Hodgkin, es posible la identificación de firmas moleculares de genes y microRNAS características de ambos componentes tumorales, células de Hodgkin y Reed Sternberg y su microambiente tumoral, mediante el análisis de muestras tumorales.

2. Perfiles específicos de expresión de genes y microRNAs pueden ser identificados por comparación de datos de expresión de fenómenos biológicos opuestos, tales como respuesta a tratamiento (pacientes respondedores vs pacientes no respondedores), constituyendo la herramienta bioinformática GSEA de gran utilidad cara la caracterización de las rutas y fenómenos biológicos implicados.

3. La respuesta a terapia en pacientes con linfoma de Hodgkin está determinada por una combinación de factores relacionados con las propias células neoplásicas y su complejo microambiente. Dicho microentorno constituye una parte funcional de la enfermedad con un papel relevante en la promoción del crecimiento tumoral y la supervivencia del tumor. Así, la respuesta a terapia en linfoma de Hodgkin clásico se asocia con alteraciones en:

   a. Rutas biológicas y genes atribuibles al componente tumoral (células HRS) l: puntos de control mitótico G2/M, G1/S y fase S), modificación de histonas, metabolismo de drogas, chaperonas y la ruta de señalización de las MAP quinasas.

   b. Poblaciones celulares específicas del microambiente tumoral: poblaciones de células T, células T citotóxicas, células Natural Killer (NK) y distintos estados de activación de macrófagos.

4. Los ensayos basados en PCR cuantitativa constituyen una aproximación robusta y fiable para el análisis de expresión de genes y microRNAs en muestras parafinadas con potencial aplicabilidad clínica.

5. Se ha desarrollado un modelo predictivo de fallo terapéutico basado en la expresión de 11 genes agrupados en sus correspondientes grupos funcionales (Ciclo celular, apoptosis, activación de macrófagos y factor regulador de interferon (IRF4)). El algoritmo desarrollado (Riesgo Molecular) es capaz de identificar subgrupos de pacientes con diferentes probabilidades de fallo. Su combinación con la variable clínica estadio IV permite la identificación de un subgrupo de pacientes de muy mal pronóstico.

6. Se ha identificado una firma molecular de 4 microRNAs asociada a pronóstico en linfoma de Hodgkin avanzado. La expresión miR-21, miR-30d, miR-30e y miR-92b* permite la estratificación de pacientes en dos grupos de riesgo (alto versus bajo riesgo) acorde con su mediana de expresión.

REFERENCES

Aldinucci, D., Lorenzon, D., Olivo, K., Rapana, B., and Gattei, V. (2004). Interactions between tissue fibroblasts in lymph nodes and Hodgkin/Reed-Sternberg cells. Leuk Lymphoma *45*, 1731-1739.

Aldinucci, D., Rapana, B., Olivo, K., Lorenzon, D., Gloghini, A., Colombatti, A., and Carbone, A. IRF4 is modulated by CD40L and by apoptotic and anti-proliferative signals in Hodgkin lymphoma. Br J Haematol *148*, 115-118.

Alizadeh, A.A., Ross, D.T., Perou, C.M., and van de Rijn, M. (2001). Towards a novel classification of human malignancies based on gene expression patterns. J Pathol *195*, 41-52.

Alvaro, T., Lejeune, M., Salvado, M.T., Bosch, R., Garcia, J.F., Jaen, J., Banham, A.H., Roncador, G., Montalban, C., and Piris, M.A. (2005). Outcome in Hodgkin's lymphoma can be predicted from the presence of accompanying cytotoxic and regulatory T cells. Clin Cancer Res *11*, 1467-1473.

Antonov, J., Goldstein, D.R., Oberli, A., Baltzer, A., Pirotta, M., Fleischmann, A., Altermatt, H.J., and Jaggi, R. (2005). Reliable gene expression measurements from degraded RNA by quantitative real-time PCR depend on short amplicons and a proper normalization. Lab Invest *85*, 1040-1050.

April, C., Klotzle, B., Royce, T., Wickham-Garcia, E., Boyaniwsky, T., Izzo, J., Cox, D., Jones, W., Rubio, R., Holton, K.*, et al.* (2009). Whole-genome gene expression profiling of formalin-fixed, paraffin-embedded tissue samples. PloS one *4*, e8162.

Azambuja, D., Natkunam, Y., Biasoli, I., Lossos, I.S., Anderson, M.W., Morais, J.C., and Spector, N. (2011). Lack of association of tumor-associated macrophages with clinical outcome in patients with classical Hodgkin's lymphoma. Ann Oncol.

Bai, M., Papoudou-Bai, A., Horianopoulos, N., Grepi, C., Agnantis, N.J., and Kanavaros, P. (2007). Expression of bcl2 family proteins and active caspase 3 in classical Hodgkin's lymphomas. Hum Pathol *38*, 103-113.

Bai, M., Tsanou, E., Agnantis, N.J., Kamina, S., Grepi, C., Stefanaki, K., Rontogianni, D., Galani, V., and Kanavaros, P. (2004). Proliferation profile of classical Hodgkin's lymphomas. Increased expression of the protein cyclin D2 in Hodgkin's and Reed-Sternberg cells. Mod Pathol *17*, 1338-1345.

Bargou, R.C., Leng, C., Krappmann, D., Emmerich, F., Mapara, M.Y., Bommert, K., Royer, H.D., Scheidereit, C., and Dorken, B. (1996). High-level nuclear NF-kappa B and Oct-2 is a common feature of cultured Hodgkin/Reed-Sternberg cells. Blood *87*, 4340-4347.

Barrington, R.A., Schneider, T.J., Pitcher, L.A., Mempel, T.R., Ma, M., Barteneva, N.S., and Carroll, M.C. (2009). Uncoupling CD21 and CD19 of the B-cell coreceptor. Proc Natl Acad Sci U S A *106*, 14490-14495.

Barth, S., Pfuhl, T., Mamiani, A., Ehses, C., Roemer, K., Kremmer, E., Jaker, C., Hock, J., Meister, G., and Grasser, F.A. (2008). Epstein-Barr virus-encoded microRNA miR-BART2 down-regulates the viral DNA polymerase BALF5. Nucleic Acids Res *36*, 666-675.

Bechtel, D., Kurth, J., Unkel, C., and Kuppers, R. (2005). Transformation of BCR-deficient germinal-center B cells by EBV supports a major role of the virus in the pathogenesis of Hodgkin and posttransplantation lymphomas. Blood *106*, 4345-4350.

Bednarski, B.K., Baldwin, A.S., Jr., and Kim, H.J. (2009). Addressing reported pro-apoptotic functions of NF-kappaB: targeted inhibition of canonical NF-kappaB enhances the apoptotic effects of doxorubicin. PloS one *4*, e6992.

Blum, K.A. (2010). Upcoming diagnostic and therapeutic developments in classical Hodgkin's lymphoma. Hematology Am Soc Hematol Educ Program *2010*, 93-100.

Bonadonna, G., Viviani, S., Bonfante, V., Gianni, A.M., and Valagussa, P. (2005). Survival in Hodgkin's disease patients--report of 25 years of experience at the Milan Cancer Institute. Eur J Cancer *41*, 998-1006.

Budhu, A., Jia, H.L., Forgues, M., Liu, C.G., Goldstein, D., Lam, A., Zanetti, K.A., Ye, Q.H., Qin, L.X., Croce, C.M.*, et al.* (2008). Identification of metastasis-related microRNAs in hepatocellular carcinoma. Hepatology *47*, 897-907.

Cabannes, E., Khan, G., Aillet, F., Jarrett, R.F., and Hay, R.T. (1999). Mutations in the IkBa gene in Hodgkin's disease suggest a tumour suppressor role for IkappaBalpha. Oncogene *18*, 3063-3070.

Cai, X., Schafer, A., Lu, S., Bilello, J.P., Desrosiers, R.C., Edwards, R., Raab-Traub, N., and Cullen, B.R. (2006). Epstein-Barr virus microRNAs are evolutionarily conserved and differentially expressed. PLoS Pathog *2*, e23.

Calin, G.A., Liu, C.G., Sevignani, C., Ferracin, M., Felli, N., Dumitru, C.D., Shimizu, M., Cimmino, A., Zupo, S., Dono, M.*, et al.* (2004). MicroRNA profiling reveals distinct signatures in B cell chronic lymphocytic leukemias. Proc Natl Acad Sci U S A *101*, 11755-11760.

Canellos, G.P., and Niedzwiecki, D. (2002). Long-term follow-up of Hodgkin's disease trial. N Engl J Med *346*, 1417-1418.

Carbone, A., Gloghini, A., Gaidano, G., Franceschi, S., Capello, D., Drexler, H.G., Falini, B., and Dalla-Favera, R. (1998). Expression status of BCL-6 and syndecan-1 identifies distinct histogenetic subtypes of Hodgkin's disease. Blood *92*, 2220-2228.

Carbone, A., Gloghini, A., Larocca, L.M., Antinori, A., Falini, B., Tirelli, U., Dalla-Favera, R., and Gaidano, G. (1999). Human immunodeficiency virus-associated Hodgkin's disease derives from post-germinal center B cells. Blood *93*, 2319-2326.

Carde, P., Koscielny, S., Franklin, J., Axdorph, U., Raemaekers, J., Diehl, V., Aleman, B., Brosteanu, O., Hasenclever, D., Oberlin, O.*, et al.* (2002). Early response to chemotherapy: a surrogate for final outcome of Hodgkin's disease patients that should influence initial treatment length and intensity? Ann Oncol *13 Suppl 1*, 86-91.

Casasnovas, R.O., Mounier, N., Brice, P., Divine, M., Morschhauser, F., Gabarre, J., Blay, J.Y., Voillat, L., Lederlin, P., Stamatoullas, A.*, et al.* (2007). Plasma cytokine and soluble receptor signature predicts outcome of patients with classical Hodgkin's lymphoma: a study from the Groupe d'Etude des Lymphomes de l'Adulte. J Clin Oncol *25*, 1732-1740.

Chang, D.Z., Zhang, J.X., Filippa, D.A., and Portlock, C.S. (2004). Unusual abdominal tumors, case 2. Localized amyloid associated with gastric adenocarcinoma. J Clin Oncol *22*, 1520-1522.

Chen, Y.T., Kitabayashi, N., Zhou, X.K., Fahey, T.J., 3rd, and Scognamiglio, T. (2008). MicroRNA analysis as a potential diagnostic tool for papillary thyroid carcinoma. Mod Pathol *21*, 1139-1146.

Chetaille, B., Bertucci, F., Finetti, P., Esterni, B., Stamatoullas, A., Picquenot, J.M., Copin, M.C., Morschhauser, F., Casasnovas, O., Petrella, T.*, et al.* (2009). Molecular profiling of classical Hodgkin lymphoma tissues uncovers variations in the tumor microenvironment and correlations with EBV infection and outcome. Blood *113*, 2765-3775.

Choy, E.Y., Siu, K.L., Kok, K.H., Lung, R.W., Tsang, C.M., To, K.F., Kwong, D.L., Tsao, S.W., and Jin, D.Y. (2008). An Epstein-Barr virus-encoded microRNA targets PUMA to promote host cell survival. J Exp Med *205*, 2551-2560.

Chu, F., and Wang, L. (2005). Applications of support vector machines to cancer classification with microarray data. Int J Neural Syst *15*, 475-484.

Ciotti, P., Garuti, A., Ballestrero, A., Cirmena, G., Chiaramondia, M., Baccini, P., Bellone, E., and Mandich, P. (2009). Reliability and reproducibility of a RNA preamplification method for low-density array analysis from formalin-fixed paraffin-embedded breast cancer samples. Diagn Mol Pathol *18*, 112-118.

Connors, J.M. (2005). State-of-the-art therapeutics: Hodgkin's lymphoma. J Clin Oncol *23*, 6400-6408.

Cossman, J., Annunziata, C.M., Barash, S., Staudt, L., Dillon, P., He, W.W., Ricciardi-Castagnoli, P., Rosen, C.A., and Carter, K.C. (1999). Reed-Sternberg cell genome expression supports a B-cell lineage. Blood *94*, 411-416.

Dave, S.S., Wright, G., Tan, B., Rosenwald, A., Gascoyne, R.D., Chan, W.C., Fisher, R.I., Braziel, R.M., Rimsza, L.M., Grogan, T.M.*, et al.* (2004). Prediction of survival in follicular lymphoma based on molecular features of tumor-infiltrating immune cells. N Engl J Med *351*, 2159-2169.

Derenzini, E., and Younes, A. (2011). Predicting treatment outcome in classical Hodgkin lymphoma: genomic advances. Genome Med *3*, 26.

Devilard, E., Bertucci, F., Trempat, P., Bouabdallah, R., Loriod, B., Giaconia, A., Brousset, P., Granjeaud, S., Nguyen, C., Birnbaum, D.*, et al.* (2002). Gene expression profiling defines molecular subtypes of classical Hodgkin's disease. Oncogene *21*, 3095-3102.

Doleshal, M., Magotra, A.A., Choudhury, B., Cannon, B.D., Labourier, E., and Szafranska, A.E. (2008). Evaluation and validation of total RNA extraction methods for microRNA expression analyses in formalin-fixed, paraffin-embedded tissues. J Mol Diagn *10*, 203-211.

Dutton, A., Reynolds, G.M., Dawson, C.W., Young, L.S., and Murray, P.G. (2005). Constitutive activation of phosphatidyl-inositide 3 kinase contributes to the survival of Hodgkin's lymphoma cells through a mechanism involving Akt kinase and mTOR. J Pathol *205*, 498-506

.

Falzetti, D., Crescenzi, B., Matteuci, C., Falini, B., Martelli, M.F., Van Den Berghe, H., and Mecucci, C. (1999). Genomic instability and recurrent breakpoints are main cytogenetic findings in Hodgkin's disease. Haematologica *84*, 298-305.

Farinha, P., Masoudi, H., Skinnider, B.F., Shumansky, K., Spinelli, J.J., Gill, K., Klasa, R., Voss, N., Connors, J.M., and Gascoyne, R.D. (2005). Analysis of multiple biomarkers shows that lymphoma-associated macrophage (LAM) content is an independent predictor of survival in follicular lymphoma (FL). Blood *106*, 2169-2174.

Farrugia, D.C., Ford, H.E., Cunningham, D., Danenberg, K.D., Danenberg, P.V., Brabender, J., McVicar, A.D., Aherne, G.W., Hardcastle, A., McCarthy, K*., et al.* (2003). Thymidylate synthase expression in advanced colorectal cancer predicts for response to raltitrexed. Clin Cancer Res *9*, 792-801.

Fiumara, P., Snell, V., Li, Y., Mukhopadhyay, A., Younes, M., Gillenwater, A.M., Cabanillas, F., Aggarwal, B.B., and Younes, A. (2001). Functional expression of receptor activator of nuclear factor kappaB in Hodgkin disease cell lines. Blood *98*, 2784-2790.

Friedman, R.C., Farh, K.K., Burge, C.B., and Bartel, D.P. (2009). Most mammalian mRNAs are conserved targets of microRNAs. Genome Res *19*, 92-105.

Gao, W., Shen, H., Liu, L., Xu, J., and Shu, Y. (2011). MiR-21 overexpression in human primary squamous cell lung carcinoma is associated with poor patient prognosis. J Cancer Res Clin Oncol *137*, 557-566.

Garcia, J.F., Camacho, F.I., Morente, M., Fraga, M., Montalban, C., Alvaro, T., Bellas, C., Castano, A., Diez, A., Flores, T*., et al.* (2003). Hodgkin and Reed-Sternberg cells harbor alterations in the major tumor suppressor pathways and cell-cycle checkpoints: analyses using tissue microarrays. Blood *101*, 681-689.

Garcia, J.F., Villuendas, R., Algara, P., Saez, A.I., Sanchez-Verde, L., Martinez-Montero, J.C., Martinez, P., and Piris, M.A. (1999). Loss of p16 protein expression associated with methylation of the p16INK4A gene is a frequent finding in Hodgkin's disease. Lab Invest *79*, 1453-1459.

Garzon, R., Marcucci, G., and Croce, C.M. (2010). Targeting microRNAs in cancer: rationale, strategies and challenges. Nat Rev Drug Discov *9*, 775-789.

Gibcus, J.H., Tan, L.P., Harms, G., Schakel, R.N., de Jong, D., Blokzijl, T., Moller, P., Poppema, S., Kroesen, B.J., and van den Berg, A. (2009). Hodgkin lymphoma cell lines are characterized by a specific miRNA expression profile. Neoplasia *11*, 167-176.

Glenn, S.T., Jones, C.A., Liang, P., Kaushik, D., Gross, K.W., and Kim, H.L. (2007). Expression profiling of archival renal tumors by quantitative PCR to validate prognostic markers. Biotechniques *43*, 639-640, 642-633, 647.

Gobbi, P.G., Zinzani, P.L., Broglia, C., Comelli, M., Magagnoli, M., Federico, M., Merli, F., Iannitto, E., Tura, S., and Ascari, E. (2001). Comparison of prognostic models in patients with advanced Hodgkin disease. Promising results from integration of the best three systems. Cancer *91*, 1467-1478.

Godfrey, T.E., Kim, S.H., Chavira, M., Ruff, D.W., Warren, R.S., Gray, J.W., and Jensen, R.H. (2000). Quantitative mRNA expression analysis from formalin-fixed, paraffin-embedded tissues using 5' nuclease quantitative reverse transcription-polymerase chain reaction. J Mol Diagn *2*, 84-91.

Green, G.H., and Diggle, P.J. (2007). On the operational characteristics of the Benjamini and Hochberg False Discovery Rate procedure. Stat Appl Genet Mol Biol *6*, Article27.
Grenert, J.P., Smith, A., Ruan, W., Pillai, R., and Wu, A.H. (2011). Gene expression profiling from formalin-fixed, paraffin-embedded tissue for tumor diagnosis. Clin Chim Acta *412*, 1462-1464.

Griffiths-Jones, S. (2004). The microRNA Registry. Nucleic Acids Res *32*, D109-111.

Griffiths-Jones, S., Grocock, R.J., van Dongen, S., Bateman, A., and Enright, A.J. (2006). miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res *34*, D140-144.

Griffiths-Jones, S., Saini, H.K., van Dongen, S., and Enright, A.J. (2008). miRBase: tools for microRNA genomics. Nucleic Acids Res *36*, D154-158.

Grimson, A., Farh, K.K., Johnston, W.K., Garrett-Engele, P., Lim, L.P., and Bartel, D.P. (2007). MicroRNA targeting specificity in mammals: determinants beyond seed pairing. Mol Cell *27*, 91-105.

Gutensohn, N.M. (1982). Social class and age at diagnosis of Hodgkin's disease: new epidemiologic evidence for the "two-disease hypothesis". Cancer Treat Rep *66*, 689-695.

Hammond, S.M. (2006). MicroRNAs as oncogenes. Curr Opin Genet Dev *16*, 4-9.

Hammond, S.M. (2007). MicroRNAs as tumor suppressors. Nat Genet *39*, 582-583.

Hasenclever, D., and Diehl, V. (1998). A prognostic score for advanced Hodgkin's disease. International Prognostic Factors Project on Advanced Hodgkin's Disease. N Engl J Med *339*, 1506-1514.

Herbst, H., Dallenbach, F., Hummel, M., Niedobitek, G., Pileri, S., Muller-Lantzsch, N., and Stein, H. (1991). Epstein-Barr virus latent membrane protein expression in Hodgkin and Reed-Sternberg cells. Proc Natl Acad Sci U S A *88*, 4766-4770.

Hertel, C.B., Zhou, X.G., Hamilton-Dutoit, S.J., and Junker, S. (2002). Loss of B cell identity correlates with loss of B cell-specific transcription factors in Hodgkin/Reed-Sternberg cells of classical Hodgkin lymphoma. Oncogene *21*, 4908-4920.

Hilliker, C., Delabie, J., Speleman, F., Bilbe, G., Bruggen, J., Van Leuven, F., and Van den Berghe, H. (1994). Localization of the gene (RSN) coding for restin, a marker for Reed-Sternberg cells in Hodgkin's disease, to human chromosome band 12q24.3 and YAC cloning of the locus. Cytogenet Cell Genet *65*, 172-176.

Horie, R., Watanabe, T., Morishita, Y., Ito, K., Ishida, T., Kanegae, Y., Saito, I., Higashihara, M., Mori, S., and Kadin, M.E. (2002). Ligand-independent signaling by overexpressed CD30 drives NF-kappaB activation in Hodgkin-Reed-Sternberg cells. Oncogene *21*, 2493-2503.

Hummel, R., Hussey, D.J., Michael, M.Z., Haier, J., Bruewer, M., Senninger, N., and Watson, D.I. (2011). MiRNAs and their association with locoregional staging and survival following surgery for esophageal carcinoma. Ann Surg Oncol *18*, 253-260.

Hwang, J.H., Voortman, J., Giovannetti, E., Steinberg, S.M., Leon, L.G., Kim, Y.T., Funel, N., Park, J.K., Kim, M.A., Kang, G.H*., et al.* (2010). Identification of microRNA-21 as a biomarker for chemoresistance and clinical outcome following adjuvant therapy in resectable pancreatic cancer. PloS one *5*, e10630.

Jaffe, E.S., Harris, N.L., Stein, H., and Vardiman, J.W. (2001). Pathology and genetics of tumours of haematopoietic and lymphoid tissues. World health organization classification of tumours (Lyon, IARC Press).

Joos, S., Kupper, M., Ohl, S., von Bonin, F., Mechtersheimer, G., Bentz, M., Marynen, P., Moller, P., Pfreundschuh, M., Trumper, L*., et al.* (2000). Genomic imbalances including amplification of the tyrosine kinase gene JAK2 in CD30+ Hodgkin cells. Cancer Res *60*, 549-552
.
Joos, S., Menz, C.K., Wrobel, G., Siebert, R., Gesk, S., Ohl, S., Mechtersheimer, G., Trumper, L., Moller, P., Lichter, P*., et al.* (2002). Classical Hodgkin lymphoma is characterized by recurrent copy number gains of the short arm of chromosome 2. Blood *99*, 1381-1387.

Jundt, F., Kley, K., Anagnostopoulos, I., Schulze Probsting, K., Greiner, A., Mathas, S., Scheidereit, C., Wirth, T., Stein, H., and Dorken, B. (2002). Loss of PU.1 expression is associated with defective immunoglobulin transcription in Hodgkin and Reed-Sternberg cells of classical Hodgkin disease. Blood *99*, 3060-3062.

Kallioniemi, O.P., Wagner, U., Kononen, J., and Sauter, G. (2001). Tissue microarray technology for high-throughput molecular profiling of cancer. Hum Mol Genet *10*, 657-662.

Kamper, P., Bendix, K., Hamilton-Dutoit, S., Honore, B., Nyengaard, J., and d'Amore, F. Tumor-infiltrating macrophages correlate with adverse prognosis and Epstein-Barr virus status in classical Hodgkin lymphoma. Haematologica.

Kamper, P., Bendix, K., Hamilton-Dutoit, S., Honore, B., Nyengaard, J.R., and d'Amore, F. (2011). Tumor-infiltrating macrophages correlate with adverse prognosis and Epstein-Barr virus status in classical Hodgkin's lymphoma. Haematologica *96*, 269-276.

Kanzler, H., Hansmann, M.L., Kapp, U., Wolf, J., Diehl, V., Rajewsky, K., and Kuppers, R. (1996a). Molecular single cell analysis demonstrates the derivation of a peripheral blood-derived cell line (L1236) from the Hodgkin/Reed-Sternberg cells of a Hodgkin's lymphoma patient. Blood *87*, 3429-3436.

Kanzler, H., Kuppers, R., Hansmann, M.L., and Rajewsky, K. (1996b). Hodgkin and Reed-Sternberg cells in Hodgkin's disease represent the outgrowth of a dominant tumor clone derived from (crippled) germinal center B cells. J Exp Med *184*, 1495-1505.

Kapp, U., Yeh, W.C., Patterson, B., Elia, A.J., Kagi, D., Ho, A., Hessel, A., Tipsword, M., Williams, A., Mirtsos, C*., et al.* (1999). Interleukin 13 is secreted by and stimulates the growth of Hodgkin and Reed-Sternberg cells. J Exp Med *189*, 1939-1946.

Kashkar, H., Haefs, C., Shin, H., Hamilton-Dutoit, S.J., Salvesen, G.S., Kronke, M., and Jurgensmeier, J.M. (2003). XIAP-mediated caspase inhibition in Hodgkin's lymphoma-derived B cells. J Exp Med *198*, 341-347.

Kelley, T.W., Pohlman, B., Elson, P., and Hsi, E.D. (2007). The ratio of FOXP3+ regulatory T cells to granzyme B+ cytotoxic T/NK cells predicts prognosis in classical Hodgkin lymphoma and is independent of bcl-2 and MAL expression. Am J Clin Pathol *128*, 958-965.

Kent, O.A., and Mendell, J.T. (2006). A small piece in the cancer puzzle: microRNAs as tumor suppressors and oncogenes. Oncogene *25*, 6188-6196.

Kewalramani, T., Zelenetz, A.D., Nimer, S.D., Portlock, C., Straus, D., Noy, A., O'Connor, O., Filippa, D.A., Teruya-Feldstein, J., Gencarelli, A.*, et al.* (2004). Rituximab and ICE as second-line therapy before autologous stem cell transplantation for relapsed or primary refractory diffuse large B-cell lymphoma. Blood *103*, 3684-3688.

Kluiver, J., Poppema, S., de Jong, D., Blokzijl, T., Harms, G., Jacobs, S., Kroesen, B.J., and van den Berg, A. (2005). BIC and miR-155 are highly expressed in Hodgkin, primary mediastinal and diffuse large B cell lymphomas. J Pathol *207*, 243-249.

Koch, I., Slotta-Huspenina, J., Hollweck, R., Anastasov, N., Hofler, H., Quintanilla-Martinez, L., and Fend, F. (2006). Real-time quantitative RT-PCR shows variable, assay-dependent sensitivity to formalin fixation: implications for direct comparison of transcript levels in paraffin-embedded tissues. Diagn Mol Pathol *15*, 149-156.

Kononen, J., Bubendorf, L., Kallioniemi, A., Barlund, M., Schraml, P., Leighton, S., Torhorst, J., Mihatsch, M.J., Sauter, G., and Kallioniemi, O.P. (1998). Tissue microarrays for high-throughput molecular profiling of tumor specimens. Nat Med *4*, 844-847.

Kulda, V., Pesta, M., Topolcan, O., Liska, V., Treska, V., Sutnar, A., Rupert, K., Ludvikova, M., Babuska, V., Holubec, L., Jr.*, et al.* (2010). Relevance of miR-21 and miR-143 expression in tissue samples of colorectal carcinoma and its liver metastases. Cancer genetics and cytogenetics *200*, 154-160.

Kumar, M., Lu, Z., Takwi, A.A., Chen, W., Callander, N.S., Ramos, K.S., Young, K.H., and Li, Y. Negative regulation of the tumor suppressor p53 gene by microRNAs. Oncogene *30*, 843-853.

Kupper, M., Joos, S., von Bonin, F., Daus, H., Pfreundschuh, M., Lichter, P., and Trumper, L. (2001). MDM2 gene amplification and lack of p53 point mutations in Hodgkin and Reed-Sternberg cells: results from single-cell polymerase chain reaction and molecular cytogenetic studies. Br J Haematol *112*, 768-775.

Kuppers, R., Klein, U., Schwering, I., Distler, V., Brauninger, A., Cattoretti, G., Tu, Y., Stolovitzky, G.A., Califano, A., Hansmann, M.L.*, et al.* (2003). Identification of Hodgkin and Reed-Sternberg cell-specific genes by gene expression profiling. J Clin Invest *111*, 529-537.

Kuppers, R., Rajewsky, K., Zhao, M., Simons, G., Laumann, R., Fischer, R., and Hansmann, M.L. (1994). Hodgkin disease: Hodgkin and Reed-Sternberg cells picked from histological sections show clonal immunoglobulin gene rearrangements and appear to be derived from B cells at various stages of development. Proc Natl Acad Sci U S A *91*, 10962-10966.

Kuppers, R., Zhao, M., Hansmann, M.L., and Rajewsky, K. (1993). Tracing B cell development in human germinal centres by molecular analysis of single cells picked from histological sections. EMBO J *12*, 4955-4967.

Lehmann, U., and Kreipe, H. (2001). Real-time PCR analysis of DNA and RNA extracted from formalin-fixed and paraffin-embedded biopsies. Methods *25*, 409-418.

Lewis, B.P., Burge, C.B., and Bartel, D.P. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell *120*, 15-20.
Llave, C., Xie, Z., Kasschau, K.D., and Carrington, J.C. (2002). Cleavage of Scarecrow-like mRNA targets directed by a class of Arabidopsis miRNA. Science *297*, 2053-2056.

Lo, A.K., To, K.F., Lo, K.W., Lung, R.W., Hui, J.W., Liao, G., and Hayward, S.D. (2007). Modulation of LMP1 protein expression by EBV-encoded microRNAs. Proc Natl Acad Sci U S A *104*, 16164-16169.

Lu, Y., Ryan, S.L., Elliott, D.J., Bignell, G.R., Futreal, P.A., Ellison, D.W., Bailey, S., and Clifford, S.C. (2009). Amplification and overexpression of Hsa-miR-30b, Hsa-miR-30d and KHDRBS3 at 8q24.22-q24.23 in medulloblastoma. PloS one *4*, e6159.

Lung, R.W., Tong, J.H., Sung, Y.M., Leung, P.S., Ng, D.C., Chau, S.L., Chan, A.W., Ng, E.K., Lo, K.W., and To, K.F. (2009). Modulation of LMP2A expression by a newly identified Epstein-Barr virus-encoded microRNA miR-BART22. Neoplasia *11*, 1174-1184.

Ma, M.M., Ding, J.W., and Xu, N. (2009). Odd-even width effect on persistent current in zigzag hexagonal graphene rings. Nanoscale *1*, 387-390.

Mack, T.M., Cozen, W., Shibata, D.K., Weiss, L.M., Nathwani, B.N., Hernandez, A.M., Taylor, C.R., Hamilton, A.S., Deapen, D.M., and Rappaport, E.B. (1995). Concordance for Hodgkin's disease in identical twins suggesting genetic susceptibility to the young-adult form of the disease. N Engl J Med *332*, 413-418.

Marafioti, T., Hummel, M., Foss, H.D., Laumen, H., Korbjuhn, P., Anagnostopoulos, I., Lammert, H., Demel, G., Theil, J., Wirth, T*., et al.* (2000). Hodgkin and reed-sternberg cells represent an expansion of a single clone originating from a germinal center B-cell with functional immunoglobulin gene rearrangements but defective immunoglobulin transcription. Blood *95*, 1443-1450.

Marshall, N.A., Christie, L.E., Munro, L.R., Culligan, D.J., Johnston, P.W., Barker, R.N., and Vickers, M.A. (2004). Immunosuppressive regulatory T cells are abundant in the reactive lymphocytes of Hodgkin lymphoma. Blood *103*, 1755-1762.

Martin-Subero, J.I., Gesk, S., Harder, L., Sonoki, T., Tucker, P.W., Schlegelberger, B., Grote, W., Novo, F.J., Calasanz, M.J., Hansmann, M.L*., et al.* (2002). Recurrent involvement of the REL and BCL11A loci in classical Hodgkin lymphoma. Blood *99*, 1474-1477
.
Martin-Subero, J.I., Klapper, W., Sotnikova, A., Callet-Bauchu, E., Harder, L., Bastard, C., Schmitz, R., Grohmann, S., Hoppner, J., Riemke, J*., et al.* (2006). Chromosomal breakpoints affecting immunoglobulin loci are recurrent in Hodgkin and Reed-Sternberg cells of classical Hodgkin lymphoma. Cancer Res *66*, 10332-10338.

Martin-Subero, J.I., Knippschild, U., Harder, L., Barth, T.F., Riemke, J., Grohmann, S., Gesk, S., Hoppner, J., Moller, P., Parwaresch, R.M.*, et al.* (2003). Segmental chromosomal aberrations and centrosome amplifications: pathogenetic mechanisms in Hodgkin and Reed-Sternberg cells of classical Hodgkin's lymphoma? Leukemia *17*, 2214-2219.

Mathas, S., Hinz, M., Anagnostopoulos, I., Krappmann, D., Lietz, A., Jundt, F., Bommert, K., Mechta-Grigoriou, F., Stein, H., Dorken, B.*, et al.* (2002). Aberrantly expressed c-Jun and JunB are a hallmark of Hodgkin lymphoma cells, stimulate proliferation and synergize with NF-kappa B. EMBO J *21*, 4104-4113.

Mathas, S., Lietz, A., Anagnostopoulos, I., Hummel, F., Wiesner, B., Janz, M., Jundt, F., Hirsch, B., Johrens-Leder, K., Vornlocher, H.P.*, et al.* (2004). c-FLIP mediates resistance of Hodgkin/Reed-Sternberg cells to death receptor-induced apoptosis. J Exp Med *199*, 1041-1052.

Medina, P.P., Nolde, M., and Slack, F.J. OncomiR addiction in an in vivo model of microRNA-21-induced pre-B-cell lymphoma. Nature *467*, 86-90.

Molin, D., Fischer, M., Xiang, Z., Larsson, U., Harvima, I., Venge, P., Nilsson, K., Sundstrom, C., Enblad, G., and Nilsson, G. (2001). Mast cells express functional CD30 ligand and are the predominant CD30L-positive cells in Hodgkin's disease. Br J Haematol *114*, 616-623.

Montalban, C., Garcia, J.F., Abraira, V., Gonzalez-Camacho, L., Morente, M.M., Bello, J.L., Conde, E., Cruz, M.A., Garcia-Sanz, R., Garcia-Larana, J.*, et al.* (2004). Influence of biologic markers on the outcome of Hodgkin's lymphoma: a study by the Spanish Hodgkin's Lymphoma Study Group. J Clin Oncol *22*, 1664-1673.

Morrissey, E.R., and Diaz-Uriarte, R. (2009). Pomelo II: finding differentially expressed genes. Nucleic Acids Res *37*, W581-586.

Mueller, N. (1992). Hodgkin's disease. In Cancer Epidemiology and Prevention, vol 2, D. Schnottenfeld, and J. Fraumeni, eds. (New York, Oxford University Press), p. 877.

Muschen, M., Re, D., Brauninger, A., Wolf, J., Hansmann, M.L., Diehl, V., Kuppers, R., and Rajewsky, K. (2000). Somatic mutations of the CD95 gene in Hodgkin and Reed-Sternberg cells. Cancer Res *60*, 5640-5643.

Navarro, A., Gaya, A., Martinez, A., Urbano-Ispizua, A., Pons, A., Balague, O., Gel, B., Abrisqueta, P., Lopez-Guillermo, A., Artells, R.*, et al.* (2007). MicroRNA expression profiling in classical Hodgkin lymphoma. Blood.

Navarro, A., Gaya, A., Martinez, A., Urbano-Ispizua, A., Pons, A., Balague, O., Gel, B., Abrisqueta, P., Lopez-Guillermo, A., Artells, R.*, et al.* (2008). MicroRNA expression profiling in classic Hodgkin lymphoma. Blood *111*, 2825-2832.

Ota, A., Tagawa, H., Karnan, S., Tsuzuki, S., Karpas, A., Kira, S., Yoshida, Y., and Seto, M. (2004). Identification and characterization of a novel gene, C13orf25, as a target for 13q31-q32 amplification in malignant lymphoma. Cancer Res *64*, 3087-3095.

Papadopoulos, G.L., Alexiou, P., Maragkakis, M., Reczko, M., and Hatzigeorgiou, A.G. (2009). DIANA-mirPath: Integrating human and mouse microRNAs in pathways. Bioinformatics *25*, 1991-1993.

Pileri, S.A., Ascani, S., Leoncini, L., Sabattini, E., Zinzani, P.L., Piccaluga, P.P., Pileri, A., Jr., Giunti, M., Falini, B., Bolis, G.B*., et al.* (2002). Hodgkin's lymphoma: the pathologist's viewpoint. J Clin Pathol *55*, 162-176.

Poppema, S., and van den Berg, A. (2000). Interaction between host T cells and Reed-Sternberg cells in Hodgkin lymphomas. Semin Cancer Biol *10*, 345-350.

Portlock, C.S., Donnelly, G.B., Qin, J., Straus, D., Yahalom, J., Zelenetz, A., Noy, A., O'Connor, O., Horwitz, S., Moskowitz, C*., et al.* (2004). Adverse prognostic significance of CD20 positive Reed-Sternberg cells in classical Hodgkin's disease. Br J Haematol *125*, 701-708.

Rajewsky, K. (1996). Clonal selection and learning in the antibody system. Nature *381*, 751-758.

Rassidakis, G.Z., Medeiros, L.J., Vassilakopoulos, T.P., Viviani, S., Bonfante, V., Nadali, G., Herling, M., Angelopoulou, M.K., Giardini, R., Chilosi, M*., et al.* (2002a). BCL-2 expression in Hodgkin and Reed-Sternberg cells of classical Hodgkin disease predicts a poorer prognosis in patients treated with ABVD or equivalent regimens. Blood *100*, 3935-3941.

Rassidakis, G.Z., Medeiros, L.J., Viviani, S., Bonfante, V., Nadali, G.P., Vassilakopoulos, T.P., Mesina, O., Herling, M., Angelopoulou, M.K., Giardini, R*., et al.* (2002b). CD20 expression in Hodgkin and Reed-Sternberg cells of classical Hodgkin's disease: associations with presenting features and clinical outcome. J Clin Oncol *20*, 1278-1287.

Re, D., Muschen, M., Ahmadi, T., Wickenhauser, C., Staratschek-Jox, A., Holtick, U., Diehl, V., and Wolf, J. (2001). Oct-2 and Bob-1 deficiency in Hodgkin and Reed Sternberg cells. Cancer Res *61*, 2080-2084.

Renna, S., Mocciaro, F., Perricone, G., Orlando, A., Virdone, R., Speciale, A., Lima, G., Stella, M., and Cottone, M. (2009). Is splenectomy a treatment option for aseptic abscesses in patients with Crohn's disease? Eur J Gastroenterol Hepatol *21*, 1314-1316.

Rossi, S., Shimizu, M., Barbarotto, E., Nicoloso, M.S., Dimitri, F., Sampath, D., Fabbri, M., Lerner, S., Barron, L.L., Rassenti, L.Z*., et al.* (2010). microRNA fingerprinting of CLL patients with chromosome 17p deletion identify a miR-21 score that stratifies early survival. Blood *116*, 945-952.

Sanchez-Aguilera, A., Montalban, C., de la Cueva, P., Sanchez-Verde, L., Morente, M.M., Garcia-Cosio, M., Garcia-Larana, J., Bellas, C., Provencio, M., Romagosa, V*., et al.* (2006). Tumor microenvironment and mitotic checkpoint are key factors in the outcome of classic Hodgkin lymphoma. Blood *108*, 662-668.

Sanchez-Espiridion, B., Montalban, C., Lopez, A., Menarguez, J., Sabin, P., Ruiz-Marcellan, C., Ramos, R., Rodriguez, J., Canovas, A., Camarero, C*., et al.* (2010). A molecular risk score based on 4 functional pathways for advanced classical Hodgkin lymphoma. Blood *116*, e12-17.

Sanchez-Espiridion, B., Sanchez-Aguilera, A., Montalban, C., Martin, C., Martinez, R., Gonzalez-Carrero, J., Poderos, C., Bellas, C., Fresno, M.F., Morante, C*., et al.* (2009). A TaqMan low-density array to predict outcome in advanced Hodgkin's lymphoma using paraffin-embedded samples. Clin Cancer Res *15*, 1367-1375.

Schwering, I., Brauninger, A., Distler, V., Jesdinsky, J., Diehl, V., Hansmann, M.L., Rajewsky, K., and Kuppers, R. (2003a). Profiling of Hodgkin's lymphoma cell line L1236 and germinal center B cells: identification of Hodgkin's lymphoma-specific genes. Mol Med *9*, 85-95.

Schwering, I., Brauninger, A., Klein, U., Jungnickel, B., Tinguely, M., Diehl, V., Hansmann, M.L., Dalla-Favera, R., Rajewsky, K., and Kuppers, R. (2003b). Loss of the B-lineage-specific gene expression program in Hodgkin and Reed-Sternberg cells of Hodgkin lymphoma. Blood *101*, 1505-1512.

Sethupathy, P., Corda, B., and Hatzigeorgiou, A.G. (2006). TarBase: A comprehensive database of experimentally supported animal microRNA targets. RNA *12*, 192-197.

Si, M.L., Zhu, S., Wu, H., Lu, Z., Wu, F., and Mo, Y.Y. (2007). miR-21-mediated tumor growth. Oncogene *26*, 2799-2803.

Sinha, N., Adhikari, N., and D, K.S. (2001). Effect of endosulfan during fetal gonadal differentiation on spermatogenesis in rats. Environ Toxicol Pharmacol *10*, 29-32.
Skinnider, B.F., and Mak, T.W. (2002). The role of cytokines in classical Hodgkin lymphoma. Blood *99*, 4283-4297.

Solberg, H.E. (1978). Discriminant analysis. CRC Crit Rev Clin Lab Sci *9*, 209-242.

Specht, K., Richter, T., Muller, U., Walch, A., Werner, M., and Hofler, H. (2001). Quantitative gene expression analysis in microdissected archival formalin-fixed and paraffin-embedded tumor tissue. Am J Pathol *158*, 419-429.

Steg, A., Wang, W., Blanquicett, C., Grunda, J.M., Eltoum, I.A., Wang, K., Buchsbaum, D.J., Vickers, S.M., Russo, S., Diasio, R.B.*, et al.* (2006). Multiple gene expression analyses in paraffin-embedded tissues by TaqMan low-density array: Application to hedgehog and Wnt pathway analysis in ovarian endometrioid adenocarcinoma. J Mol Diagn *8*, 76-83.

Steidl, C., Lee, T., Shah, S.P., Farinha, P., Han, G., Nayar, T., Delaney, A., Jones, S.J., Iqbal, J., Weisenburger, D.D.*, et al.* Tumor-associated macrophages and survival in classic Hodgkin's lymphoma. N Engl J Med *362*, 875-885.

Steidl, C., Lee, T., Shah, S.P., Farinha, P., Han, G., Nayar, T., Delaney, A., Jones, S.J., Iqbal, J., Weisenburger, D.D.*, et al.* (2010). Tumor-associated macrophages and survival in classic Hodgkin's lymphoma. N Engl J Med *362*, 875-885.

Stein, H., Marafioti, T., Foss, H.D., Laumen, H., Hummel, M., Anagnostopoulos, I., Wirth, T., Demel, G., and Falini, B. (2001). Down-regulation of BOB.1/OBF.1 and Oct2 in classical Hodgkin disease but not in lymphocyte predominant Hodgkin disease correlates with immunoglobulin transcription. Blood *97*, 496-501.

Stewart, G.D., Baird, J., Rae, F., Nanda, J., Riddick, A.C., and Harrison, D.J. (2011). Utilizing mRNA extracted from small, archival formalin-fixed paraffin-embedded prostate samples for translational research: assessment of the effect of increasing sample age and storage temperature. Int Urol Nephrol.

Straus, S.E., Jaffe, E.S., Puck, J.M., Dale, J.K., Elkon, K.B., Rosen-Wolff, A., Peters, A.M., Sneller, M.C., Hallahan, C.W., Wang, J.*, et al.* (2001). The development of lymphomas in families with

autoimmune lymphoproliferative syndrome with germline Fas mutations and defective lymphocyte apoptosis. Blood *98*, 194-200.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S.*, et al.* (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A *102*, 15545-15550.

Sup, S.J., Alemany, C.A., Pohlman, B., Elson, P., Malhi, S., Thakkar, S., Steinle, R., and Hsi, E.D. (2005). Expression of bcl-2 in classical Hodgkin's lymphoma: an independent predictor of poor outcome. J Clin Oncol *23*, 3773-3779.

Szymanowska, N., Klapper, W., Gesk, S., Kuppers, R., Martin-Subero, J.I., and Siebert, R. (2008). BCL2 and BCL3 are recurrent translocation partners of the IGH locus. Cancer genetics and cytogenetics *186*, 110-114.

Thomas, R.K., Kallenborn, A., Wickenhauser, C., Schultze, J.L., Draube, A., Vockerodt, M., Re, D., Diehl, V., and Wolf, J. (2002). Constitutive expression of c-FLIP in Hodgkin and Reed-Sternberg cells. Am J Pathol *160*, 1521-1528.

Torlakovic, E., Tierens, A., Dang, H.D., and Delabie, J. (2001). The transcription factor PU.1, necessary for B-cell development is expressed in lymphocyte predominance, but not classical Hodgkin's disease. Am J Pathol *159*, 1807-1814.

Tzankov, A., Krugmann, J., Fend, F., Fischhofer, M., Greil, R., and Dirnhofer, S. (2003). Prognostic significance of CD20 expression in classical Hodgkin lymphoma: a clinicopathological study of 119 cases. Clin Cancer Res *9*, 1381-1386.

Ushmorov, A., Leithauser, F., Sakk, O., Weinhausel, A., Popov, S.W., Moller, P., and Wirth, T. (2006). Epigenetic processes play a major role in B-cell-specific gene silencing in classical Hodgkin lymphoma. Blood *107*, 2493-2500.

Ushmorov, A., Ritz, O., Hummel, M., Leithauser, F., Moller, P., Stein, H., and Wirth, T. (2004). Epigenetic silencing of the immunoglobulin heavy-chain gene in classical Hodgkin lymphoma-derived cell lines contributes to the loss of immunoglobulin expression. Blood *104*, 3326-3334.

Van Vlierberghe, P., De Weer, A., Mestdagh, P., Feys, T., De Preter, K., De Paepe, P., Lambein, K., Vandesompele, J., Van Roy, N., Verhasselt, B.*, et al.* (2009). Comparison of miRNA profiles of microdissected Hodgkin/Reed-Sternberg cells and Hodgkin cell lines versus CD77+ B-cells reveals a distinct subset of differentially expressed miRNAs. Br J Haematol *147*, 686-690.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol *3*, RESEARCH0034.
von Ahlfen, S., Missel, A., Bendrat, K., and Schlumpberger, M. (2007). Determinants of RNA quality from FFPE samples. PloS one *2*, e1261.

Wang, T.L., Diaz, L.A., Jr., Romans, K., Bardelli, A., Saha, S., Galizia, G., Choti, M., Donehower, R., Parmigiani, G., Shih Ie, M.*, et al.* (2004). Digital karyotyping identifies thymidylate synthase amplification as a mechanism of resistance to 5-fluorouracil in metastatic colorectal cancer patients. Proc Natl Acad Sci U S A *101*, 3089-3094.

Weiss, L.M., Strickler, J.G., Warnke, R.A., Purtilo, D.T., and Sklar, J. (1987). Epstein-Barr viral DNA in tissues of Hodgkin's disease. Am J Pathol *129*, 86-91.

Wolowiec, D., Berger, F., Ffrench, P., Bryon, P.A., and Ffrench, M. (1999). CDK1 and cyclin A expression is linked to cell proliferation and associated with prognosis in non-Hodgkin's lymphomas. Leuk Lymphoma *35*, 147-157.

Wong, R.K., and Altekruse, S.F. (2009). Where are you and how do we find you? The dilemma of identifying Barrett's epithelium before adenocarcinoma of the esophagus. Am J Gastroenterol *104*, 1363-1365.

Yao, J., Liang, L., Huang, S., Ding, J., Tan, N., Zhao, Y., Yan, M., Ge, C., Zhang, Z., Chen, T.*, et al.* (2010). MicroRNA-30d promotes tumor invasion and metastasis by targeting Galphai2 in hepatocellular carcinoma. Hepatology *51*, 846-856.

Yekta, S., Shih, I.H., and Bartel, D.P. (2004). MicroRNA-directed cleavage of HOXB8 mRNA. Science *304*, 594-596.

Zanotti, R., Trolese, A., Ambrosetti, A., Nadali, G., Visco, C., Ricetti, M.M., Benedetti, F., and Pizzolo, G. (2002). Serum levels of soluble CD30 improve International Prognostic Score in predicting the outcome of advanced Hodgkin's lymphoma. Ann Oncol *13*, 1908-1914.

Zheng, B., Fiumara, P., Li, Y.V., Georgakis, G., Snell, V., Younes, M., Vauthey, J.N., Carbone, A., and Younes, A. (2003). MEK/ERK pathway is aberrantly active in Hodgkin disease: a signaling pathway shared by CD30, CD40, and RANK that regulates cell proliferation and survival. Blood *102*, 1019-1027.

Zhi, F., Chen, X., Wang, S., Xia, X., Shi, Y., Guan, W., Shao, N., Qu, H., Yang, C., Zhang, Y.*, et al.* (2010). The use of hsa-miR-21, hsa-miR-181b and hsa-miR-106a as prognostic indicators of astrocytoma. European journal of cancer *46*, 1640-1649.

Zweig, M.H., and Campbell, G. (1993). Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clin Chem *39*, 561-577.

APPENDIX

PUBLICATIONS

Sánchez-Espiridión B, Sánchez-Aguilera A, Montalbán C, Martin C, Martinez R, González-Carrero J, Poderos C, Bellas C, Fresno MF, Morante C, Mestre MJ, Mendez M, Mazorra F, Conde E, Castaño A, Sánchez-Godoy P, Tomas JF, Morente MM, Piris MA, García JF; Spanish Hodgkin's Lymphoma Study Group.
A TaqMan low-density array to predict outcome in advanced Hodgkin's lymphoma using paraffin-embedded samples.
Clin Cancer Res. 2009 Feb 15;15(4):1367-75. PubMed PMID: 19228737.

Sánchez-Espiridión B, Montalbán C, López A, Menárguez J, Sabín P, Ruiz-Marcellán C, Lopez A, Ramos R, Rodríguez J, Cánovas A, Camarero C, Canales M, Alves J, Arranz R, Acevedo A, Salar A, Serrano S, Bas A, Moraleda JM, Sánchez-Godoy P, Burgos F, Rayón C, Fresno MF, Laraña JG, García-Cosío M, Santonja C, López JL, Llanos M, Mollejo M, González-Carrero J, Marín A, Forteza J, García-Sanz R, Tomás JF, Morente MM, Piris MA, García JF; Spanish Hodgkin Lymphoma Study Group.
A molecular risk score based on 4 functional pathways for advanced classical Hodgkin lymphoma.
Blood. 2010 Aug 26;116(8):e12-7. Epub 2010 May 17. PubMed PMID: 20479282.

Montes-Moreno S, Martinez N, **Sanchez-Espiridión B**, Díaz Uriarte R, Rodriguez ME, Saez A, Montalbán C, Gomez G, Pisano DG, García JF, Conde E, Gonzalez-Barca E, Lopez A, Mollejo M, Grande C, Martinez MA, Dunphy C, Hsi ED, Rocque GB, Chang J, Go RS, Visco C, Xu-Monette Z, Young KH, Piris MA.
MicroRNA expression in diffuse large B-cell lymphoma treated with chemoimmunotherapy.
Blood. 2011 Jul 28;118(4):1034-40. Epub 2011 Jun 1. PubMed PMID: 21633089.

Greaves WO, Kim JE, Singh RR, Drakos E, Kunkalla K, **Sánchez-Espiridión B**, Garcia JF, Medeiros LJ, Vega F.
Glioma-associated oncogene homologue 3, a hedgehog transcription factor, is highly expressed in Hodgkin and Reed-Sternberg cells of classical Hodgkin lymphoma.
Hum Pathol. 2011 Apr 28. [Epub ahead of print] PubMed PMID: 21531006.

Albizua E, Gallardo M, Barrio S, Rapado I, Jimenez A, Ayala R, Rueda D, **Sanchez-Espiridion B,** Puigdecanet E, Espinet B, Florensa L, Besses C, Martinez-Lopez J.
Differential expression of JAK2 and Src kinase genes in response to hydroxyurea treatment in polycythemia vera and essential thrombocythemia.
Ann Hematol. 2011 Aug;90(8):939-46. Epub 2011 Feb 18. PubMed PMID: 21331593.

Martín-Sánchez E, Sánchez-Beato M, Rodríguez ME, **Sánchez-Espiridión B,** Gómez-Abad C, Bischoff JR, Piris MA, García-Orad A, García JF.
HDAC inhibitors induce cell cycle arrest, activate the apoptotic extrinsic pathway and synergize with a novel PIM inhibitor in Hodgkin lymphoma-derived cell lines.
Br J Haematol. 2011 Feb;152(3):352-6. doi: 10.1111/j.1365-2141.2010.08401.x. Epub 2010 Oct 19.PubMed PMID: 20955406.

BOOK CHAPTER

Sánchez-Espiridión, García JF, Sánchez-Beato M.

Chapter Tittle: Advances in classical Hodgkin lymphoma biology: new prognostic factors and outcome prediction using gene expression signatures:

Hodgkin´s Lymphoma , ISBN 979-953-307-065-9

DOI:10.1158/1078-0432.CCR-08-1119

A TaqMan Low-Density Array to Predict Outcome in Advanced Hodgkin's Lymphoma Using Paraffin-Embedded Samples

Beatriz Sánchez-Espiridión,[1] Abel Sánchez-Aguilera,[1] Carlos Montalbán,[5] Carmen Martin,[6] Rafael Martinez,[7] Joaquín González-Carrero,[12] Concepción Poderos,[13] Carmen Bellas,[5] Manuel F. Fresno,[14] Cesar Morante,[15] Maria J. Mestre,[8] Miguel Mendez,[9] Francisco Mazorra,[16] Eulogio Conde,[17] Angel Castaño,[10] Pedro Sánchez-Godoy,[11] José F. Tomas,[4] Manolo M. Morente,[2] Miguel A. Piris,[1] and Juan F. García[1,3] for the Spanish Hodgkin's Lymphoma Study Group

Abstract

Purpose: Despite major advances in the treatment of classic Hodgkin's lymphoma (cHL), ~30% of patients in advanced stages may eventually die as result of the disease, and current methods to predict prognosis are rather unreliable. Thus, the application of robust techniques for the identification of biomarkers associated with treatment response is essential if new predictive tools are to be developed.

Experimental Design: We used gene expression data from advanced cHL patients to identify transcriptional patterns from the tumoral cells and their nonneoplastic microenvironment, associated with lack of maintained treatment response. Gene-Set Enrichment Analysis was used to identify functional pathways associated with unfavorable outcome that were significantly enriched in either the Hodgkin's and Reed-Sternberg cells (regulation of the G_2-M checkpoint, chaperones, histone modification, and signaling pathways) or the reactive cell microenvironment (mainly represented by specific T-cell populations and macrophage activation markers).

Results: To explore the pathways identified previously, we used a series of 52 formalin-fixed paraffin-embedded advanced cHL samples and designed a real-time PCR-based low-density array that included the most relevant genes. A large majority of the samples (82.7%) and all selected genes were analyzed successfully with this approach.

Conclusions: The results of this assay can be combined in a single risk score integrating these biological pathways associated with treatment response and eventually used in a larger series to develop a new molecular outcome predictor for advanced cHL.

Classic Hodgkin's lymphoma (cHL) is considered to be a monoclonal proliferation of the characteristic Hodgkin's and Reed-Sternberg (HRS) cells. It has a defective B-cell immunophenotype and a characteristic paucity of neoplastic cells within the tumor, diluted in a reactive inflammatory background composed of nonneoplastic B and T cells, macrophages, eosinophils, neutrophils, and plasma cells, which comprise the bulk of the infiltrate. The B-cell origin and monoclonality of the HRS cells have been clearly established in the last two decades (1–3). Likewise, progress has been made in recent years to clarify the particular composition of the enigmatic cell microenvironment (4, 5). This is commonly made up of a characteristic TH_2 environment (6) that is involved in the production of survival signals.

Although cHL is usually a curable tumor, ~20% to 30% of patients relapse and eventually die due to progressive disease or complications of therapy (7, 8). Patients with advanced disease and clinical indicators of poor prognosis, and those with disease that persists despite optimized primary treatment, may need intensified treatment (9). In contrast, another fraction of patients could benefit from reduced treatment. Current predictive systems are based on clinical and analytic variables, such as the International Prognostic Score developed for advanced cHL (10), but this still fails to identify a significant fraction of patients with very short failure-free survival (11).

In this context, the application of robust molecular techniques to identify molecular events and biological processes associated with treatment response is a necessary requisite for developing new predictive tools that enhance the

Authors' Affiliations: [1]The Lymphoma Group and [2]Tumour Bank Network, Department of Molecular Pathology, Spanish National Cancer Centre; Departments of [3]Pathology and [4]Hematology, M. D. Anderson International; [5]Department of Internal Medicine and Pathology, Hospital Ramon y Cajal; Departments of [6]Pathology and [7]Hematology, Hospital Clínico Universitario San Carlos; Departments of [8]Pathology and [9]Oncology, Hospital Móstoles; Departments of [10]Pathology and [11]Hematology, Hospital Severo Ochoa de Leganes, Madrid, Spain; Departments of [12]Pathology and [13]Hematology, Complejo Hospitalario Xeral-Cies, Vigo, Spain; Departments of [14]Pathology and [15]Hematology, Hospital Central de Asturias, Oviedo, Spain; and Departments of [16]Pathology and [17]Oncology, Hospital Marqués de Valdecilla, Santander, Spain

Received 5/6/08; revised 8/29/08; accepted 11/7/08.

Grant support: Ministerio de Sanidad y Consumo grants PI051623, PI052800, PI052327, RETICS, Acción Transversal and Ministerio de Ciencia y Tecnología grant SAF2005-00221.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Note: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

Requests for reprints: Miguel A. Piris, Centro Nacional de Investigaciones Oncológicas, Melchor Fernández Almagro 3, Madrid 28029, Spain. Phone: 34-91-224-69-00; Fax: 34-91-224-69-23; E-mail: mapiris@cnio.es.

doi:10.1158/1078-0432.CCR-08-1119

DOI:10.1158/1078-0432.CCR-08-1119

Translational Relevance

Current predictive systems, based on clinical and analytic variables, fail to accurately identify a significant fraction of advanced Hodgkin's lymphoma patients with short failure-free survival. The purpose of this work was the identification of biological processes underlying treatment failure in advanced Hodgkin's lymphoma patients and the subsequent development of a quantitative real-time PCR-based assay to be applied to routine formalin-fixed, paraffin-embedded samples. This study identifies gene subsets expressed by the tumoral cells and the Hodgkin's microenvironment and shows that robust methodologies based on quantitative real-time PCR are suitable for expression profiling of tumors and can be easily applied to paraffin-embedded samples allowing interrogating a limited number of selected genes in a single sample. In addition, we have identified functional signatures associated with treatment response and showed the potential prognostic capacity of our assay finding a positive correlation between the expression of the proposed genes and treatment response.

accuracy of classic clinical variables. Reliable prognostic markers could allow subsets of patients to be identified who might benefit from alternative approaches. Several biological markers identified in tumor tissues, alone (12–15) or in combination (16), have been associated with clinical outcome and treatment response. Not surprisingly, most of these variables reflect functional characteristics of the neoplastic cells in tumor tissues, revealed by proteins with deregulated expression in HRS cells.

The HRS cells and the inflammatory infiltrate secrete cytokines, creating an elaborate cross-talk that contributes to the survival, proliferation, and immune evasion of the tumor cells in many interacting ways (17–20). Recent studies have indicated that the composition of this background is also associated with the clinical outcome of the patients (21–23). Indeed, previous work by our group identified specific gene signatures associated with treatment response that are attributable to the nonneoplastic component of the tumors (24).

In the study reported here, we used gene expression data from the analysis of advanced cHL samples and Hodgkin's lymphoma-derived cell lines (24) to identify specific markers from the tumoral cells and their nonneoplastic microenvironment that were associated with lack of maintained treatment response. Gene-Set Enrichment Analysis (GSEA) was used to recognize specific functional pathways related to outcome that were significantly enriched in both tumor and microenvironment components. We then assigned the signatures generated from both tumoral and stroma cells to specific pathways in an attempt to ascribe a biological significance to these findings. To investigate these cell subpopulations and chosen pathways, we selected 64 genes that could be assayed using quantitative real-time PCR [RT-PCR; TaqMan low-density array (TLDA) assays] on formalin-fixed, paraffin-embedded (FFPE) tissues.

Materials and Methods

Samples. The initial gene data set has been generated previously (24) by gene expression profiling of tumor samples from 29 patients with advanced cHL, including 14 responders and 15 nonresponders to doxorubicin, bleomycin, vinblastine, and dacarbazine treatment and 5 Hodgkin's lymphoma-derived cell lines (L428, L540, L1236, HDLM2, and HDMYZ; Fig. 1). This data set was used here to define Tumor Database (TDB) and Microenvironment Database (MDB) and subsequently in GSEA analysis.

To validate specific genes to be analyzed in FFPE samples, we selected an independent series of 52 cHL cases that did not overlap with the initial series used for expression profiling. Patients were eligible if they fulfilled the same stringent criteria as described previously: ages between 18 and 65 years, advanced cHL (Ann Arbor stage IV, III, or IIB with bulky masses), proven HIV-negative status, and receiving a standard chemotherapy regimen that included doxorubicin (doxorubicin, bleomycin, vinblastine, and dacarbazine or variants). All tissue samples consisted of representative specimens of pretreatment lymph node biopsies and were collected after revision by the institutional review board.

The endpoint of this study was maintained response to therapy. Good response (favorable outcome) was considered to be that where patients sustained complete response (18 months), whereas a bad response (unfavorable outcome) was concluded for cases without complete response or in patients with early relapses following previously published criteria (25). Because the main aim of this study was to analyze the biological variables associated with response to the initial therapy, data from second-line and salvage therapies and/or bone marrow transplantation were not considered.

Supervised analysis and identification of tumoral and microenvironment signatures. The analysis of the initial data set (24) was done using the GEPAS (26) tool, which is available free at http://www.gepas.org. To identify genes that are differentially expressed between tumoral samples and cHL cell lines, we used a supervised method based on Student's t test with a correction for multiple testing. Unadjusted P values were obtained from 100,000 permutations of the data set and false discovery rates were calculated by the method of Benjamini and Hochberg (27). Genes with false discovery rate < 0.15 were considered to be differentially expressed and used to construct the TDB and MDB employed in subsequent analyses.

Gene expression profiles of cHL-derived cell lines and tumor tissues were compared to facilitate the separation and recognition of the signatures attributable to both the neoplastic cells and the microenvironment, thus making possible the building of two databases (TDB and MDB) for further analysis and selection of the genes analyzed in the TLDA assay. We first compared gene expression profiles from the tumoral samples and from cHL-derived cell lines. Genes overexpressed in the cell lines were included in the TDB, whereas those genes overexpressed in the tumor samples were assigned to the MDB. To avoid loss of information, genes that were not differentially expressed were included in both databases used in further analyses. This first analysis was done without consideration of clinical outcomes of the patients.

Immunohistochemical validation of selected genes in TDB and MDB. Additionally, we selected a limited number of genes for validation at the protein level using tissue microarray-based immunohistochemical techniques on an independent series of 142 cHL cases as described previously (24). Primary chosen antibodies were as follows: anti-BCL2 and anti-LYZ (DAKO); anti-BCLXL (Zymed Laboratories); anti-CASP3 and anti-CCNA (Novocastra Laboratories); anti-CTSL (Alexis Biochemicals); anti-STAT1, anti-SH2D1A, anti-CDK7, anti-HSP70, and anti-MUM1 (IRF4; Santa Cruz Biotechnology); anti-CCNH (Cell Signaling Technology); and anti-HIST1H and anti-HISTH2A (Upstate Biotechnology). These markers were chosen as representative of the different signatures and for the availability of reliable antibodies for paraffin-embedded tissue.

DOI:10.1158/1078-0432.CCR-08-1119

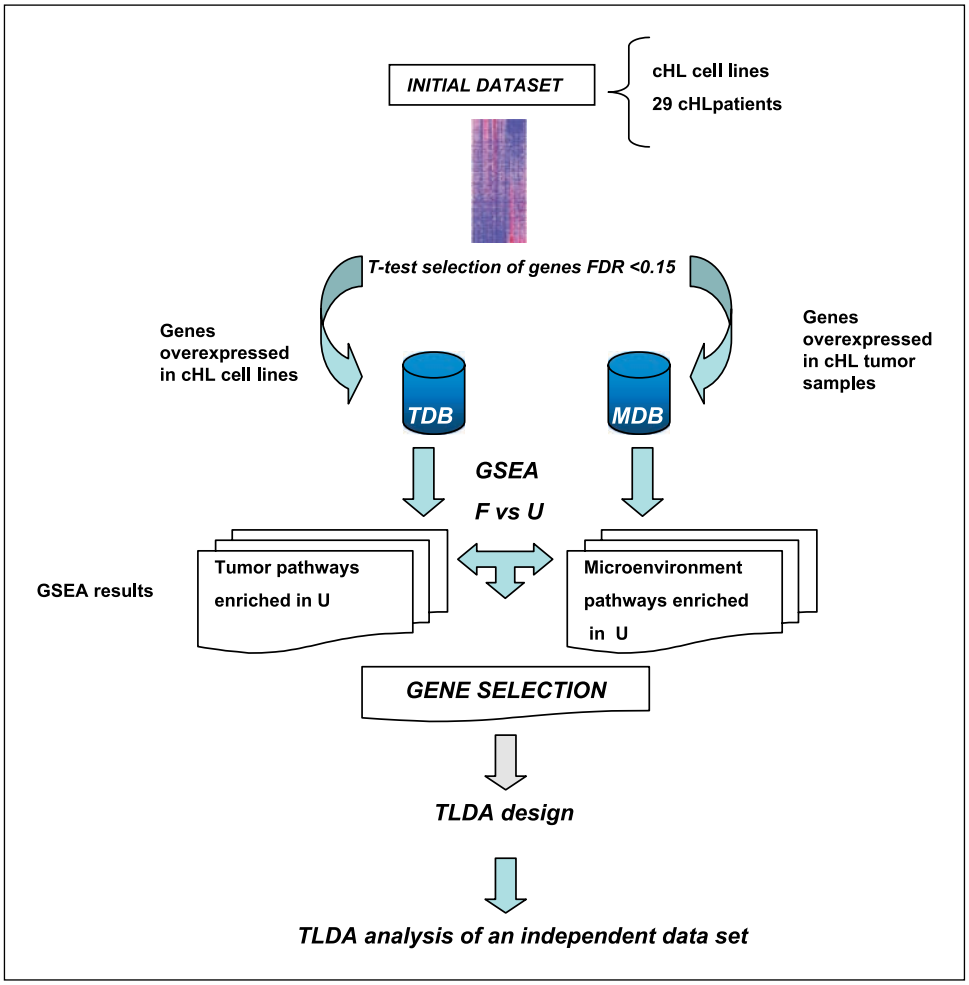


Fig. 1. Workflow analysis and enrichment plots for gene-sets identified by GSEA in unfavorable outcome group used to design TLDA assay. After characterization of the tumor and gene databases, GSEA allowed us to select the genes included in the TLDA assay. *F,* favorable outcome; *U,* unfavorable outcome.

GSEA and pathway selection. GSEA was done to identify sets of related genes that might be correlated with therapy resistance. The TDB and MDB were tested separately. Briefly, this method uses *t* statistics to search for predefined lists of gene-sets that are associated with phenotypic differences. The pathways were selected taking into consideration the particular cellular heterogeneity of cHL tumors, thus including all public pathways associated with immune response and cell-cell interactions. We used Biocarta[18] and other public sources available through the Molecular Signature Database[19] to generate the gene-set databases used for the TDB and MDB analysis. The final gene-set databases were manually curated and enriched in pathways also known to be involved in lymphoma pathogenesis and immune response.

The rank of all genes in the sets was determined and an enrichment score was calculated as described previously (28). We performed the analysis with 1,000 random class permutations. Genes with >30% of missing values were excluded, and only gene-sets meeting the gene-set size criteria (min = 10, max = 500 genes) were analyzed.

Representative genes were chosen from the most highly enriched pathways in the TDB and MDB. They were selected based on the strength of their ability to discriminate patients with a good or bad treatment response (enrichment score of the genes in each pathway) and the biological relevance of the pathway. To ensure high quality and fidelity of TLDA assays, we also considered amplicon length as a selection criterion, discarding those genes with longer amplicon length.

[18] www.biocarta.com

[19] http://www.broad.mit.edu/gsea/msigdb/index.jsp

RNA extraction and cDNA synthesis. Total mRNA was extracted from three 10 μm FFPE sections as described previously (29), with minor modifications. Briefly, paraffin sections were deparaffinized by incubation with 1 mL xylene for 10 min. Samples were then centrifuged and the supernatant was removed. After an ethanol wash, tissue pellets were dried, resuspended in 200 μL RNA lysis buffer containing 50 mmol/L Tris, 0.5 mmol/L EDTA (pH 8.0), and 10% SDS, and incubated overnight at 65°C with 10 μL proteinase K (20 mg/mL; Qiagen). RNA was purified by phenol-chloroform extraction followed by precipitation in an equal volume of isopropanol in the presence of 1 μL lineal acrylamide (Ambion) at -20°C. The RNA pellet was washed once in 70% ethanol, dried, and resuspended in 30 μL Tris/EDTA (pH 8). For genomic DNA removal, DNase digestion was carried out by treating the total RNA with 5 μL DNase I (1 units/μL; Epicentre Biotechnologies). The final RNA concentration (A260: A280) and purity (A260:A230 ratio) was measured using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies).

First-strand cDNA was synthesized from total RNA using the High Capacity cDNA Archive kit (Applied Biosystems) in 50 μL reactions using these cDNA samples for TLDA analysis according to the manufacturer's instructions.

Limited amounts of tissue and RNA and cDNA were available from FFPE samples, which restricted the number of genes and cases that could be analyzed in some samples, so we were able to test the application of the TaqMan PreAmp Master Mix (T-PreAmp; Applied Biosystems) in 13 cases and 32 genes. The pooled assays were diluted to a final concentration of one-fifth that of the PreAmp primer/assay pool. Initial experiments comparing the volume of PreAmp reaction recommended by the manufacturer (total PreAmp reaction volume:

DOI:10.1158/1078-0432.CCR-08-1119

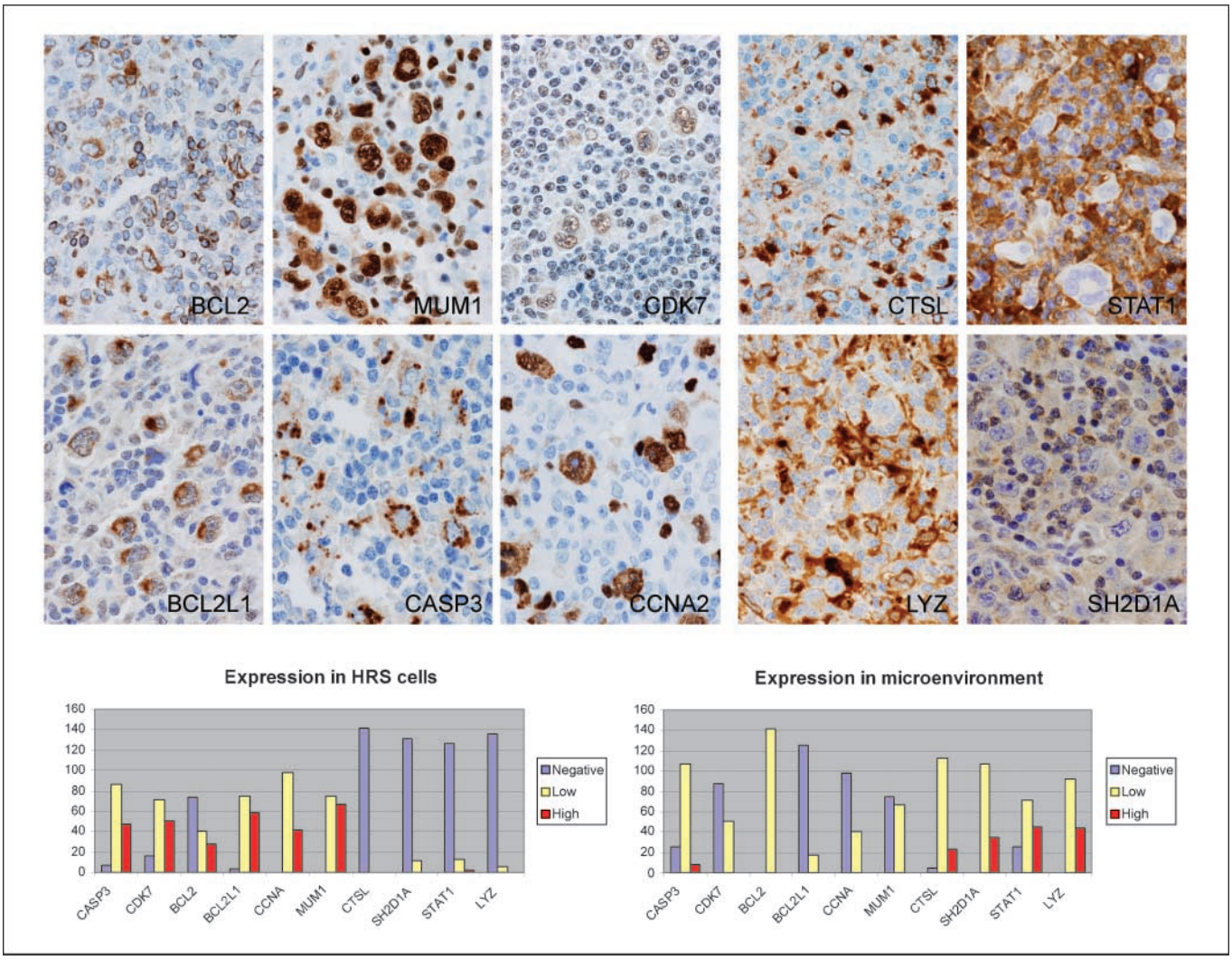


Fig. 2. Immunohistochemical analyses of selected markers. Proteins identified in the six panels on the *left* correspond to genes included in the TDB, and their expression is mainly restricted to tumoral HRS cells (BCL2, MUM1, CDK7, BCL2L1, CASP3, and CCNA2). The four panels on the *right* correspond to genes included in the MDB, and the expression of the respective proteins is mainly restricted to fibroblasts, macrophages, and reactive T cells (CTSL, STAT1, LYZ, and SH2D1A), whereas tumor cells are negative. The histograms represents the number of postive cHL samples for each marker expressed by the HRS cells of the microenvironment.

50 μL) and a tenth of that volume (5 μL PreAmp reaction volume) gave identical PreAmp results, so we subsequently used the latter volume for reasons of cost-effectiveness. The chosen PreAmp reaction involved 14 cycles of preamplification with 15 s at 95°C and 4 min at 60°C. The preamplified products were diluted at a ratio 1:10 and used as templates for RT-PCR analysis.

TLDA assays. To explore the different pathways, we designed a quantitative RT-PCR assay based on TLDA technology (Applied Biosystems) to measure the expression of each selected gene in triplicate.

Each cDNA sample (30 μL) was added to 20 μL RNase-free water and 50 μL of 2× TaqMan Universal PCR Master Mix (No AmpErase UNG; Applied Biosystems). The mixture was then transferred into a loading port on a TLDA card. The card was centrifuged twice and sealed and PCR amplification was done using Applied Biosystems Prism 7900HT Sequence Detection System under the following thermal cycler conditions: 2 min at 50°C and 10 min at 94.5°C for 40 cycles (30 s at 97°C and 1 min at 59.7°C).

Reference gene selection. The TLDAs included HMBS, GUSB, TBP, and 18S as reference genes based on their proven role as housekeeping genes (30, 31) and their uniform expression in preliminary TLDA assays in FFPE tumor samples from this series of cHL (data not shown).

To determine the stability of the selected reference genes, the geNORM Visual Basic application available in RealTime StatMiner (Integromics), also known as the pairwise approach, was used as described previously by Vandesompele et al. (32). This tool exploits the principle that the expression ratio of two perfect reference genes should be identical in all samples regardless of the experimental condition or cell type. The program calculates the gene stability measure M by determining the average pairwise variation between a particular reference gene and all other control genes. Genes with higher values of M have greater variation in RNA expression. By stepwise exclusion of the least stable gene and recalculation of the M values, the most stable reference genes are identified. Finally, a normalization factor was calculated based on the geometric mean of the expression levels of the best-performing reference genes. The two best candidates, HMBS and GUSB, were chosen for normalization of gene expression levels.

Statistical analysis and prediction algorithm. Association between standard clinical variables (age, gender, stage, and International Prognostic Score) were assessed by the Pearson χ^2 test.

DOI:10.1158/1078-0432.CCR-08-1119

To illustrate the relationship between gene expression data and treatment response, we derived an integrated risk score, which was defined as the logarithmic mean of the expression levels of all the genes included in the analysis.

Associations between gene expression levels and the probability of treatment response were estimated by logistic regression analysis. Normalized expression levels were used as independent continuous variables and phenotype-favorable (responders) versus phenotype-unfavorable (nonresponders) as the dependent variable. Receiver operating characteristic curves were used to compare the predictive sensitivity and specificity of individual genes (33) and to select the best predictive markers.

All statistical analyses were two-sided and done using SPSS 13.0 (SPSS).

Results

Clinical variables of selected samples. Differences in the distribution of classic clinical variables (age, gender, and stage) between the two groups of patients (responders and non-responders) were not statistically significant (χ^2 test). Interestingly, the International Prognostic Score did not differ significantly between the two groups either because the sample was too small to conclude statistical significance for the magnitude of the difference measured or because the analysis was restricted to advanced Hodgkin's lymphoma cases (data not shown).

Supervised analysis and identification of tumor and microenvironment signatures. Established cHL-derived cell lines are commonly used as model systems for characterizing the biology of HRS cells, whereas tumoral samples represent a complex mixture of tumoral cells and reactive microenvironment (34). To identify genes expressed by either HRS cells or the microenvironment, we first compared gene expression profiles from the tumoral samples and gene profiles from cHL-derived cell lines. We identified 3,463 of the 5,134 genes that were differentially expressed (false discovery rate ≤ 0.15). Those genes overexpressed in the cell lines were included in the TDB (1,705 genes), whereas those overexpressed in the tumor samples were assigned to the MDB (1,758 genes). Genes that were not differentially expressed (false discovery rate > 0.15; 1,671 genes) were considered as potentially being expressed by both populations and were included in both databases.

The gene composition of the two databases confirms this hypothesis: the TDB contains genes known to be expressed by HRS cells. These include cell cycle regulators, signaling, surface receptors, and transcription factors. In addition, genes reported previously as being expressed by HRS cells in studies done by other groups (35), such as TNFRSF8 (CD30), GATA3, the tumor necrosis factor receptor family member RANK, and the metalloproteinase TIMP1 (36, 37), were incorporated within the TDB. The MDB is mainly composed of genes involved in the immune response. We performed an additional validation by showing that genes attributed to the TDB (BCL2, BCLXL, CASP3, CDK7, MUM1, and CCNA) were indeed expressed by the tumoral cells, whereas genes attributed to the microenvironment (STAT1, LYZ, SH2D1A, and CTSL) were expressed by macrophages, reactive T cells, and fibroblasts (Fig. 2). This confirms and extends previous observations on the expression of ALDH1A1, RRM2, CDC2, MAD2L1, TOP2A, and PCNA (35).

These gene signatures were used to generate two different data sets with the gene expression data of the previously

Table 1. GSEA-enriched pathways associated with unfavorable outcome

	Normalized enrichment score	Nominal probability	Genes included in TLDA assays		
Gene-sets enriched in TDB					
G_2-M pathway	-1.67	0.016	*AURKA*	*CENPE*	*MAD2L1*
			BUB1B	*CENPF*	*MAPRE1*
			BUB3	*CHEK1*	*NBS1*
			CCNH	*CSE1L*	*NUMA1*
			CDC2	*HMMR*	*RSN*
G_s pathway	-1.50	0.052	*RAMP*	*CCNH*	*CDK7*
			CCNA2	*CDC6*	
			CCNE2	*CDC2*	
G_1 pathway	-1.35	0.145	*BCCIP*	*CCNH*	*CDKN2C*
Histone pathway	-1.26	0.151	*H1F0*	*H2AFX*	*HIST1H3D*
Chaperone pathway	-1.24	0.269	*DNAJA2*	*HSPA4*	*HSPA9B*
			HSP90AA1		
Drug resistance pathway	-1.23	0.204	*DCK*	*RRM2*	*TYMS*
			MLH1	*TOP2A*	
Mitogen-activated protein kinase pathway	-0.83	0.736	*GRB2*	*MAPK14*	*MAPK9*
			MAP3K7	MAPK6	
Mitochondrial pathway	-0.77	0.883	*BCL2*	*BCL2L1*	*CASP3*
			CYCS		
Gene-sets enriched in MDB					
T-cell pathway	-1.50	0.063	*CD3D*	*CD8B1*	*CTSL*
			SH2D1A	*IFI16*	
Monocyte/macrophage	-1.39	0.161	*ALDH1A1*	*ITGA4*	*LCP1*
			LYZ	*STAT1*	

NOTE: The most significantly enriched pathways identified by GSEA in TDB and MDB associated with the unfavorable outcome group and the highest-scoring genes from each pathway selected for investigation in the TLDA assay.

DOI:10.1158/1078-0432.CCR-08-1119

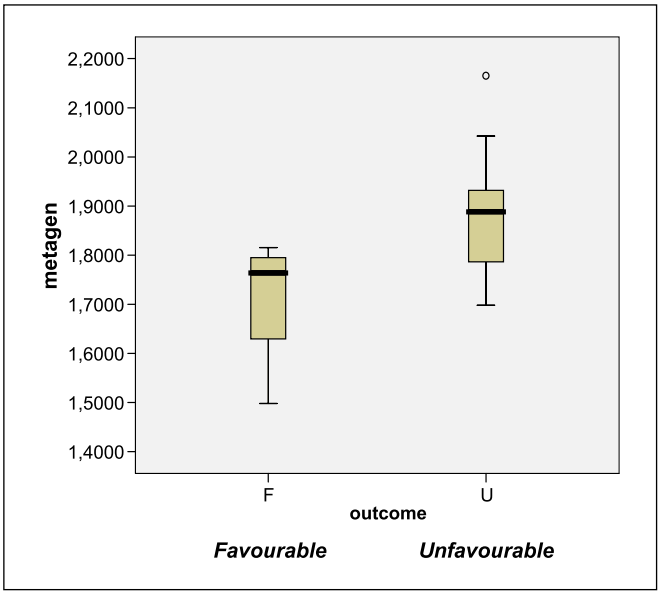


Fig. 3. Box-plots of the integrated risk score, defined as the logarithmic mean of expression levels of all genes included in theTLDA assays (metagen), in each patient outcome group. *Black line within the box,* median of each group. Box group values are within the first and third quartiles. *Whiskers* (error bars) extend from each end of the box to the adjacent values in the data and represent the most extreme values within 1.5 times the interquartile range from the ends of the box. *Symbols,* outliers.

described 29 cHL patients containing the genes characteristic of the reactive microenvironment and the HRS cells, respectively.

GSEA. We hypothesized that response to treatment in cHL might be associated with concurrent alterations in biological pathways or coregulated gene-sets rather than with randomly identified individual genes. However, comparisons made on the level of gene lists indicated considerable divergence when different statistical methods or data sets were used. Thus, we applied GSEA to the two previously generated databases (TDB and MDB) taking into consideration that group testing techniques and pathways recurrently identified by this method are more likely to be reliable than those identified by other approaches (38).

We used independent gene-sets for each database (the complete list of gene-sets used in the analysis is presented in Supplementary Table S1). Taking into consideration the low number of samples and their uneven phenotypic distribution, we used GSEA for testing hypotheses and designing the TLDA assay rather than for final statistical analysis.

GSEA analysis highlighted 15 of the 46 functional pathways associated with unfavorable outcome in TDB and 13 of 48 in MDB. The TDB analysis identified two gene-sets (G_2-M and G_1-S) that were significantly enriched at nominal probabilities of <5% in HRS cell signature-containing genes such as AURKA, MAD2L1, BUB1B, BUB3, CHEK1, and CDK1 (CDC2) associated with the regulation of the spindle checkpoint, in agreement with previous results (24). Also, G_1 pathway (CCNH, CDKN2C, and BCCIP), histone (H1F0, H2AFX, and HIST1H3D), chaperones (DNAJA2, HSPA4, HSPA9B, and HSP90AA1), and other signaling pathways such as mitogen-activated protein kinase pathway and mitochondrial signaling were identified as being enriched in the unfavorable outcome group by GSEA analysis (Table 1; Fig. 2). Additionally, in TDB, the drug resistance pathway was enriched in the unfavorable outcome group of patients. This pathway contains genes such as TOP2A and MLH1, which are known to be involved in transport and metabolism of doxorubicin.

On the other hand, in the analysis of the MDB, T cells, monocyte/macrophages, and dendritic cell pathways were the most enriched gene-sets among those associated with unfavorable outcomes. These included genes reporting from specific T-cell populations, highlighting the important role of the microenvironment.

TLDA analysis and relationship with treatment response. The results of GSEA analysis prompted us to design a TLDA assay with the highest-scoring genes from the previously identified pathways reporting on functions known to be altered in cHL disease that might be related to outcome and treatment response.

The selected genes for the TLDA assays are listed in Table 1. We selected 56 representative genes from the most enriched pathways in the TDB and MDB based on the strength of their performance in discriminating different outcomes and the biological relevance of the pathway. Three additional genes, BCL2, IRF4, and FOXP3, were included based on their published prognostic significance in cHL (12, 23, 39). TNFRSF8 (CD30) was also included because of its common expression in HRS cells (the complete list of genes is available in Supplementary Table S1).

Fifty-two FFPE tumor samples were analyzed and adequate RT-PCR profiles (Ct < 35 for endogenous genes; ref. 40) were obtained for 82.7% of cases (43 of 52) and normalized using the two most stable endogenous genes (HMBS and GUSB) identified by geNORM analysis.

Normalized expression levels (ΔCt) of the individual genes varied considerably among samples. However, consistent with previous results (24), most of these genes were, on average, overexpressed in the unfavorable group. As shown in Fig. 3, the integrated risk score was significantly higher in the bad outcome group ($P = 0.025$), although with some overlap in the values, probably due to the limited number of cases. This result confirms that the functional pathways selected by GSEA are also associated with treatment response in an independent series of cHL patients.

The area under the receiver operating characteristic curves, with a cutoff of 0.65, which corresponds to P values < 0.05, was used to identify the best predictive genes, of which there were 14: BCCIP, CASP3, CCNE2, CSEL1, CTSL, CYCS, DCK, DNAJA2, HSP90AA1, HSPA4, ITGA4, LYZ, RSN, and TYMS (Table 2). With this set of genes, the logistic regression model had an overall 100% predictive accuracy for the whole series ($\chi^2 = 51.049$; $P < 0.001$), indicating that individual genes could be included in a single model of outcome prediction. Due to the small number of cases analyzed, leave-one-out cross-validation gave a value of 69.5% accurate classification. Thus, these initial results need to be confirmed in a larger series to develop a predictive model of sufficient accuracy that can be of general application.

Reanalysis of the cases with inadequate RT-PCR data. Although RT-PCR is a sensitive, precise, and reproducible tool for determining gene expression in tissues, the quality and amount of RNA that can be extracted from FFPE tissues are low, which

DOI:10.1158/1078-0432.CCR-08-1119

limits the number cases and target genes that can be analyzed. PreAmplification techniques (PreAmp) could enhance the sensitivity and quality of the LDA assay. Thus, we also tested the application of the TaqMan PreAmp Master Mix technique in a subset of cases and genes in the study and were thereby able to recover some missing data and improve the quality of the assay.

The study of a few cases yielded RT-PCR results considered of low quality, with Ct < 35. With the aim of increasing the proportion of analyzable cases in further studies, we tested a preamplification step in a smaller set of 13 cases, including 2 previously considered to be low quality. This allowed us to analyze all the genes and cases and yielded a mean improvement of 4.49 cycles (mean Ct value = 25.62 ± 4.76 using PreAmp versus 30.11 ± 3.40 without PreAmp), which represents more reliable and better quality data.

Discussion

The increasing knowledge of Hodgkin's lymphoma pathogenesis and the complex relationship existing between the neoplastic HRS cells and their microenvironment have not yet been translated into the development of new predictive tools that could accurately identify Hodgkin's lymphoma patients with different risks of treatment response and failure.

Transcriptional analysis of cancer is proving to be a powerful and increasingly useful tool in biomedical research. A goal of these studies is to use gene expression patterns revealed by transcriptional profiling to understand the pathogenesis of the disease and to predict prognosis and responsiveness to therapy. Indeed, previous gene expression analysis studies on Hodgkin's lymphoma samples (41) suggested that these techniques could potentially be applied for outcome prediction. However, the practical applications of these advances in the routine clinical care of the patients are still rare mainly due to the limitations arising from the relatively few cases that can be studied using frozen specimens. Thus, it is essential to simplify experimental methods, use them for paraffin-embedded samples, and reduce the gene composition of the predictive signatures to develop new predictive models of potential clinical application.

In this study, we have shown that robust methodologies, based on quantitative RT-PCR, are suitable for expression profiling of tumors and can be applied to routine FFPE tissue samples and that new platforms such as TLDAs allow the robust analysis of a limited number of genes. Hundreds of reactions can be done simultaneously, enabling large numbers of cases to be rapidly assessed. These technologies would be especially useful for identifying gene signatures of prognostic significance (42) and have already been shown to be useful in other tumor types (43). Thus, we have built a novel TLDA assay that includes genes known to be involved in Hodgkin's lymphoma pathogenesis and that could be related to treatment response. To select the genes for inclusion in our assay, we applied GSEA to the data from our previous study using gene expression profiling of cHL tumor samples to identify a reduced number of genes.

The amount of tissue and RNA and cDNA available from FFPE samples are limited, which restricts the number of analyzable genes and cases, so we tested an alternative approach that could be used in further studies. Preamplification techniques (PreAmp) enhance sensitivity of RT-PCR especially for low-abundance target genes and low RNA quality, which increases the number of cases and target genes that can be analyzed. PreAmp maintenance of gene expression profiles has been already validated (44) in previous studies.

With the present approach, we first identify genes putatively expressed by the neoplastic HRS cells and the microenvironment by comparing gene signatures from cHL-derived cell lines and profiles from tumor samples. A similar bioinformatic approach has recently been adopted with this tumor type (34), and the gene composition of the two databases confirms that it yields informative and consistent results.

GSEA identifies functional pathways overrepresented in specific phenotypes, thus permitting a reduction in the number of redundant genes and avoiding the loss of information from important biological processes. The final set of genes includes a balanced representation of those associated with the tumor cells and the microenvironment and that are intrinsically associated with the pathogenetic characteristics of cHL.

Table 2. Summary of the best discriminant genes (14 genes with area under receiver operating characteristic curve > 0.65) used for the logistic regression analysis

Gene symbol	Gene name	Area under receiver operating characteristic curve	*P*
BCCIP	*BRCA2 and CDKN1A interacting protein*	0.756	0.008
CASP3	*Caspase-3, apoptosis-related cysteine peptidase*	0.765	0.006
CCNE2	*Cyclin E2*	0.656	0.051
CSEL1	*CSE1 chromosome segregation 1-like (yeast)*	0.657	0.050
CTSL	*Cathepsin L*	0.678	0.046
CYCS	*Cytochrome **c**, somatic*	0.753	0.009
DCK	*Deoxycytidine kinase*	0.747	0.010
DNAJA2	*DNAJ (Hsp40) homologue, subfamily A, member 2*	0.671	0.049
HSP90AA1	*Heat shock protein 90 kDa α (cytosolic), class A member 1*	0.662	0.050
HSPA4	*Heat shock 70 kDa protein 4*	0.659	0.050
ITGA4	*Integrin, α_4 (antigen CD49D, α_4 subunit of VLA-4 receptor)*	0.651	0.055
LYZ	*Lysozyme (renal amyloidosis)*	0.674	0.047
RSN	*Restin (Reed-Sternberg cell-expressed intermediate filament-associated protein)*	0.706	0.033
TYMS	*Thymidylate synthetase*	0.750	0.010

DOI:10.1158/1078-0432.CCR-08-1119

Selected genes of the tumor signature highlight significant biological functions that are known to be involved in Hodgkin's lymphoma pathogenesis (24) and, more specifically, cell proliferation, consistent with the increased proliferation shown by HRS cells (45). The G_2-M pathway includes several important genes associated with the regulation of the spindle checkpoint, such as AURKA, MAD2L1, BUB1B, BUB3, CHEK1, and CDK1 (CDC2). Most of these genes are known to be involved in chemoresistance (46, 47) and could be therapeutic targets. RSN is another interesting marker of HRS cells in cHL (48), which is able to multimerize tubulin acting as polymerization chaperones and is probably also associated with the regulation of the spindle checkpoint (49).

There is also a representation of cell cycle regulators specifically related with drug metabolism, including topoisomerase 2α (cellular target of Adriamycin), TYMS, and RRM2, which have been associated with drug resistance in different tumor models (50, 51).

Another interesting observation is the presence of the chaperone pathway in the selected genes. Among the genes included in the TLDA assay with a significantly higher risk score in the bad treatment response group, we found histones (HSP90AA1 and HSPA4), which are frequently overexpressed in cancer cells and may play a role in malignant transformations (52). HSP90A acting as a chaperone regulates proteins that promote HRS survival, such as AKT, MEK, and components of the nuclear factor-κB pathway (53, 54). It is of note that there are small molecules such as 17-allylamino-17-demethoxygeldamycin that inhibit HSP90 function (55) and could offer an alternative approach to the treatment of cHL.

On the other hand, the microenvironment signature is overrepresented by genes involved in the Th2 immune response, which has recently been described as being characteristic of Hodgkin's lymphoma tumors and associated with tumor survival (20). Other interesting T-cell populations present in the selected genes are cytotoxic and regulatory T cells, also associated with outcome in cHL (23). A macrophage pathway is also included, consistent with the association between macrophage activation and Th2-type immune responses. Recent observations also indicate a relationship between tumor-infiltrating macrophages and length of survival among patients with follicular lymphoma (56, 57). The inclusion of these pathways in our gene selection is consistent with the idea that treatment response and survival in cHL (and other lymphoma types) correlate with the molecular features of nonmalignant immune cells present in the tumor at diagnosis.

Although the main aim of this study was the development of an exploratory tool for this tumor model, the results also confirm its potential prognostic capacity, because higher-risk scores were obtained in patients with poor treatment response.

We expect these results to pave the way for a more comprehensive analysis of a larger series that could allow a predictive model to be developed in which different genes and pathways could be integrated with specific scores.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Appendix A

The following centers and investigators participate and contribute with the Spanish Hodgkin's Lymphoma Study Group: P. Domínguez and C. Jara (Fundación Hospital Alcorcón); M.J. Mestre, R. Quibén, M. Méndez, and L. Borbolla (Hospital de Móstoles); M.A. Martínez and C. Grande (Hospital 12 de Octubre); M. García-Cosío, C. Montalbán, and J. García-Laraña (Hospital Ramón y Cajal); C. Bellas and M. Provencio (Hospital Puerta de Hierro); A. Castaño and P. Sánchez-Godoy (Hospital Severo Ochoa); C. Martín and R. Martínez (Hospital Clinico Universitario San Carlos); J. Menárguez, P. Sabín, and E. Flores (Hospital Gregorio Marañón); J. González-Carrero and C. Ponderós (Hospital Xeral-Cies); T. Álvaro and L. Font (Hospital Verge de la Cinta); V. Romagosa and A. Fernández de Sevilla (Institut Catala d'Oncologia); M. Mollejo and M.A. Cruz (Hospital Virgen de la Salud); H. Álvarez-Arguelles and M. Llanos (Hospital Universitario Canarias); C. Morante (Hospital Cabueñes); F. Mazorra and E. Conde (Hospital Marqués de Valdecilla); M.F. Fresno, C. Rayón, and C. Nicolás (Hospital Central de Asturias); T. Flores and R. Garca-Sanz (Hospital Universitario de Salamanca); J. Guma (Hospital Sant Joan); P. Gonzalvo (Hospital Comarcal de Jarrio); G. Fernández (Hospital Alvarez Buyllas); J. Forteza, M. Fraga, and J.L. Bello (F Med Santiago de Compostela); J.R. Méndez (Hospital Valle de Nalón); and J.F. García, M.M. Morente, and M.A. Piris (Spanish National Cancer Centre).

Acknowledgments

We thank Gonzalo Gómez for help with bioinformatics tools, the Tumour Bank Network of the Spanish National Cancer Centre for providing all frozen and paraffin-embedded tissue used in this study, and all the hospitals, pathologists, and clinicians that have collaborated in this work.

References

1. Marafioti T, Hummel M, Foss HD, et al. Hodgkin and Reed-Sternberg cells represent an expansion of a single clone originating from a germinal center B-cell with functional immunoglobulin gene rearrangements but defective immunoglobulin transcription. Blood 2000;95:1443–50.
2. Kuppers R, Rajewsky K, Zhao M, et al. Hodgkin disease: Hodgkin and Reed-Sternberg cells picked from histological sections show clonal immunoglobulin gene rearrangements and appear to be derived from B cells at various stages of development. Proc Natl Acad Sci U S A 1994;91:10962–6.
3. Kanzler H, Hansmann ML, Kapp U, et al. Molecular single cell analysis demonstrates the derivation of a peripheral blood-derived cell line (L1236) from the Hodgkin/Reed-Sternberg cells of a Hodgkin's lymphoma patient. Blood 1996;87:3429–36.
4. Skinnider BF, Mak TW. The role of cytokines in classical Hodgkin lymphoma. Blood 2002;99:4283–97.
5. Maggio E, van den Berg A, Diepstra A, Kluiver J, Visser L, Poppema S. Chemokines, cytokines and their receptors in Hodgkin's lymphoma cell lines and tissues. Ann Oncol 2002;13 Suppl 1:52–6.
6. Van den Berg A, Visser L, Poppema S. High expression of the CC chemokine TARC in Reed-Sternberg cells. A possible explanation for the characteristic T-cell infiltrate in Hodgkin's lymphoma. Am J Pathol 1999;154:1685–91.
7. Canellos GP, Niedzwiecki D. Long-term follow-up of Hodgkin's disease trial. N Engl J Med 2002;346:1417–8.
8. Bonadonna G, Viviani S, Bonfante V, Gianni AM, Valagussa P. Survival in Hodgkin's disease patients—report of 25 years of experience at the Milan Cancer Institute. Eur J Cancer 2005;41:998–1006.
9. Connors JM. State-of-the-art therapeutics: Hodgkin's lymphoma. J Clin Oncol 2005;23:6400–8.
10. Hasenclever D, Diehl V. A prognostic score for ad-

DOI:10.1158/1078-0432.CCR-08-1119

vanced Hodgkin's disease. International Prognostic Factors Project on advanced Hodgkin's disease. N Engl J Med 1998;339:1506–14.

11. Gobbi PG, Zinzani PL, Broglia C, et al. Comparison of prognostic models in patients with advanced Hodgkin disease. Promising results from integration of the best three systems. Cancer 2001;91:1467–78.

12. Rassidakis GZ, Medeiros LJ, Vassilakopoulos TP, et al. BCL-2 expression in Hodgkin and Reed-Sternberg cells of classical Hodgkin disease predicts a poorer prognosis in patients treated with ABVD or equivalent regimens. Blood 2002;100:3935–41.

13. Rassidakis GZ, Medeiros LJ, McDonnell TJ, et al. BAX expression in Hodgkin and Reed-Sternberg cells of Hodgkin's disease: correlation with clinical outcome. Clin Cancer Res 2002;8:488–93.

14. Garcia JF, Camacho FI, Morente M, et al. Hodgkin and Reed-Sternberg cells harbor alterations in the major tumor suppressor pathways and cell-cycle checkpoints: analyses using tissue microarrays. Blood 2003;101:681–9.

15. Herling M, Rassidakis GZ, Vassilakopoulos TP, Medeiros LJ, Sarris AH. Impact of LMP-1 expression on clinical outcome in age-defined subgroups of patients with classical Hodgkin lymphoma. Blood 2006;107:1240; author reply 1241.

16. Montalban C, Garcia JF, Abraira V, et al. Influence of biologic markers on the outcome of Hodgkin's lymphoma: a study by the Spanish Hodgkin's Lymphoma Study Group. J Clin Oncol 2004;22:1664–73.

17. Juszczynski P, Ouyang J, Monti S, et al. The AP1-dependent secretion of galectin-1 by Reed Sternberg cells fosters immune privilege in classical Hodgkin lymphoma. Proc Natl Acad Sci U S A 2007;104:13134–9.

18. Ishida T, Ishii T, Inagaki A, et al. Specific recruitment of CC chemokine receptor 4-positive regulatory T cells in Hodgkin lymphoma fosters immune privilege. Cancer Res 2006;66:5716–22.

19. Marshall NA, Christie LE, Munro LR, et al. Immunosuppressive regulatory T cells are abundant in the reactive lymphocytes of Hodgkin lymphoma. Blood 2004;103:1755–62.

20. Re D, Kuppers R, Diehl V. Molecular pathogenesis of Hodgkin's lymphoma. J Clin Oncol 2005;23:6379–86.

21. Oudejans JJ, Jiwa NM, Kummer JA, et al. Activated cytotoxic T cells as prognostic marker in Hodgkin's disease. Blood 1997;89:1376–82.

22. Ten Berge RL, Oudejans JJ, Dukers DF, Meijer JW, Ossenkoppele GJ, Meijer CJ. Percentage of activated cytotoxic T-lymphocytes in anaplastic large cell lymphoma and Hodgkin's disease: an independent biological prognostic marker. Leukemia 2001;15:458–64.

23. Alvaro T, Lejeune M, Salvado MT, et al. Outcome in Hodgkin's lymphoma can be predicted from the presence of accompanying cytotoxic and regulatory T cells. Clin Cancer Res 2005;11:1467–73.

24. Sanchez-Aguilera A, Montalban C, de la Cueva P, et al. Tumor microenvironment and mitotic checkpoint are key factors in the outcome of classic Hodgkin lymphoma. Blood 2006;108:662–8.

25. Carde P, Koscielny S, Franklin J, et al. Early response to chemotherapy: a surrogate for final outcome of Hodgkin's disease patients that should influence initial treatment length and intensity? Ann Oncol 2002;13 Suppl 1:86–91.

26. Montaner D, Tarraga J, Huerta-Cepas J, et al. Next station in microarray data analysis: GEPAS. Nucleic Acids Res 2006;34:W486–91.

27. Green GH, Diggle PJ. On the operational characteristics of the Benjamini and Hochberg false discovery rate procedure. Stat Appl Genet Mol Biol 2007;6:Article27.

28. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;102:15545–50.

29. Koch I, Slotta-Huspenina J, Hollweck R, et al. Real-time quantitative RT-PCR shows variable, assay-dependent sensitivity to formalin fixation: implications for direct comparison of transcript levels in paraffin-embedded tissues. Diagn Mol Pathol 2006;15:149–56.

30. Bonanomi A, Kojic D, Giger B, et al. Quantitative cytokine gene expression in human tonsils at excision and during histoculture assessed by standardized and calibrated real-time PCR and novel data processing. J Immunol Methods 2003;283:27–43.

31. Ohl F, Jung M, Radonic A, Sachs M, Loening SA, Jung K. Identification and validation of suitable endogenous reference genes for gene expression studies of human bladder cancer. J Urol 2006;175:1915–20.

32. Vandesompele J, De Preter K, Pattyn F, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 2002;3:RESEARCH0034.

33. Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clin Chem 1993;39:561–77.

34. Navarro A, Gaya A, Martinez A, et al. MicroRNA expression profiling in classical Hodgkin lymphoma. Blood 2008;111:2825–32.

35. Kuppers R, Klein U, Schwering I, et al. Identification of Hodgkin and Reed-Sternberg cell-specific genes by gene expression profiling. J Clin Invest 2003;111:529–37.

36. Fiumara P, Snell V, Li Y, et al. Functional expression of receptor activator of nuclear factor κB in Hodgkin disease cell lines. Blood 2001;98:2784–90.

37. Oelmann E, Herbst H, Zuhlsdorf M, et al. Tissue inhibitor of metalloproteinases 1 is an autocrine and paracrine survival factor, with additional immune-regulatory functions, expressed by Hodgkin/Reed-Sternberg cells. Blood 2002;99:258–67.

38. Manoli T, Gretz N, Grone HJ, Kenzelmann M, Eils R, Brors B. Group testing for pathway analysis improves comparability of different microarray datasets. Bioinformatics 2006;22:2500–6.

39. Van Imhoff GW, Boerma EJ, van der Holt B, et al. Prognostic impact of germinal center-associated proteins and chromosomal breakpoints in poor-risk diffuse large B-cell lymphoma. J Clin Oncol 2006;24:4135–42.

40. Gillis AJ, Stoop HJ, Hersmus R, et al. High-throughput microRNAome analysis in human germ cell tumours. J Pathol 2007;213:319–28.

41. Devilard E, Bertucci F, Trempat P, et al. Gene expression profiling defines molecular subtypes of classical Hodgkin's disease. Oncogene 2002;21:3095–102.

42. Hagemann T, Bozanovic T, Hooper S, et al. Molecular profiling of cervical cancer progression. Br J Cancer 2007;96:321–8.

43. Glenn ST, Jones CA, Liang P, Kaushik D, Gross KW, Kim HL. Expression profiling of archival renal tumors by quantitative PCR to validate prognostic markers. Biotechniques 2007;43:639–40, 642–633, 647.

44. Noutsias M, Rohde M, Block A, et al. Preamplification techniques for real-time RT-PCR analyses of endomyocardial biopsies. BMC Mol Biol 2008;9:3.

45. Morente MM, Piris MA, Abraira V, et al. Adverse clinical outcome in Hodgkin's disease is associated with loss of retinoblastoma protein expression, high Ki67 proliferation index, and absence of Epstein-Barr virus-latent membrane protein 1 expression. Blood 1997;90:2429–36.

46. Weaver BA, Cleveland DW. Decoding the links between mitosis, cancer, and chemotherapy: The mitotic checkpoint, adaptation, and cell death. Cancer Cell 2005;8:7–12.

47. Duxbury MS, Ito H, Zinner MJ, Ashley SW, Whang EE. RNA interference targeting the M2 subunit of ribonucleotide reductase enhances pancreatic adenocarcinoma chemosensitivity to gemcitabine. Oncogene 2004;23:1539–48.

48. Hilliker C, Delabie J, Speleman F, et al. Localization of the gene (RSN) coding for restin, a marker for Reed-Sternberg cells in Hodgkin's disease, to human chromosome band 12q24.3 and YAC cloning of the locus. Cytogenet Cell Genet 1994;65:172–6.

49. Slep KC, Vale RD. Structural basis of microtubule plus end tracking by XMAP215, CLIP-170, and EB1. Mol Cell 2007;27:976–91.

50. Wang TL, Diaz LA, Jr., Romans K, et al. Digital karyotyping identifies thymidylate synthase amplification as a mechanism of resistance to 5-fluorouracil in metastatic colorectal cancer patients. Proc Natl Acad Sci U S A 2004;101:3089–94.

51. Farrugia DC, Ford HE, Cunningham D, et al. Thymidylate synthase expression in advanced colorectal cancer predicts for response to raltitrexed. Clin Cancer Res 2003;9:792–801.

52. Kamal A, Boehm MF, Burrows FJ. Therapeutic and diagnostic implications of Hsp90 activation. Trends Mol Med 2004;10:283–90.

53. Sato S, Fujita N, Tsuruo T. Modulation of Akt kinase activity by binding to Hsp90. Proc Natl Acad Sci U S A 2000;97:10832–7.

54. Broemer M, Krappmann D, Scheidereit C. Requirement of Hsp90 activity for IκB kinase (IKK) biosynthesis and for constitutive and inducible IKK and NF-κB activation. Oncogene 2004;23:5378–86.

55. Georgakis GV, Li Y, Rassidakis GZ, Martinez-Valdez H, Medeiros LJ, Younes A. Inhibition of heat shock protein 90 function by 17-allylamino-17-demethoxy-geldanamycin in Hodgkin's lymphoma cells down-regulates Akt kinase, dephosphorylates extracellular signal-regulated kinase, and induces cell cycle arrest and cell death. Clin Cancer Res 2006;12:584–90.

56. Alvaro T, Lejeune M, Camacho FI, et al. The presence of STAT1-positive tumor-associated macrophages and their relation to outcome in patients with follicular lymphoma. Haematologica 2006;91:1605–12.

57. Dave SS, Wright G, Tan B, et al. Prediction of survival in follicular lymphoma based on molecular features of tumor-infiltrating immune cells. N Engl J Med 2004;351:2159–69.

e-*Blood*

A molecular risk score based on 4 functional pathways for advanced classical Hodgkin lymphoma

*Beatriz Sánchez-Espiridión,[1] *Carlos Montalbán,[2] Ángel López,[1] Javier Menárguez,[3] Pilar Sabín,[3] Carmen Ruiz-Marcellán,[4] Andrés Lopez,[4] Rafael Ramos,[5] Jose Rodríguez,[5] Araceli Cánovas,[6] Carmen Camarero,[6] Miguel Canales,[7] Javier Alves,[7] Reyes Arranz,[8] Agustín Acevedo,[8] Antonio Salar,[9] Sergio Serrano,[9] Águeda Bas,[10] Jose M. Moraleda,[10] Pedro Sánchez-Godoy,[11] Fernando Burgos,[11] Concepción Rayón,[12] Manuel F. Fresno,[12] José García Laraña,[10] Mónica García-Cosío,[2] Carlos Santonja,[13] Jose L. López,[13] Marta Llanos,[14] Manuela Mollejo,[15] Joaquín González-Carrero,[16] Ana Marín,[17] Jerónimo Forteza,[18] Ramón García-Sanz,[19] Jose F. Tomás,[20] Manuel M. Morente,[1] Miguel A. Piris,[1] and Juan F. García,[1,20] on behalf of the Spanish Hodgkin Lymphoma Study Group

[1]Molecular Pathology Programme, the Genetic and Molecular Epidemiology Group, and the Tumour Bank, Centro Nacional de Investigaciones Oncológicas (CNIO), Madrid; [2]Hospital Ramón y Cajal, Madrid; [3]Hospital Gregorio Marañón, Madrid; [4]Hospital Vall d'Hebron, Barcelona; [5]Hospital Son Dureta, Palma de Mallorca; [6]Hospital de Cruces, Baracaldo; [7]Hospital La Paz, Madrid; [8]Hospital La Princesa, Madrid; [9]Hospital del Mar, Barcelona; [10]Hospital Universitario Virgen de la Arrixaca, Murcia; [11]Hospital Severo Ochoa, Madrid; [12]Hospital Central de Asturias, Oviedo; [13]Fundación Jiménez Díaz, Madrid; [14]Hospital Universitario de Canarias, Tenerife; [15]Hospital Virgen de la Salud, Toledo; [16]Hospital Xeral-Cíes, Vigo; [17]Hospital Virgen del Rocío, Sevilla; [18]Hospital Clínico Universitario, Santiago Compostela; [19]Hospital Universitario de Salamanca, Salamanca; and [20]M. D. Anderson España, Madrid, Spain

Despite improvement in the treatment of advanced classical Hodgkin lymphoma, approximately 30% of patients relapse or die as result of the disease. Current predictive systems, determined by clinical and analytical parameters, fail to identify these high-risk patients accurately. We took a multistep approach to design a quantitative reverse-transcription polymerase chain reaction assay to be applied to routine formalin-fixed paraffin-embedded samples, integrating genes expressed by the tumor cells and their microenvironment. The significance of 30 genes chosen on the basis of previously published data was evaluated in 282 samples (divided into estimation and validation sets) to build a molecular risk score to predict failure. Adequate reverse-transcription polymerase chain reaction profiles were obtained from 262 of 282 cases (92.9%). Best predictor genes were integrated into an 11-gene model, including 4 functional pathways (cell cycle, apoptosis, macrophage activation, and interferon regulatory factor 4) able to identify low- and high-risk patients with different rates of 5-year failure-free survival: 74% versus 44.1% in the estimation set ($P < .001$) and 67.5% versus 45.0% in the validation set ($P = .022$). This model can be combined with stage IV into a final predictive model able to identify a group of patients with very bad outcome (5-year failure-free survival probability, 25.2%). (*Blood*. 2010;116(8): e12-e17)

Introduction

Classical Hodgkin lymphoma (cHL) is assumed to be a curable tumor, but an important fraction of patients with advanced disease do not respond favorably to the current standard chemotherapy regimens whose base is adriamycin. The most widely used and reproducible prognostic score is the product of clinical and analytical parameters integrated in the International Prognostic Score (IPS), but it still fails to identify accurately, at the moment of diagnosis, a significant fraction of patients with very poor prognosis.[1-3] Thus, the identification of biomarkers that, at diagnosis, may be consistently associated with nonresponse is essential for the recognition of patients at high risk of treatment failure to establish a more rational risk-adapted treatment strategy.

cHL represents a distinctive model of histologic complexity, with a minor population of the neoplastic Hodgkin and Reed-Sternberg (HRS) cells diluted in a reactive inflammatory background composed of nonneoplastic B and T cells, macrophages, eosinophils, neutrophils, and plasma cells. The complex relationship between the HRS cells and their microenvironment is only partially understood; however, important, if fragmentary, advances in our understanding are steadily being made.[4] The clinical outcome of cHL has been found to be related to the expression of multiple biologic markers alone[5-8] or in combination,[9] expressed either by the tumor HRS cells, macrophages, regulatory T cells, or other nonneoplastic cell subpopulations.[10-13]

Some of these previous analyses rely on array-based gene expression analyses, which use frozen tissue and in most cases can only provide retrospective information. Investigators of other studies have used immunohistochemical staining, with some inherent limitations to the reproducibility of the data thus generated. It is now feasible to apply multigenic predictive molecular tests in advanced cHL patients in a routine setting by the use of a quantitative reverse-transcription polymerase chain reaction

Submitted February 12, 2010; accepted April 10, 2010. Prepublished online as *Blood* First edition paper, May 17, 2010; DOI 10.1182/blood-2010-02-270009.

*B.S.-E. and C.M. contributed equally to this work.

The online version of this article contains a data supplement.

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

(RT-PCR) assay as we here describe, incorporating a selected number of genes that capture information from tumor and microenvironment cell components, designed for application to routine formalin-fixed paraffin-embedded (FFPE) samples and that can be used at the moment of the initial diagnosis.

Methods

Patients and samples

Previous studies allowed us to identify a group of genes whose expression was associated with the response of patients with advanced cHL to standard first-line treatment. Thus, the selection of genes to be analyzed was determined by results previously obtained in 2 independent series of 29 and 52 advanced cHL patients.[10,11]

The patients included in this study fulfilled stringent, previously described criteria (ie, age older than 16 years; advanced cHL; Ann Arbor stage IV, III, or IIB with bulky masses[1]; proven HIV-negative status) who have been treated with a first-line standard chemotherapy regimen that included adriamycin—ABVD (adriamycin, bleomycin, vinblastine, and dacarbazine) or ABVD variants—and for whom information was available about the achievement of complete remission (CR) and a follow-up of at least 12 months thereafter, which is a well-known and accepted surrogate indication of the course of the disease. With respect to the latter, patients were considered to have had a favorable course if they had achieved CR and maintained it for at least 12 months or with an unfavorable course if they had either not achieved CR or if they had once had it but had relapsed during the following 12 months. All tissue samples consisted of representative pretreatment lymph node biopsies collected after revision and approved by the institutional review board of the participating institutions of the Spanish Hodgkin Lymphoma Study Group. The study initially included 282 FFPE patients, who were randomly split and assigned to the training (194 cases) or validation sets (88 cases) on the basis of the minimum estimated sample size to derive the final model (Table 1).

Additional exclusion criteria were insufficient RNA quality (purity ratio A260:A230 < 1.7) or a weak RT-PCR signal (average cycle threshold > 35) for the reference genes or in more than 10 genes of the assay. As a result, 20 patients were excluded, and the remaining 262 patients (183 in the training group and 79 in the validation group) meeting these criteria were included in the statistical analysis (Table 1).

Table 1. Clinical characteristics of the cHL series

Characteristic	Estimation, n (%)	Validation, n (%)	Total, n (%)	*P*
Age, y	(n = 183)	(n = 79)	(n = 262)	
Younger than 45	133 (72.67)	56 (70.88)	189 (72.14)	.883
45 or older	50 (27.32)	23 (29.11)	73 (27.86)	
Sex	(n = 183)	(n = 79)	(n = 262)	
Male	99 (54.10)	51 (64.55)	150 (57.25)	.151
Female	84 (45.90)	28 (35.44)	112 (42.75)	
Stage IV	(n = 182)	(n = 79)	(n = 261)	
No	132 (72.52)	47 (59.49)	179	.053
Yes	50 (37.87)	32 (40.50)	82	
IPS code	(n = 182)	(n = 79)	(n = 261)	
Less than 3	109 (59.89)	41 (51.90)	150 (57.47)	.288
3 or greater	73 (40.11)	38 (48.10)	111	
Outcome	(n = 183)	(n = 79)	(n = 262)	
F	132 (72.14)	57 (72.16)	189 (72.14)	> .999
U	51 (27.86)	22 (27.84)	73 (27.86)	

Clinical characteristics of patients with adequate RT-PCR profiles. Summary of the clinical characteristics of the patients in the estimation and validation sets that yielded suitable analyzable data (262 of 282; 92.90%). Differences in distribution of standard clinical parameters (age, sex, stage, IPS and outcome) between estimation and validation datasets tested by Pearson chi-square with Yates correction were not statistically significant (IPS values of 0-2 classified as low IPS; IPS values > 2 classified as high IPS).

cHL indicates classical Hodgkin lymphoma; F, favorable response; IPS, International Prognostic Score; RT-PCR, reverse transcription polymerase chain reaction; U, unfavorable response.

Gene selection

The genes included in the assay initially were selected from 2 preliminary expression-profiling studies[10,11] that rendered a list of genes expressed by HRS and microenvironment cell subpopulations identified in unfavorable cHL patients. Selected genes were primarily chosen on the basis of their prognostic ability and capacity to represent biologic functions identified as relevant in cHL pathogenesis.[11] In addition, the strength and consistency of primer and probe performance also were taken into account.[11] The initial selection consisted of 30 genes, including genes expressed by the neoplastic cells involved in the cell cycle (G_2/M), apoptosis, histones, chaperones, drug metabolism, and mitogen-activated protein kinase signatures, and from microenvironment genes expressed by different cellular or functional populations of T cells, monocytes, macrophages, and dendritic cells. Details of the RT-PCR assays are available in supplemental Table 1 (available on the *Blood* Web site; see the Supplemental Materials link at the top of the online article).

Analysis of gene expression

Gene expression was analyzed by the use of a customized TaqMan low-density array platform (Micro Fluidic Cards; Applied Biosystems) on FFPE as previously described.[11,14] A preamplification step (PreAmp; Applied Biosystems) was used to improve the sensitivity of our assay for low-abundance target genes available from FFPE samples.[15-17] Reactions were performed by use of the ABI PRISM 7900HT Sequence Detection system (Applied Biosystems), and we measured the expression of each gene in triplicate and then normalized it with a set of 2 reference genes (*HMBS* and *GUSB*) whose uniform expression in cHL tumor samples was tested in previous studies.[11] Missing values were imputed using the K-nearest neighbor algorithm.[18]

Statistical analysis

Differences in the distributions of standard clinical parameters (age, sex, stage, IPS, the individual variables contained in IPS, and outcome) in the estimation and validation datasets were tested by the Pearson χ^2 test (Table 1).

The first end point of this study was the response to standard first-line treatment considering favorable response (F) and unfavorable response (U), as mentioned previously. Data from second-line and salvage therapies and/or bone-marrow transplantation were not considered.

The selection of the best predictive genes and the logistic regression model was on the basis only of the data from the training group of 183 patients, without any previous survival analysis that used information from the validation group. Univariate regression analysis was performed with treatment response (F vs U) as the dependent variable to identify genes significantly associated ($P < .05$) with outcome. In addition, final gene selection analysis was performed by cross-validation with the use of 3 prediction algorithms (http://tnasas.bioinfo.cnio.es/): diagonal linear discriminant analysis,[19] support vector machines,[20] and K-nearest neighbor.[21] Cross validation was used to test the classification ability of the initial set of significant genes to choose the strongest predictor genes, which were classified into functional groups on the basis of their known biologic relationship and their coregulated expression as estimated by the Pearson correlation coefficient. Individual genes from each functional group were weighted by the use of linear discriminant analysis.[22] Finally, these functional gene clusters associated with cHL outcome were analyzed in a multivariate logistic regression model with response to therapy (F vs U) as a dependent variable. In this way, an algorithm was derived that combines these measurements into a quantitative "molecular risk score" (MRS), which can be used as a continuous variable to estimate the probability of treatment response. The MRS cut-off points were prespecified by the use of an area under the receiver operating characteristic curve (ROC) analysis to define different risk groups. (See the supplemental Appendix for details of the statistical analysis and methods.) Finally, performance

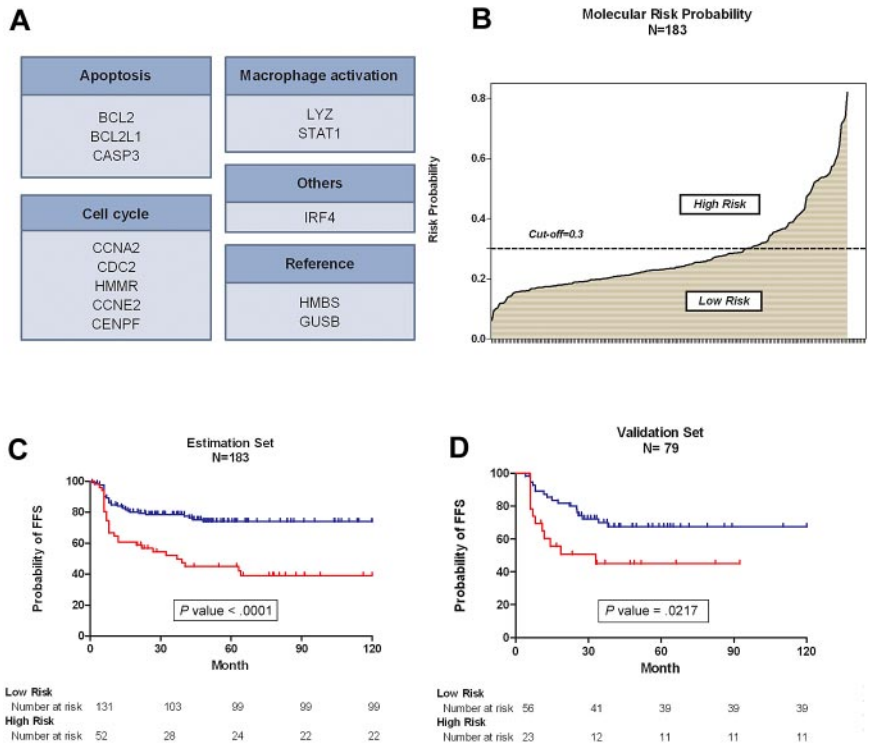


Figure 1. Panel of 11 genes and the molecular risk algorithm. (A) The molecular risk algorithm is determined on the basis of the relative contributions of each of the 4 gene functional groups from the tumoral HRS and their reactive microenvironment as follows: MRS = exp (fx)/(1 + exp [fx]), where fx = (−0.913) + (0.401 × apoptosis) + (0.284 × cell cycle) + (−0.301 × monocyte) + (−0.143 × IRF4). Coefficients were derived from a multivariate analysis in which positive values indicate that a greater level of expression is correlated with a worse outcome, and negative coefficients indicate that a greater level of expression of the pathways is associated with a better outcome. (B) MRS as a continuous function was used to set a threshold for stratifying patients by ROC analysis. Patients were stratified according to the levels of the molecular risk score into low-risk (< 0.3) and high-risk (≥ 0.3) groups. (C-D) Survival estimates of FFS in patients from estimation (n = 183) and validation (n = 79) sets after classification into risk groups. Kaplan-Meier analysis and the log-rank test gave significant results in both estimation and validation sets, indicating the potential prognostic capacity of the algorithm developed here.

of the logistic regression model was tested in the validation group of patients (n = 79).

For graphical representation, survival analyses were performed with the Kaplan-Meier method and long-rank test separately in the training and validation series and in the entire series. Because the primary objective of the study was to identify patients at high risk of treatment failure, we used failure-free survival (FFS) as the fundamental end point for survival analysis. FFS was defined as the time interval between treatment initiation and treatment failure or last follow-up. Failure was defined as either the failure to achieve CR or the occurrence of progressive disease, irrespective of whether there had been an initial CR. Overall survival (OS), an end point whose significance is imperfect because it is conditioned by the effect of subsequent eventual treatments and complication of treatment, was included as a secondary end point in the survival analyses, defined as the time interval between diagnosis and death caused by the lymphoma.

Finally, in the whole series, a multivariate Cox proportional hazards model, including the data at diagnosis, the IPS stratified as previously defined (0-2 vs ≥ 3),[1] and its 7 individual variables (hemoglobin < 1.5 g/dL; albumin < 4 g/dL; leukocytosis ≥ 15 000/mm^3; lymphopenia < 600/mm^3; age ≥ 45 years; male sex; stage IV), was applied to test the independence of the MRS, including the remaining significant variables in a final integrative model.

All statistical analyses were 2-sided; values of P less than .05 were considered to be significant. These were performed with SPSS 15.0 (SPSS Inc). Survival curves were assessed by the Kaplan-Meier method, and risk groups were compared by the log-rank test. Plots were generated with the use of GraphPad Prism Version 5 (GraphPad Software, Inc).

Results

Gene selection and development of the predictor model

In training series, univariate regression analysis of the expression data for the 30 initially selected genes revealed 20 genes to significantly predict failure to first-line treatment (supplemental Table 3). When cross validation was applied, the genes most frequently found in prognostic models consisted of a panel of 11 genes that were included in the final model: *BCL2*, *BCL2L1*, *CASP3*, *HMMR*, *CENPF*, *CCNA2*, *CCNE2*, *CDC2*, *LYZ*, *STAT1*, and *IRF4*.

To derive the model, we took a 2-step approach, first combining individual gene-expression patterns into precise functional pathways and then subsequently correlating these functional groups with the clinical outcome by using multivariate logistic regression. Final selected genes were weighted by the use of linear discriminant analysis and clustered into their corresponding functional pathways defined as macrophage activation (*LYZ*, *STAT1*), cell cycle (*HMMR*, *CENPF*, *CCNA2*, *CCNE2*, *CDC2*), and apoptosis (*BCL2*, *BCL2L1*, *CASP3*; Figure 1A). The Pearson correlation coefficient was significant for the genes included in each of the signatures (P < .001) apart from *IRF4*, which was included as an independent predictive gene because there were neither distinct functional relationships nor statistically significant correlations with other genes or pathways (supplemental Table 4). These functional groups captured information about the tumoral HRS cells and their nontumoral microenvironment, in agreement with previous studies.[10] The multivariate logistic regression analysis integrating these pathways showed that cell cycle and apoptosis terms were associated with an unfavorable outcome of patients, whereas macrophage activation and *IRF4* signatures had protective effects.

Thus, the optimized final model was determined on the basis of the relative contributions of each of the 4 functional terms, as described in the following equation: constant (−0.913) + (0.401 × apoptosis) + (0.284 × cell cycle) + (−0.301 × macrophage activation) + (−0.143 × *IRF4*). The continuous probability function generated by the logistic regression was defined as the MRS to treatment failure and ranged from 0.06 to 0.813 (Figure 1B). ROC analysis was used to define a threshold for stratifying patients, and the largest area under the curve was obtained by the use of 0.3 as the threshold, thus dividing the series into high-risk (MRS ≥ 0.3) and low-risk (MRS < 0.3) cases (Figure 1B-C).

Validation of the MRS

The MRS differed significantly between the various outcome groups (supplemental Figure 2) and predicted treatment response with an accuracy of 68.9% in the estimation dataset and 70% in the validation dataset. Because treatment failure was the main end point used to derive the logistic regression function, FFS as previously defined, was used for validation of the model and Kaplan-Meier analysis of survival. Predicted probabilities also identified 2 risk groups associated with FFS in both the estimation and validation sets (Figure 1C; supplemental Figure 3). FFS probabilities at 5 years were 74.0% versus 44.1% ($P < .001$) in the training set and 67.5% versus 45.0% ($P = .022$) in the validation set.

In addition, the analyses of OS showed different risk groups identified by the MRS in the estimation group of patients. The differences were not significant in the validation series, probably because of the limited number of events (supplemental Figure 3).

There were no significant statistical differences in the MRS distributions of the 2 histologic subtypes, nodular sclerosis and mixed-cellularity cHL ($P = .186$, t test), and the differences in survival between risk groups identified by the MRS remained significant in both the estimation and validation sets stratified by histologic subtype (supplemental Figure 4).

Integrative model using MRS and clinical variables

In the whole series, a multivariate Cox proportional hazards model with FFS as the dependent variable and including the MRS and the IPS, we found that only the MRS was significant (Table 2). No interaction was observed between the IPS or the individual IPS variables and the MRS low- and high-risk groups (supplemental Table 5). Thus, to compare the molecular risk algorithm and the individual IPS components, a backward stepwise selection Cox model was tested, with FFS as the dependent variable and including the MRS and the individual components of the IPS in the global series of samples. Only MRS and stage IV were statistically significant (Table 2) and so were retained in the final Cox regression model. Patient stratification into quartiles on the basis of the Cox model identified a subgroup of advanced cHL patients (fourth quartile) with a very poor outcome: 5-year FFS of 24.3% ($P < .001$; Figure 2).

Table 2. MRS and the IPS variables

	P	Hazard ratio (95% CI)
IPS and MRS (n = 262)*		
MRS	< .001§	24.362 (6.268-94.691)
IPS	.205	1.111 (0.944-1.309)
IPS variables and MRS (n = 262)†		
MRS	< .001§	24.715 (6.804-89.768)
Hemoglobin < 10.5 g/dL	.352	1.243 (0.787-1.963)
Albumin < 4 g/dL	.504	1.157 (0.755-1.772)
Leucocytosis ≥ 15 000/mm^3	.256	1.312 (0.821-2.095)
Lymphopenia < 600/mm^3	.555	0.820 (0.425-1.583)
Age ≥ 45 y	.369	1.227 (0.785-1.916)
Stage IV	.041§	1.552 (1.018-2.360)
Male sex	.753	0.936 (0.618-1.416)
PT-BR Integrative Cox model (n = 262)‡		
MRS	< .001§	23.782 (6.041-94.340)
Stage IV	.025§	1.409 (1.044-1.900)

95% CI indicates 95% confidence interval; IPS, International Prognostic Score; and MRS, molecular risk score.

*Multivariate Cox regression analysis considering MRS and the IPS.

†Univariate Cox regression analyses considering individual IPS variables and MRS.

‡Cox regression analysis of the final variables included in the integrative model (molecular risk plus stage IV) obtained by backward stepwise selection.

§Significant.

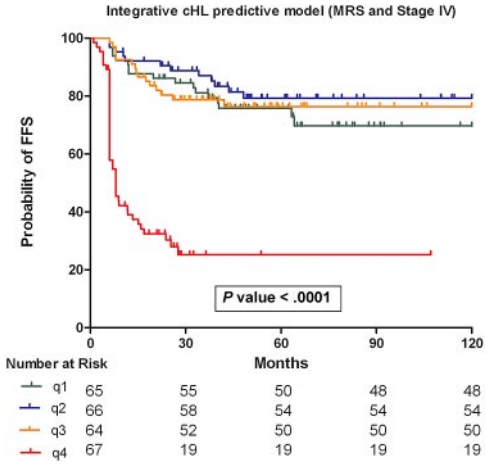


Figure 2. Integrative risk model of cHL. The final Cox model integrates the MRS and clinical variable stage IV. Patients in quartiles 1, 2, and 3 have comparable FFS rates at 5 years (76.4%, 79.3%, and 69.7%, respectively) whereas patients in quartile 4 show a 5-year FFS of 24.3% ($P < .001$).

Discussion

Here we describe, in a series of patients with advanced cHL, a 4-cluster/11-gene model derived from an initial selection of 30 potentially predictive markers that can be detected by RT-PCR and integrated into a molecular risk algorithm that can identify patient subgroups with very different probabilities of treatment failure. This approach is determined on the basis of reliable quantitative RT-PCR techniques applicable to paraffin-embedded diagnostic samples and follows similar approaches taken in breast cancer and other tumor types.[23,24] This benefits from previous expression profiling studies performed in cHL,[10] and the improved knowledge about the role of the tumoral cell and the microenvironment in the pathogenesis and outcome of this disease.[13,25,26] This MRS, calculated in an initial estimation set, was confirmed in an independent set. Genes included in the score were selected on the basis of previous gene expression profiling data generated in independent sets of patients[11,27] and can be classified in 4 functional pathways.

The final 4-cluster/11-gene model can additionally incorporate one of the well-established clinical variables (stage IV), thus integrating the main molecular characteristics of the tumors related with treatment response and tumor burden estimation in a single scoring system. The multivariate Cox model indicates that most patients with stage IV cHL and with a high MRS (≥ 0.3) will have a very poor outcome, with 5-year FFS probability of 24.3% and OS probability of 76.3%. Therefore, this combination of stage IV and high MRS identifies a group of patients with very bad outcome who could have been initially missed by consideration of the IPS alone.[1,3]

It is of note that in the present series the IPS did not show any significant prognostic influence on FFS. From the individual components of the IPS, only stage IV disease remained significant in multivariate analyses. This finding is in agreement with recent studies in which the authors demonstrated that IPS is of limited utility in advanced HL cases treated in the modern era, where more

accurate pathologic diagnosis, improved control of therapy, use of growth factors, and enhanced supportive care are yielding better outcomes compared with historic results.[3] Thus, deviations from the standard therapy cannot be justified on the basis of the survival prediction capacity of the IPS, and identification of high-risk populations needs to be supplemented with molecular markers.

In the model, the expression of *BCL2*, *BCL2L1*, *CASP3*, *HMMR*, *CENPF*, *CCNA2*, *CCNE2*, and *CDC2* included in the apoptosis and cell-cycle pathways, respectively, were correlated with short FFS. The expression of various antiapoptotic *BCL2* family proteins has been repeatedly reported in HRS cells,[7,28,29] thereby contributing to the survival of HRS cells. *BCL2* and *BCL2L1* (*BCL-XL*) both frequently are expressed by HRS cells in cHL, and their levels have been associated with inferior FFS in patients treated with ABVD or equivalent regimens,[5] thus confirming the importance of this group of apoptotic regulators for cHL outcome prediction. The second signature, cell cycle, is mainly composed of genes coding for regulatory proteins of the S and G_2/M phases of the cell cycle, thus directly related with cell proliferation. Again, expression of some of these markers has been previously described at the protein level in cHL,[7,30,31] and a significant prognostic value for CDC2 (CDK1) and CCNA2 (cyclin A2) protein expression in non-Hodgkin lymphomas was also found for both disease-free survival and OS. Interestingly, this aberrant association between increased expressions of antiapoptotic proteins and growth fraction-associated proteins in HRS cells provides further evidence that cell cycle and apoptosis regulation are profoundly disturbed and closely related in the disease, further justifying the inclusion of these 2 pathways in the predictive model.

Moreover, pathways involved in cell cycle and apoptosis regulation are rational therapeutic targets. Indeed, inhibitors of Cyclin-Cdk complexes (including Cyclin E-Cdk2, Cyclin A-Cdk2, and Cyclin B-Cdk1 complexes) are currently under preclinical and clinical investigation in different cancer types,[32-34] and these drugs could be considered for the treatment of advanced and refractory cHL patients. Likewise, the potential for targeting BCL2-related proteins in lymphoma is promising. Small molecule inhibitors of the Bcl-2 family have demonstrated high target affinity and an improved toxicity profile, and clinical trials of these agents are yielding interesting results.[35]

Confirming previous observations about the importance of the reactive microenvironment for cHL patient outcome,[9,11,12] *LYZ* and *STAT1* genes, expressed at high levels in a subset of tissue monocytes and activated macrophages, also are included in this model, and correlated with prolonged FFS and better outcome. The relevance of the cell composition of the reactive background in cHL has been reinforced by the data recently reported by Steidl et al.[36] They used gene expression profiling to identify a gene signature of tumor-associated macrophages that is associated with treatment failure, in an approach methodologically similar to previous reports from our group and others.[10,37] The discrepancy in the results concerning the role of macrophages may have arisen from technical differences (RT-PCR vs microarray gene-expression and immunohistochemistry) or the selection of markers such as *LYZ* and *STAT1* in this study, reflecting a specific functional status of the monocyte-macrophages.[38]

In addition, *IRF4* (*MUM1*) expression was associated with longer FFS. This gene is an interferon regulatory factor, lymphocyte specific, induced after nuclear factor-κB (NF-κB) activation, which controls B-cell proliferation and differentiation and has recently been shown to be up-modulated by CD40 engagement in HL cells.[39] Interestingly, the lack of IRF4 protein has been previously associated with outcome in cHL,[40] representing a potential adverse prognostic factor. In addition, both *IRF4* and *BCL2L1* represent well-known NF-κB target genes[41,42] whose expression is induced after NF-κB pathway activation. Thus, our final model includes important subrogates from the NF-κB activation, which is thought to be an essential pathogenetic mechanism in this disease. It is of note that Bednarski et al[43] recently described how the inhibition of the canonical NF-κB pathway enhances the proapoptotic effects of adriamycin, thus also identifying NF-κB inhibition as an interesting therapeutic approach.

In conclusion, we have developed a molecular risk algorithm, on the basis of feasible and reproducible molecular techniques, that is capable of stratifying at the moment of diagnosis advanced cHL patients with different outcome. Moreover, a combination of this algorithm with the presence of clinical stage IV disease can be used to identify a group of cHL patients with a very bad outcome who could benefit from more intensive therapeutic approaches.

These results are promising, but further validation in larger and independent series of patient is needed for the model to become established as part of the clinical routine. Also the predictive value of this model should also be tested in patients treated with modern intensive chemotherapy.

Acknowledgments

We thank Nuria Malats, from the Genetic and Molecular Epidemiology Group at the CNIO, for her useful advice and supervision of the statistical analysis. We also thank Marién Castillo and Laura Cereceda, at the CNIO Tumor Bank, for collecting the human tumor samples and for their excellent assistance with data management. Members of the Spanish Hodgkin Lymphoma Study Group are cited in the supplemental Appendix.

This work was supported by grants from the Fondo de Investigaciones Sanitarias (PI08/1985, PI05/1623, PI05/2800, PI05/2327, RETIC RD06/0020/0107) and the Ministerio de Ciencia y Tecnología (SAF2008-03 871), Spain. B.S.-E. is supported by a grant from the Ministerio de Ciencia e Innovación (FIS), Spain.

Authorship

Contribution: B.S.E., C.M., M.A.P., and J.F.G. contributed to the conception and design of the study, the analysis and interpretation of the data, and the drafting of the article; B.S.E. and A.L. performed statistical analysis; M.M.M. managed tissue banking; J.M., P.S., C.R.M., A.L., R.R., J.R., A.C., C.C., M.C., J.A., R.A., A.A., A.S., S.S., A.B., J.M.M., P.S.-G., F.B., C.R., M.F.F., J.G.L., M.G.-C., C.S., J.L.L., M.L.I., M.M., J.G.C., A.M., J.F., R.G.C, and J.F.T. managed patient databases and contributed with tumor samples and clinical follow-up; and all the authors read and approved the final version of the manuscript.

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

Members of the Spanish Hodgkin Lymphoma Study Group are listed in the supplemental Appendix.

Correspondence: Miguel A Piris, MD, Lymphoma Group, Spanish National Cancer Centre (CNIO). E-28029 Madrid, Spain; e-mail: mapiris@cnio.es.

References

1. Hasenclever D, Diehl V. A prognostic score for advanced Hodgkin's disease. International Prognostic Factors Project on Advanced Hodgkin's Disease. *N Engl J Med.* 1998;339(21):1506-1514.
2. Canellos GP, Anderson JR, Propert KJ, et al. Chemotherapy of advanced Hodgkin's disease with MOPP, ABVD, or MOPP alternating with ABVD. *N Engl J Med.* 1992;327(21):1478-1484.
3. Moccia AA, Donaldson J, Chhanabhai M, et al. The International Prognostic Project Score (IPS) in advanced-stage Hodgkin lymphoma has limited utility in patients treated in the modern era. *Blood.* 2009;114(22). Abstract 1554.
4. Herreros B, Sanchez-Aguilera A, Piris MA. Lymphoma microenvironment: culprit or innocent? *Leukemia.* 2008;22(1):49-58.
5. Rassidakis GZ, Medeiros LJ, Vassilakopoulos TP, et al. BCL-2 expression in Hodgkin and Reed-Sternberg cells of classical Hodgkin disease predicts a poorer prognosis in patients treated with ABVD or equivalent regimens. *Blood.* 2002; 100(12):3935-3941.
6. Rassidakis GZ, Medeiros LJ, McDonnell TJ, et al. BAX expression in Hodgkin and Reed-Sternberg cells of Hodgkin's disease: correlation with clinical outcome. *Clin Cancer Res.* 2002;8(2):488-493.
7. García JF, Camacho FI, Morente M, et al. Hodgkin and Reed-Sternberg cells harbor alterations in the major tumor suppressor pathways and cell-cycle checkpoints: analyses using tissue microarrays. *Blood.* 2003;101(2):681-689.
8. Herling M, Rassidakis GZ, Vassilakopoulos TP, Medeiros LJ, Sarris AH. Impact of LMP-1 expression on clinical outcome in age-defined subgroups of patients with classical Hodgkin lymphoma. *Blood.* 2006;107(3):1240; author reply 1241.
9. Montalbán C, Garcia JF, Abraira V, et al. Influence of biologic markers on the outcome of Hodgkin's lymphoma: a study by the Spanish Hodgkin's Lymphoma Study Group. *J Clin Oncol.* 2004; 22(9):1664-1673.
10. Sánchez-Aguilera A, Montalban C, de la Cueva P, et al. Tumor microenvironment and mitotic checkpoint are key factors in the outcome of classic Hodgkin lymphoma. *Blood.* 2006;108(2):662-668.
11. Sánchez-Espiridión B, Sanchez-Aguilera A, Montalban C, et al. A TaqMan low-density array to predict outcome in advanced Hodgkin's lymphoma using paraffin-embedded samples. *Clin Cancer Res.* 2009;15(4):1367-1375.
12. ten Berge RL, Oudejans JJ, Dukers DF, Meijer JW, Ossenkoppele GJ, Meijer CJ. Percentage of activated cytotoxic T-lymphocytes in anaplastic large cell lymphoma and Hodgkin's disease: an independent biological prognostic marker. *Leukemia.* 2001;15(3):458-464.
13. Alvaro T, Lejeune M, Salvado MT, et al. Outcome in Hodgkin's lymphoma can be predicted from the presence of accompanying cytotoxic and regulatory T cells. *Clin Cancer Res.* 2005;11(4):1467-1473.
14. Gloghini A, Canal B, Dal Maso L, Carbone A. Multiple gene expression analyses in human lymphoid tissues by Taqman low-density array using amplified RNA isolated from paraffin-embedded samples. *Diagn Mol Pathol.* 2009;18(3):156-164.
15. Noutsias M, Rohde M, Block A, et al. Preamplification techniques for real-time RT-PCR analyses of endomyocardial biopsies. *BMC Mol Biol.* 2008; 9:3.
16. Ciotti P, Garuti A, Ballestrero A, et al. Reliability and reproducibility of a RNA preamplification method for low-density array analysis from formalin-fixed paraffin-embedded breast cancer samples. *Diagn Mol Pathol.* 2009;18(2):112-118.
17. Li J, Smyth P, Cahill S, et al. Improved RNA quality and TaqMan Pre-amplification method (PreAmp) to enhance expression analysis from formalin fixed paraffin embedded (FFPE) materials. *BMC Biotechnol.* 2008;8:10.
18. Troyanskaya O, Cantor M, Sherlock G, et al. Missing value estimation methods for DNA microarrays. *Bioinformatics.* 2001;17(6):520-525.
19. Solberg HE. Discriminant analysis. *CRC Crit Rev Clin Lab Sci.* 1978;9(3):209-242.
20. Chu F, Wang L. Applications of support vector machines to cancer classification with microarray data. *Int J Neural Syst.* 2005;15(6):475-484.
21. Díaz-Uriarte R, Alibes A, Morrissey ER, Canada A, Rueda OM, Neves ML. Asterias: integrated analysis of expression and aCGH data using an open-source, web-based, parallelized software suite. *Nucleic Acids Res.* 2007;35(Web Server issue):W75-80.
22. Kohlmann M, Held L, Grunert VP. Classification of therapy resistance based on longitudinal biomarker profiles. *Biom J.* 2009;51(4):610-626.
23. Paik S, Shak S, Tang G, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med.* 2004; 351(27):2817-2826.
24. van de Vijver MJ, He YD, van't Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med.* 2002; 347(25):1999-2009.
25. Tzankov A, Meier C, Hirschmann P, Went P, Pileri SA, Dirnhofer S. Correlation of high numbers of intratumoral FOXP3+ regulatory T cells with improved survival in germinal center-like diffuse large B-cell lymphoma, follicular lymphoma and classical Hodgkin's lymphoma. *Haematologica.* 2008;93(2):193-200.
26. Skinnider BF, Mak TW. The role of cytokines in classical Hodgkin lymphoma. *Blood.* 2002;99(12): 4283-4297.
27. Sánchez-Aguilera A, Garcia JF, Sanchez-Beato M, Piris MA. Hodgkin's lymphoma cells express alternatively spliced forms of HDM2 with multiple effects on cell cycle control. *Oncogene.* 2006; 25(18):2565-2574.
28. Bai M, Papoudou-Bai A, Horianopoulos N, Grepi C, Agnantis NJ, Kanavaros P. Expression of bcl2 family proteins and active caspase 3 in classical Hodgkin's lymphomas. *Hum Pathol.* 2007;38(1): 103-113.
29. Dürkop H, Hirsch B, Hahn C, Stein H. cIAP2 is highly expressed in Hodgkin-Reed-Sternberg cells and inhibits apoptosis by interfering with constitutively active caspase-3. *J Mol Med.* 2006; 84(2):132-141.
30. Bai M, Tsanou E, Agnantis NJ, et al. Proliferation profile of classical Hodgkin's lymphomas. Increased expression of the protein cyclin D2 in Hodgkin's and Reed-Sternberg cells. *Mod Pathol.* 2004;17(11):1338-1345.
31. Wolowiec D, Berger F, Ffrench P, Bryon PA, Ffrench M. CDK1 and cyclin A expression is linked to cell proliferation and associated with prognosis in non-Hodgkin's lymphomas. *Leuk Lymphoma.* 1999;35(1-2):147-157.
32. Boss DS, Schwartz GK, Middleton MR, et al. Safety, tolerability, pharmacokinetics and pharmacodynamics of the oral cyclin-dependent kinase inhibitor AZD5438 when administered at intermittent and continuous dosing schedules in patients with advanced solid tumours. *Ann Oncol.* 2010;21(4):884-894.
33. Gilmartin AG, Bleam MR, Richter MC, et al. Distinct concentration-dependent effects of the polo-like kinase 1-specific inhibitor GSK461364A, including differential effect on apoptosis. *Cancer Res.* 2009;69(17):6969-6977.
34. Johnson N, Bentley J, Wang LZ, et al. Pre-clinical evaluation of cyclin-dependent kinase 2 and 1 inhibition in anti-estrogen-sensitive and resistant breast cancer cells. *Br J Cancer.* 2010;102(2): 342-350.
35. Jayanthan A, Howard SC, Trippett T, et al. Targeting the Bcl-2 family of proteins in Hodgkin lymphoma: in vitro cytotoxicity, target modulation and drug combination studies of the Bcl-2 homology 3 mimetic ABT-737. *Leuk Lymphoma.* 2009;50(7): 1174-1182.
36. Steidl C, Lee T, Shah SP, et al. Tumor-associated macrophages and survival in classic Hodgkin's lymphoma. *N Engl J Med.* 2010;362(10):875-885.
37. Devilard E, Bertucci F, Trempat P, et al. Gene expression profiling defines molecular subtypes of classical Hodgkin's disease. *Oncogene.* 2002; 21(19):3095-3102.
38. Alvaro T, Lejeune M, Camacho FI, et al. The presence of STAT1-positive tumor-associated macrophages and their relation to outcome in patients with follicular lymphoma. *Haematologica.* 2006; 91(12):1605-1612.
39. Aldinucci D, Rapana B, Olivo K, et al. IRF4 is modulated by CD40L and by apoptotic and anti-proliferative signals in Hodgkin lymphoma. *Br J Haematol.* 2010;148(1):115-118.
40. Valsami S, Pappa V, Rontogianni D, et al. A clinicopathological study of B-cell differentiation markers and transcription factors in classical Hodgkin's lymphoma: a potential prognostic role of MUM1/IRF4. *Haematologica.* 2007;92(10): 1343-1350.
41. Grumont RJ, Gerondakis S. Rel induces interferon regulatory factor 4 (IRF-4) expression in lymphocytes: modulation of interferon-regulated gene expression by rel/nuclear factor kappaB. *J Exp Med.* 2000;191(8):1281-1292.
42. Sevilla L, Zaldumbide A, Pognonec P, Boulukos KE. Transcriptional regulation of the bcl-× gene encoding the anti-apoptotic Bcl-xL protein by Ets, Rel/NFkappaB, STAT and AP1 transcription factor families. *Histol Histopathol.* 2001;16(2):595-601.
43. Bednarski BK, Baldwin AS Jr, Kim HJ. Addressing reported pro-apoptotic functions of NF-kappaB: targeted inhibition of canonical NF-kappaB enhances the apoptotic effects of doxorubicin. *PLoS ONE.* 2009;4(9):e6992.

miRNA expression in diffuse large B-cell lymphoma treated with chemoimmunotherapy

Santiago Montes-Moreno,[1,2] Nerea Martinez,[1] Beatriz Sanchez-Espiridión,[1] Ramon Díaz Uriarte,[3] Maria Elena Rodriguez,[1] Anabel Saez,[4] Carlos Montalbán,[5] Gonzalo Gomez,[3] David G. Pisano,[3] Juan Fernando García,[6] Eulogio Conde,[7] Eva Gonzalez-Barca,[8] Andres Lopez,[9] Manuela Mollejo,[10] Carlos Grande,[11] Miguel Angel Martinez,[12] Cherie Dunphy,[12] Eric D. Hsi,[13] Gabrielle B. Rocque,[14] Julie Chang,[14] Ronald S. Go,[14] Carlo Visco,[15] Zijun Xu-Monette,[16] Ken H. Young,[16] and Miguel A. Piris[1,2]

[1]Lymphoma Group, Spanish National Cancer Research Centre (CNIO), Madrid, Spain; [2]Pathology Department, Hospital Universitario Marques de Valdecilla, Universidad de Cantabria, IFIMAV, Santander, Spain; [3]Bioinformatics Unit, CNIO, Madrid, Spain; [4]Red de Bancos de Tumores de Andalucía, Granada, Spain; [5]Department of Oncology, Hospital Ramon y Cajal, Madrid, Spain; [6]Department of Pathology, M. D. Anderson Cancer Center, Madrid, Spain (on behalf of the Grupo Español de Linfomas y Transplante Autólogo de Médula Ósea); [7]Department of Haematology, Hospital Universitario Marques de Valdecilla, Universidad de Cantabria, Santander, Spain; [8]Department of Haematology, Institut Catalá d'Óncología, Hospital Duran I Reynals, Barcelona, Spain; [9]Department of Haematology, Hospital Vall d´Hebron, Barcelona, Spain; [10]Department of Pathology, Hospital Virgen de la Salud, Toledo, Spain; [11]Departments of Haematology and Pathology, Hospital Universitario 12 de Octubre, Madrid, Spain; [12]Department of Pathology, University of North Carolina School of Medicine, Chapel Hill, NC; [13]Department of Pathology, Cleveland Clinic, Cleveland, OH; [14]Departments of Medicine, Gundersen Lutheran Health System and University of Wisconsin, WI; [15]Department of Hematology, San Bortolo Hospital, Vicenza, Italy; and [16]Department of Hematopathology, The University of Texas M. D. Anderson Cancer Center, Houston, TX

Diffuse large B-cell lymphoma (DLBCL) prognostication requires additional biologic markers. miRNAs may constitute markers for cancer diagnosis, outcome, or therapy response. In the present study, we analyzed the miRNA expression profile in a retrospective multicenter series of 258 DLBCL patients uniformly treated with chemoimmunotherapy. Findings were correlated with overall survival (OS) and progression-free survival (PFS). miRNA and gene-expression profiles were studied using microarrays in an initial set of 36 cases. A selection of miRNAs associated with either DLBCL molecular subtypes (GCB/ABC) or clinical outcome were studied by multiplex RT-PCR in a test group of 240 cases with available formalin-fixed, paraffin-embedded (FFPE) diagnostic samples. The samples were divided into a training set (123 patients) and used to derive miRNA-based and combined (with IPI score) Cox regression models in an independent validation series (117 patients). Our model based on miRNA expression predicts OS and PFS and improves upon the predictions based on clinical variables. Combined models with IPI score identified a high-risk group of patients with a 2-year OS and a PFS probability of < 50%. In summary, a precise miRNA signature is associated with poor clinical outcome in chemoimmunotherapy-treated DLBCL patients. This information improves upon IPI-based predictions and identifies a subgroup of candidate patients for alternative therapeutic regimens. (*Blood*. 2011; 118(4):1034-1040)

Introduction

Diffuse large B-cell lymphoma (DLBCL) is the most common type of non-Hodgkin lymphoma in adults, accounting for > 80% of aggressive lymphomas.[1] DLBCL is a heterogeneous group of tumors with different genetic abnormalities, clinical features, responses to treatment, and prognosis.[2] This heterogeneity hinders outcome prediction based on clinical and/or molecular parameters.

Combination therapy that associates CHOP (cyclophosphamide, doxorubicin, vincristine, and prednisone) with rituximab (R-CHOP) has become a standard treatment for DLBCL, leading to complete remission rates of 75%-80% and a 3- to 5-year PFS of 50%-60%.[3-8] Nevertheless, patients who fail to respond to first-line therapy or relapse continue to pose a challenge, and identification at diagnosis of poor-outcome cases is crucial for deciding between alternative treatment schemes.

The International Prognostic Index (IPI) has been the primary clinical tool for predicting the outcome of patients with aggressive non-Hodgkin lymphoma.[9] Original IPI factors were redistributed in patients treated with R-CHOP to give a revised score (R-IPI) that distinguishes 3 prognostic categories, with 4-year survival rates ranging from 94%-55% for poor-risk patients.[7] Nevertheless, the R-IPI does not discriminate patients with < 50% probability of survival, which restricts its clinical value.[7]

The biologic heterogeneity of DLBCL has been shown substantially to reflect the cell origin of these tumors from germinal center or activated B cells. These differences are significant independently of IPI stratification, showing that identifying[4,8] cell or origin signatures captures features other than IPI and can refine outcome prediction.[10] These differences between GC and ABC DLBCL remain significant in patients treated with combined immunochemotherapy including rituximab.[11] This classification system can be accurately reproduced using immunohistochemistry against GCET1, CD10, bcl6, MUM1, and FOXP1 in formalin-fixed,

Submitted November 27, 2010; accepted May 13, 2011. Prepublished online as *Blood* First Edition paper, June 1, 2011; DOI 10.1182/blood-2010-11-321554.

Presented in part at the United States and Canadian Academy of Pathology Annual Meeting, March 2009, Boston, MA.[59]

The online version of this article contains a data supplement.

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

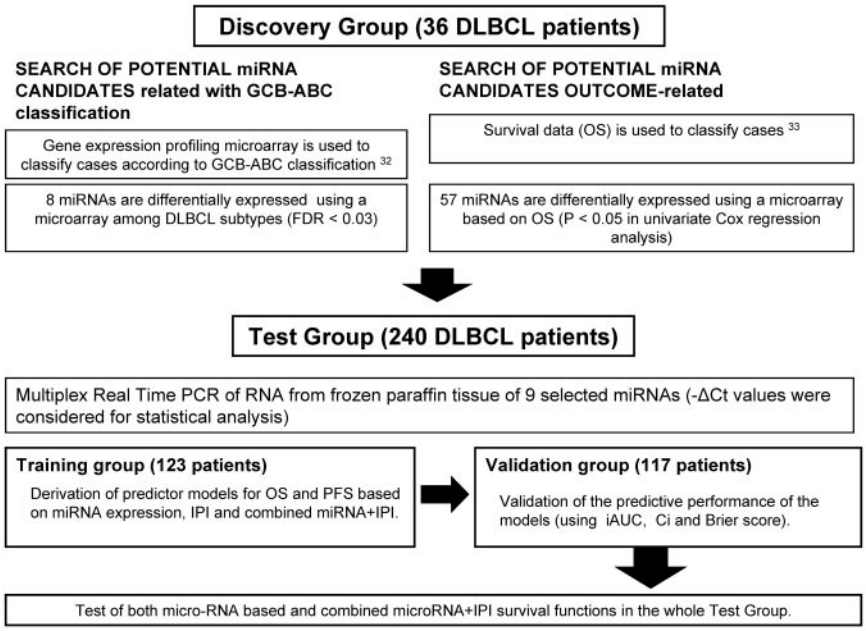


Figure 1. Flowchart with experimental design. The whole series of patients was divided into 2 major groups: a discovery group (36 patients with available frozen tissue for profiling using array technologies) and a test group (240 patients with available FFPE tissue from the diagnostic pathologic sample). The test group was further divided systematically into 2 sets of patients: the training group (123 patients) and the validation group (117 patients). The clinical characteristics according to sex, age, stage, extranodal disease, serum lactate dehydrogenase levels, electrocorticogram, and IPI score of the different sets of patients are summarized in Table 1.

paraffin-embedded (FFPE) samples,[12] thereby providing a simple tool for evaluating the protein-expression profile of the tumor at the time of diagnosis. However, although semiquantitative immunohistochemistry for subclassifying DLBCL is feasible and reproducible, the concordance rates of the different markers vary.[13] Different methods of gene-expression profiling based on quantitative RT-PCR have been developed as alternative or complementary strategies for patient subclassification.[14,15]

Recently, miRNAs, small noncoding RNAs that fine-tune the expression of multiple genes,[16,17] have been shown to be excellent biomarkers for cancer diagnosis and prognosis,[18-21] including hematologic malignancies.[22-27] Furthermore, recent evidence demonstrates that they may constitute markers of differentiation stage, malignant transformation, or sensitivity or resistance to specific drugs.[22,28-31]

The main goal of the present study was to search for an miRNA signature associated with clinical outcome in DLBCL patients treated with R-CHOP using FFPE tissue samples. We also investigated whether miRNA expression profiling can identify particular miRNA species that are differentially expressed between DLBCL subtypes according to the cell of origin (COO) classification.[10,32]

Methods

The experimental procedures are summarized in Figure 1.

Patients and treatments

The study population consisted of a retrospective series of 258 de novo cases of DLBCL obtained from various centers in Spain, 1 in Italy, and 3 in the United States. The study was reviewed and approved as being of minimal/no risk or as exempt by each of the participating institutional review boards, and the overall collaborative study was approved by the institutional review board at the Spanish National Cancer Research Center (CNIO) in Madrid, Spain. The study protocol and sampling methods were approved by the Instituto de Salud Carlos III institutional review board in de-identified anonymous format. Cases associated with HIV or HCV infections or previous immunosuppressive treatments were excluded. All histologic samples corresponded to initial diagnostic biopsies before treatment. Histologic criteria used for diagnosis and classification were those of the World Health Organization.[1] All cases positively stained for CD20. Cases diagnosed as T-cell histiocyte-rich B-cell lymphoma, primary mediastinal B-cell lymphoma cases, cutaneous LBCL, intravascular LBCL, and those histologically associated with a follicular lymphoma component were excluded.

All patients were treated as part of their routine care with standard treatment protocols using a combination of anthracycline-based regimens (6-8 cycles in most cases) and immunotherapy including rituximab; the majority were treated with R-CHOP (n = 243). Other regimens included R-EPOCH and R-MegaCHOP. Responses to treatment were determined by a computed tomography scan in most cases (as recorded in the clinical recovery data sheet) and following the response criteria for lymphoma as defined by Cheson et al.[33]

Array-based expression analysis and in silico prediction of the miRNA regulatory network

miRNA and gene-expression hybridization were carried out using Agilent Technologies microarrays. RNA and DNA extraction methods, details of microarrays and hybridization procedures, and miRNA and gene-expression profiling array normalization are described in supplemental Methods (available on the *Blood* Web site; see the Supplemental Materials link at the top of the online article).

The differential miRNA expression profile between DLBCL subtypes according to the COO signature was studied after GEP-based classification of the cases[32] (for details, see supplemental Methods).

miRNA expression data for all 36 DLBCL cases from the discovery set of patients were examined in a univariate (gene-by-gene) Cox model using SignS.[34] miRNA and gene-expression data have been deposited in the Gene Expression Omnibus[35] and are accessible through GEO series accession number GSE21849 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE21849).

Targets were predicted using available databases (miRBase Version 11.0, MICROCOSM, and TargetScan release 5.1) and a Pearson correlation test based on gene-expression and miRNA-expression data from cases in the discovery set (details and additional references are provided in supplemental Methods).

Real-time PCR for relative miRNA quantification using RNA from FFPE tissue

miRNA expression in FFPE tissues was analyzed using the Applied Biosystems 384-well multiplexed real-time PCR assay with 250 ng of total RNA. Details of the methods, including the selection of endogenous miRNAs, are described in supplemental Methods.

Table 1. Clinical features of the series

		Test set (n = 240 patients)				
	Screening set	Training	*P*	Validation	Relative risk of death, OS (95%CI)	Relative risk of progression, PFS (95% CI)
Number of patients	36	123		117		
IPI factors						
Age, y			.004		2.406 (1.298-4.462); *P* = .005	1.238 (0.758-2.021); *P* = .394
≤ 60	13	61		35		
> 60	23	62		79		
Stage			.5		2.221 (1.213-4.067); *P* = .01	3.287 (1.823-5.927); *P* = .000
I-II	8	55		43		
III-IV	28	68		65		
Lactate dehydrogenase			.112		3.13 (1.53-6.43); *P* = .001	3.051 (1.66-5.608); *P* = .000
low	10	56		28		
high	26	67		55		
Performance status			.87		2.265 (1.306-3.929); *P* = .004	1.956 (1.182-3.238); *P* = .009
Ambulatory (0-1)	29	90		56		
Not ambulatory (2-4)	7	33		22		
Extranodal site involvement			.49		2.012 (1.135-3.565); *P* = .017	2.197 (1.315-3.670); *P* = .003
≤ 1 site	28	96		57		
> 1 site	8	27		21		
IPI score (number of IPI factors)			.067		1.690 (1.378-2.073); *P* = .000	1.479 (1.244-1.758); *P* = .000
Low risk (0,1)	6	46		33		
Low to intermediate risk (2)	10	24		25		
Intermediate to high risk (3)	12	29		26		
High risk (4,5)	8	22		28		

Distribution of IPI factors in the whole series is shown. The χ^2 test was used to assess differences in the proportions of individual prognostic factors between the training and validation sets of patients. Relative risk of event (OS and PFS) estimated from univariate Cox regression is shown for each IPI factor in the test set of patients.

Immunohistochemistry and tissue microarray construction

All cases of DLBCL with available FFPE tissue were histologically reviewed. Representative areas were selected to construct tissue microarrays. Immunohistochemical staining was performed after standard automated protocols using antibodies against CD10, bcl6, MUM-1/IRF4, GCET1, and FOXP1. Immunohistochemical markers were scored on the basis of the cutoffs used by Choi et al[12] (see details in supplemental Methods).

Statistics

The statistical analyses are fully described in supplemental Methods. In brief, outcome-related indices (overall survival [OS] and PFS) were calculated as defined by Cheson et al.[33] Survival distributions were estimated using the Kaplan-Meier method[36] and compared with the log-rank test.[37] The percentage of patients alive at the median follow-up time (and 95% confidence intervals) were noted. The χ^2 test was used to assess differences in the proportions of individual prognostic factors between series.

Cox regression analysis[38] was used to derive 3 independent survival models based on IPI score, miRNA expression (of 9 selected miRNAs), and GC-ABC classification based on immunohistochemistry[12] for both OS and PFS in the training set. More complex models composed of combinations of the 3 individual predictor models were examined. We assessed the improved model fit using χ^2 values to check for significant changes in log likelihood.

Only the combination of miRNAs and IPI score gave a significantly better fit than the individual models for both OS and PFS ($P < .05$ in all comparisons), which justified our fitting a large model with these variables. In this model, significant variables were determined by backward stepwise selection using AUC as the criterion. In this way, we derived definitive combined models for both OS and PFS. These were then validated in the test series using the (integrated) area under the ROC curve, the concordance index, and the Brier score (for details, see supplemental Methods).

Both miRNA-based and combined survival models (functions) based on IPI score and miRNAs after variable selection were applied to the entire test group of patients. For details, see survival functions $h_i(t)$(OS) and $h_i(t)$(PFS) in supplemental Methods.

Models were constructed and validated using the R statistical program, specifically with the packages *survival* (T. Therneau), *pec* (T. Gerds), and *survcomp* (B. Haibe-Kains, C. Sotiriou, and G. Bontempi). Additional analyses were conducted with SPSS Version 15.0.0.

Results

Clinical characteristics of the series

A summary of the clinical characteristics of the entire set of patients used in this study can be found in Table 1. Complete clinical and histopathologic data were available for all 258 patients. The median follow-up time for all patients was 21.3 months. The median follow-up time among patients alive at last follow-up was 27 months (range, 2-105 months). The estimated 2-year OS was 74.7% ± 3% and the estimated PFS was 67.5% ± 3% (supplemental Figure 1). Because the number of events at the median follow-up time comprised 75% (44 of 58) and 92% (66 of 77) of the total number of events during follow-up for OS and PFS, respectively, we considered that the series was suitable for further statistical analysis despite the limited follow-up.

No significant differences were found between the IPI variables in the training and validation groups of patients in the test set (240 patients), with the exception of age. All clinical components of IPI were predictive for OS in the univariate analysis and all but age for PFS (relative risks and confidence intervals for each variable are shown in Table 1, and survival estimates according to IPI in supplemental Figure 1).

Confirmation of the prognostic capacity of COO classification based on immunohistochemistry

Immunohistochemistry was performed in all 240 cases with available FFPE tissue. Most cases (232 of 240) could be classified

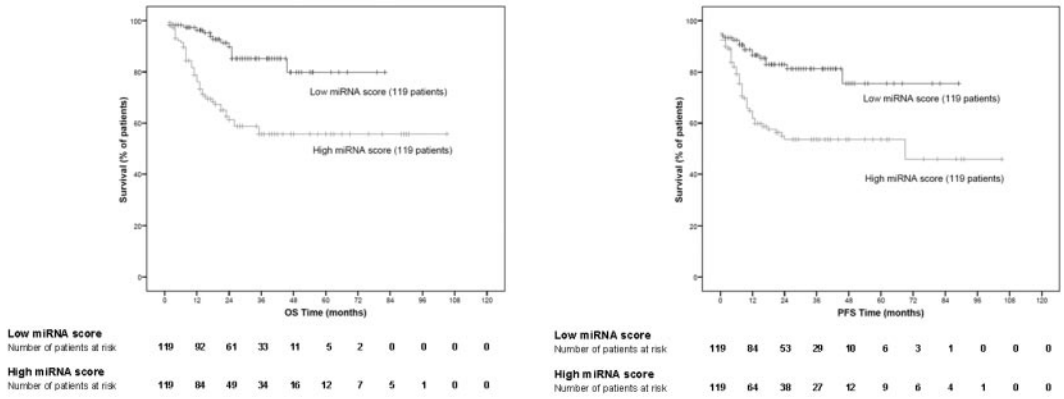


Figure 2. Kaplan-Meier representation of the miRNA-based survival model. miRNA-based survival scores were calculated for each patient in the test group according to the survival function obtained in the training set of patients. After median stratification, Kaplan-Meier estimates were plotted for OS and PFS (log-rank test, $P < .001$ for both OS and PFS).

into GC or ABC subtypes according to previously published algorithms.[12] There were 106 cases classified as the GC type (46%) and 126 as the ABC type (54%). The estimated 2-year OS for ABC-type DLBCL cases was 69.8% ± 4.5%, significantly worse than for GC-type DLCBL patients (81.4% ± 4.3%; $P < .05$). Differences were also found for PFS (60.7% ± 4.7% for the ABC type compared with 75.6% ± 4.6% for the GC type, $P < .05$; supplemental Figure 1).

Identification of a COO miRNA signature.

After gene expression–based classification using the gene set classifier of Wright et al,[32] 11 cases were classified as the GC type and 18 cases as the ABC type. Eight miRNAs were found to be differentially expressed between these subtypes (FDR < 0.03; see supplemental Table 2). These included miR-331, miR-151, miR-28, and miR-454-3p, which were up-regulated in the GC-type DLBCL, whereas miR-222, miR-144, miR-451, and miR-221 were up-regulated in the ABC-type DLBCL. We searched for the putative gene targets of the 8 miRNAs that were differentially expressed among the subtypes. Selected target pairs were miR151-5p and miR28-5p targeting *FOXP1*, miR144 targeting *LRMP1*, and miR451 targeting *MME* (CD10; see supplemental Table 2). Moreover, Gene Set Enrichment Analysis (GSEA Version 2[39]) demonstrated that the GC–B-cell pathway was the main target set of genes that could be modulated by this set of miRNAs (see supplemental Table 2).

Identification of a set of candidate miRNAs associated with outcome in DLBCL

To search for miRNAs related to outcome but not associated with the previously described COO signatures, miRNA expression data for all 36 DLBCL cases from the discovery set of patients were subjected to a univariate (gene-by-gene) Cox analysis after FCMS using SignS.[34] Fifty-seven human miRNAs were correlated positively or negatively with OS ($P < .05$; see supplemental materials). None of these 57 miRNAs formed part of the COO signature, suggesting that this method may add some complementary information to the previous approach.

After the previously described procedures, a final set of 9 miRNAs was subjected to relative RT-PCR quantification in the entire test group of 240 patients for whom FFPE tissue was available (see details in supplemental Methods). Seven of these miRNAs (miR-221, miR-222, miR-331, miR-451, miR-28, miR-151, and miR-148a) were identified using the COO-signature approach and 2 additional miRNAs (miR-93 and miR-491) were obtained from the univariate (gene-by-gene) Cox analysis using SignS.[34]

Generation of a miRNA-based predictor model using FFPE tissue samples

After Cox regression analysis,[38] 3 independent survival models based on IPI score, miRNA expression (of 9 selected miRNAs), and GC-ABC classification based on immunohistochemistry[12] were derived for OS and PFS in the training set.

miRNA expression-based models using the expression of 9 miRNAs as a continuous variable were able to predict both OS and PFS. Their predictive performance in the validation group of patients was confirmed by 3 different statistical methods including the (integrated) area under the ROC curve, the concordance index, and the Brier score (for details, see supplemental Methods). When evaluating the Brier score, we used a conditional weighting scheme from a Cox model of the censoring distribution,[40] including all miRNAs and the IPI score as predictors. The results for all tests are shown in supplemental Tables 2 and 3 and in supplemental Figure 2.

After confirming the predictive performance in the validation set of patients, we plotted Kaplan-Meier curves using the whole test group (training and validation sets) of patients. After median stratification of the continuous score obtained from the miRNA-based models, all patients from the test group were classified as having either a low miRNA score (below median) or a high miRNA score (above median). Significant differences in OS and PFS were found between the 2 groups of patients (log-rank test, $P < .001$ for both; Figure 2).

Only the combination of miRNAs and IPI score was significantly better than the individual models for both OS and PFS (χ^2 test for the change in log likelihood, $P < .05$ for all comparisons), justifying our fitting a large model with these variables. The model was derived by backward stepwise selection using area under the curve as the criterion. This yielded combined models for both OS and PFS (details of survival functions can be found in the supplemental Methods). Because the construction of the combined models involves a multivariate analysis (variable selection step), the prediction based on the expression of the miRNAs present in the combined models was independent of the IPI score (the miRNAs included in these combined models together with IPI score and its associated hazard ratios are shown in Figure 3).

Kaplan-Meier estimates for OS and PFS in all test group series using the combined models were calculated and plotted according to the distribution of terciles (Figure 4). A high-risk subgroup of patients with OS and PFS below 50% after a 2-year follow-up was identified.

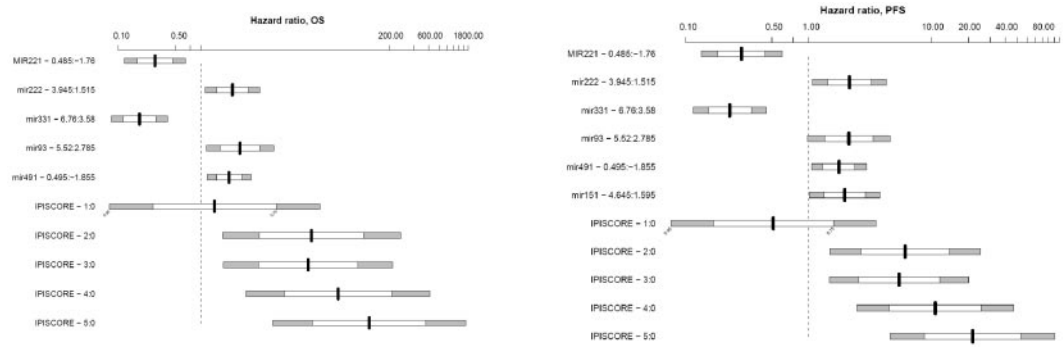


Figure 3. Hazard ratio charts. The interquartile range of the hazard ratios estimated from the models after variable selection is shown for each continuous variable. For each miRNA, we calculated the log of the hazard ratio, where the difference in that variable was that of the interquartile range (ie, the third to the first quartiles). For example, for miR-222, the first and the third quartiles are 1.515 and 3.945; the bar shows the hazard ratio. For PFS: exp(0.315 * [3.945-1.515] = 2.15) and a 75% and 95% interval. For IPISCORE, a discrete variable, we show the log hazard ratio comparing each of the values of IPISCORE (except 0) with IPISCORE = 0 as a reference.

The miRNA-regulatory network

To gain a deeper insight into the network of genes that might be regulated by the set of outcome predictor miRNAs, we interrogated the current versions of the MiRBase and TargetScan miRNA target prediction databases. The predicted pairs we found included miR-222–CDKN1A (p21), miR-93–BCL66, miR-93–MCL1, and miR-93–MAP3K14 (NIK), among others.

We also examined the correlation between gene expression and miRNA expression in the data from the 29 samples of the discovery set. Because down-regulation of the target mRNA[41,42] is considered the main mechanism by which miRNAs modulate protein expression, significant negative correlation pairs were identified, including miR-331–CARD10, miR-331–IRF4, miR-331–PIM2, and miR-331–AID. The complete correlation grid obtained is shown in supplemental Table 2. Interestingly, the absence of any significant negative correlation between many in silico–predicted pairs of miRNAs and mRNAs in this test suggests that additional mechanisms to mRNA down-regulation, such as translation inhibition, may explain the protein-expression modulation afforded by miRNAs.[41,43,44]

Discussion

miRNAs are emergent biomarkers of disease that have proved useful for cancer diagnosis and prognosis[18-21] and in hematologic malignancies.[22-27] They have been demonstrated to reflect accurately the differentiation stage of human lymphoid B cells,[22,27,28] providing valuable information for tumor classification[19,28] and prognostication[26,27] that may be added to that available from gene-expression and clinical data.

In the present study, we analyzed miRNA expression profiles in a series of homogeneously treated DLBCL patients, deriving miRNA-based models that correlate with OS and PFS, but which are independent of IPI. We used a 2-step approach to identify a candidate set of miRNAs and then validated this with multiplex RT-PCR using RNA from FFPE tissue in 2 independent sets of patients. This method (real-time PCR) has been found to give reliable measures of gene and miRNA expression[14,15] that are alternatives to the classic semiquantitative immunohistochemical approaches.[13] Furthermore, the technique can be routinely applied to FFPE samples, because the small size of miRNAs makes them relatively resistant to RNAse degradation and because they can be successfully isolated from routinely processed FFPE tissue.[26]

After the first step of candidate identification, we found a signature of miRNAs related to the differentiation stage as defined by the COO signature.[32] Specifically, we found miR-331, miR-151, miR-28, and miR-454-3p to be up-regulated in the GC-type DLBCL. Conversely, miR-222, miR-144, miR-451, and miR-221 were up-regulated in the ABC-type DLBCL. Our data from patient tissue samples classified according to their gene-expression profile

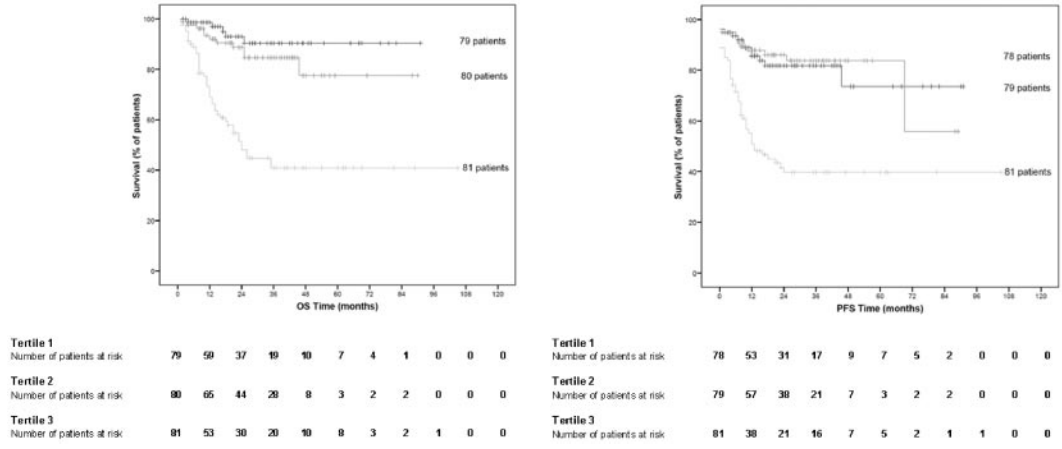


Figure 4. Kaplan-Meier representation of combined survival model based on IPI score and miRNAs. Survival scores according to IPI and miRNAs were calculated for each patient in the test group according to the combined survival function obtained in the training set of patients (see "Generation of a miRNA-based predictor model using FFPE tissue samples"). After tercile stratification, Kaplan-Meier estimates were plotted for OS and PFS (log-rank test, $P < .001$ for both OS and PFS). Hazard ratios estimated from the models are shown in Figure 3.

(using the COO classifier) confirm those from other studies, which found miR-222 and miR221 to be up-regulated in the well-known ABC-DLCL cell lines OCI-Ly-10 and OCI-Ly-3.[26,27,29,45] Furthermore, our in silico prediction indicates that genes essential as markers of the COO subtype (mainly GCB genes) are putative targets of certain differentially expressed miRNAs.

When we compared the classification of patients according to the miRNA model and the immunochemically based GC-ABC classification, the systems were shown to be nonoverlapping and complementary, thereby establishing different subsets of cases ($P < .001$, supplemental Table 4 sheet D). These results are particularly interesting because there is not complete agreement about the best method for stratifying patients with DLBCL. Different immunohistochemical algorithms are currently being tested with various results among series of patients, including the one presented here.[46,47]

Our results also identify a set of miRNAs in which expression is associated with poor clinical outcome in R-CHOP–treated DLBCL patients. miRNA expression–based models using the expression of miRNAs as a continuous variable were able to predict both OS and PFS in 2 independent sets of patients. These models predict OS and PFS and improve IPI-based predictions, which allowed us to generate combined models identifying a high-risk subgroup of patients with OS and PFS below 50% after a 2-year follow-up.

This particular signature, which contains some miRNAs previously shown to be correlated with the outcome of DLBCL patients[27,29] and other hematologic malignancies,[23] includes miRNAs that target pathways commonly deregulated in DLBCL. These include, for example, genes related to apoptosis (*MCL1*), the cell cycle (*CDKN1A*), MAPK and NFkB signaling (*MAP3K14*, *MUM1/IRF4*, *CARD10*, and *PIM2*), somatic mutation during the GC reaction (*AID*), and key transcription factors such as *BCL6*. Specifically, miR-221 and miR-222 have been found to be essential growth-regulatory mediators inhibiting *p27* (Kip1), a cell-cycle inhibitor and tumor suppressor.[48-51] Other components of the signature, such as the miR-106b-25 cluster (including miR-93), have recently been found to interfere with the expression of *CDKN1A* (p21) and *BCL2L11*, thereby impairing the TGFB tumor-suppressor pathway in other cancers.[52] Functional experiments will be required to confirm the candidate interactions identified here. Transfection experiments (introducing the miRNA and measuring changes in the target mRNA/protein) and silencing experiments (using shRNAs to inhibit the constitutive expression of the miRNAs) might both be performed to address this matter, but this was beyond the scope of the present study.

It is remarkable that the miRNA signature identified here predicted survival independently of the COO classification. Although the GC-ABC subclassification has a potentially predictive role in the identification of patients likely to respond to specific therapies for ABC-type DLBCL,[53] the development of new methods that can capture a set of DLBCL cases with particularly aggressive behavior paves the way for the design of trials using alternative therapeutic strategies (eg, untargeted therapies such as stem cell transplantation or more refined targeted therapies against the substrate genes/pathways involved).

The recent observation that *c-MYC* rearrangements in DLBCL are associated with poor prognosis in a subset (5%-15%) of rituximab-treated patients[54,55] led us to consider the possible relationship between *MYC* status and the miRNA expression classifier. The biologic basis for such an association might reside in the role of *MYC* as a transcriptional regulator of the expression of some miRNAs.[56-58] None of the miRNAs included in the prognostic signatures described here has been found to be associated with *MYC* in a wide range of functional studies in lymphoid cell lines and lymphoma animal models.[56-58] However, because of the possible combinatorial effect of both *MYC* rearrangements and miRNA deregulation, additional studies on the combination of *MYC* and miRNA predictive impact are warranted. Our data identified a set of miRNAs that could be useful outcome prognostic markers in DLBCL treated with R-CHOP, and an integrated model with IPI identified a subset of high-risk patients with a 2-year OS $< 50\%$. Therefore, our approach may serve to refine outcome prediction and to assign a risk-stratified therapy for DLBCL patients.

Acknowledgments

The authors thank María Encarnación Castillo and CNIO's Tumor Bank Unit (María Jesús Artiga, Laura Cereceda, and Manuel Morente) for their skillful retrieval and handling of clinical data and samples from different clinical institutions and Lorena di Lisio for fruitful discussions on methodologic issues. They also acknowledge all of the clinical colleagues who kindly completed the clinical data form, particularly M. Canales, F. Mazorra, M. Cruz, J. Menarguez, and T. Gerds for discussions about Brier scores.

This study was supported by grants from the Ministerio de Sanidad y Consumo (PI051623, PI052800, CP06/00002, RTICC), the Asociación Española contra el Cancer (AECC), and the Ministerio de Ciencia e Innovación Spain (SAF 2008-03871).

Authorship

Contribution: S.M.-M. designed and performed the research, analyzed the data, and wrote the manuscript; N.M. and B.S.-E. performed the research and analyzed the data; A.S., C.M., E.G.-B., A.L., M.M., C.G., and M.A.M. contributed clinical data; M.E.R. performed the research; R.D.U., D.G.P., and G.G. analyzed the data; J.F.G. contributed clinical data and performed the research; C.D., E.D.H., E.C., C.V., and R.S.G. contributed clinical data and skills; G.B.R., Z.X.-M., and J.C. performed the data analysis; K.H.Y. analyzed the data and critically revised and drafted the manuscript; and M.A.P. designed the research, analyzed the data, and wrote the manuscript.

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

Correspondence: Miguel A. Piris, Pathology Department, Hospital Universitario Margués de Valdecilla, IFIMAV, Avenida de Valdecilla n 25, 39008 Santander, Spain; e-mail: mapiris@cnio.es.

References

1. Swerdlow SH CE, Harris NL, Jaffe ES, et al. *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues.* 4th ed. Lyon, France: IARC Press; 2008.
2. Lossos IS, Morgensztern D. Prognostic biomarkers in diffuse large B-cell lymphoma. *J Clin Oncol.* 2006;24(6):995-1007.
3. Sehn LH, Donaldson J, Chhanabhai M, et al. Introduction of combined CHOP plus rituximab therapy dramatically improved outcome of diffuse large B-cell lymphoma in British Columbia. *J Clin Oncol.* 2005;23(22):5027-5033.
4. Habermann TM, Weller EA, Morrison VA, et al. Rituximab-CHOP versus CHOP alone or with maintenance rituximab in older patients with diffuse large B-cell lymphoma. *J Clin Oncol.* 2006;24(19):3121-3127.
5. Feugier P, Van Hoof A, Sebban C, et al. Long-term results of the R-CHOP study in the treatment of elderly patients with diffuse large B-cell lymphoma: a study by the Groupe d'Etude des

Lymphomes de l'Adulte. *J Clin Oncol.* 2005; 23(18):4117-4126.

6. Coiffier B, Lepage E, Briere J, et al. CHOP chemotherapy plus rituximab compared with CHOP alone in elderly patients with diffuse large-B-cell lymphoma. *N Engl J Med.* 2002;346(4):235-242.
7. Sehn LH, Berry B, Chhanabhai M, et al. The revised International Prognostic Index (R-IPI) is a better predictor of outcome than the standard IPI for patients with diffuse large B-cell lymphoma treated with R-CHOP. *Blood.* 2007;109(5):1857-1861.
8. Pfreundschuh M, Trumper L, Osterborg A, et al. CHOP-like chemotherapy plus rituximab versus CHOP-like chemotherapy alone in young patients with good-prognosis diffuse large-B-cell lymphoma: a randomised controlled trial by the MabThera International Trial (MInT) Group. *Lancet Oncol.* 2006;7(5):379-391.
9. A predictive model for aggressive non-Hodgkin's lymphoma. The International Non-Hodgkin's Lymphoma Prognostic Factors Project. *N Engl J Med.* 1993;329(14):987-994.
10. Alizadeh AA, Eisen MB, Davis RE, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature.* 2000; 403(6769):503-511.
11. Lenz G, Wright G, Dave SS, et al. Stromal gene signatures in large-B-cell lymphomas. *N Engl J Med.* 2008;359(22):2313-2323.
12. Choi WW, Weisenburger DD, Greiner TC, et al. A new immunostain algorithm classifies diffuse large B-cell lymphoma into molecular subtypes with high accuracy. *Clin Cancer Res.* 2009; 15(17):5494-5502.
13. de Jong D, Rosenwald A, Chhanabhai M, et al. Immunohistochemical prognostic markers in diffuse large B-cell lymphoma: validation of tissue microarray as a prerequisite for broad clinical applications—a study from the Lunenburg Lymphoma Biomarker Consortium. *J Clin Oncol.* 2007;25(7):805-812.
14. Malumbres R, Chen J, Tibshirani R, et al. Paraffin-based 6-gene model predicts outcome in diffuse large B-cell lymphoma patients treated with R-CHOP. *Blood.* 2008;111(12):5509-5514.
15. Alizadeh AA, Gentles AJ, Lossos IS, Levy R. Molecular outcome prediction in diffuse large-B-cell lymphoma. *N Engl J Med.* 2009;360(26):2794-2795.
16. Ambros V. The functions of animal microRNAs. *Nature.* 2004;431(7006):350-355.
17. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell.* 2004;116(2):281-297.
18. Ji J, Shi J, Budhu A, et al. MicroRNA expression, survival, and response to interferon in liver cancer. *N Engl J Med.* 2009;361(15):1437-1447.
19. Lu J, Getz G, Miska EA, et al. MicroRNA expression profiles classify human cancers. *Nature.* 2005;435(7043):834-838.
20. Yanaihara N, Caplen N, Bowman E, et al. Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. *Cancer Cell.* 2006;9(3):189-198.
21. Schetter AJ, Leung SY, Sohn JJ, et al. MicroRNA expression profiles associated with prognosis and therapeutic outcome in colon adenocarcinoma. *JAMA.* 2008;299(4):425-436.
22. Basso K, Sumazin P, Morozov P, et al. Identification of the human mature B cell miRNome. *Immunity.* 2009;30(5):744-752.
23. 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(17):1793-1801.
24. Garzon R, Volinia S, Liu CG, et al. MicroRNA signatures associated with cytogenetics and prognosis in acute myeloid leukemia. *Blood.* 2008; 111(6):3183-3189.
25. Marcucci G, Radmacher MD, Maharry K, et al. MicroRNA expression in cytogenetically normal acute myeloid leukemia. *N Engl J Med.* 2008; 358(18):1919-1928.
26. Lawrie CH, Soneji S, Marafioti T, et al. MicroRNA expression distinguishes between germinal center B cell-like and activated B cell-like subtypes of diffuse large B cell lymphoma. *Int J Cancer.* 2007; 121(5):1156-1161.
27. Malumbres R, Sarosiek KA, Cubedo E, et al. Differentiation stage-specific expression of microRNAs in B lymphocytes and diffuse large B-cell lymphomas. *Blood.* 2009;113(6):3754-3764.
28. Zhang J, Jima DD, Jacobs C, et al. Patterns of microRNA expression characterize stages of human B-cell differentiation. *Blood.* 2009;113(19): 4586-4594.
29. Lawrie CH, Chi J, Taylor S, et al. Expression of microRNAs in diffuse large B cell lymphoma is associated with immunophenotype, survival and transformation from follicular lymphoma. *J Cell Mol Med.* 2009;13(7):1248-1260.
30. Kovalchuk O, Filkowski J, Meservy J, et al. Involvement of microRNA-451 in resistance of the MCF-7 breast cancer cells to chemotherapeutic drug doxorubicin. *Mol Cancer Ther.* 2008;7(7): 2152-2159.
31. Zhu H, Wu H, Liu X, et al. Role of MicroRNA miR-27a and miR-451 in the regulation of MDR1/P-glycoprotein expression in human cancer cells. *Biochem Pharmacol.* 2008;76(5):582-588.
32. Wright G, Tan B, Rosenwald A, Hurt EH, Wiestner A, Staudt LM. A gene expression-based method to diagnose clinically distinct subgroups of diffuse large B cell lymphoma. *Proc Natl Acad Sci U S A.* 2003;100(17):9991-9996.
33. Cheson BD, Pfistner B, Juweid ME, et al. Revised response criteria for malignant lymphoma. *J Clin Oncol.* 2007;25(5):579-586.
34. Diaz-Uriarte R. SignS: a parallelized, open-source, freely available, web-based tool for gene selection and molecular signatures for survival and censored data. *BMC Bioinformatics.* 2008;9: 30.
35. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 2002;30(1):207-210.
36. Kaplan E, Meier P. Nonparametric estimation from incomplete observations. *Am Stat Assoc J.* 1958;53:457-81.
37. Peto R, Pike MC, Armitage P, et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design. *Br J Cancer.* 1976;34(6):585-612.
38. Cox D. Regression models and life-tables. *J R Stat Soc.* 1972;34:187-220.
39. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A.* 2005;102(43):15545-15550.
40. Gerds TA, Schumacher M. Consistent estimation of the expected Brier score in general survival models with right-censored event times. *Biom J.* 2006;48(6):1029-1040.
41. Xiao C, Rajewsky K. MicroRNA control in the immune system: basic principles. *Cell.* 2009;136(1): 26-36.
42. Kim DH, Saetrom P, Snove O Jr, Rossi JJ. MicroRNA-directed transcriptional gene silencing in mammalian cells. *Proc Natl Acad Sci U S A.* 2008;105(42):16230-16235.
43. Baek D, Villen J, Shin C, Camargo FD, Gygi SP, Bartel DP. The impact of microRNAs on protein output. *Nature.* 2008;455(7209):64-71.
44. Selbach M, Schwanhausser B, Thierfelder N, Fang Z, Khanin R, Rajewsky N. Widespread changes in protein synthesis induced by microRNAs. *Nature.* 2008;455(7209):58-63.
45. Roehle A, Hoefig KP, Repsilber D, et al. MicroRNA signatures characterize diffuse large B-cell lymphomas and follicular lymphomas. *Br J Haematol.* 2008;142(5):732-744.
46. Meyer PN, Fu K, Greiner TC, et al. Immunohistochemical methods for predicting cell of origin and survival in patients with diffuse large B-cell lymphoma treated with rituximab. *J Clin Oncol.* 2011; 29(2):200-207.
47. Ott G, Ziepert M, Klapper W, et al. Immunoblastic morphology but not the immunohistochemical GCB/non-GCB classifier predicts outcome in diffuse large B-cell lymphoma in the RICOVER-60 trial of the DSHNHL. *Blood.* 2010;116(23):4916-4925.
48. le Sage C, Nagel R, Egan DA, et al. Regulation of the p27(Kip1) tumor suppressor by miR-221 and miR-222 promotes cancer cell proliferation. *EMBO J.* 2007;26(15):3699-3708.
49. Visone R, Russo L, Pallante P, et al. MicroRNAs (miR)-221 and miR-222, both overexpressed in human thyroid papillary carcinomas, regulate p27Kip1 protein levels and cell cycle. *Endocr Relat Cancer.* 2007;14(3):791-798.
50. Medina R, Zaidi SK, Liu CG, et al. MicroRNAs 221 and 222 bypass quiescence and compromise cell survival. *Cancer Res.* 2008;68(8):2773-2780.
51. Fornari F, Gramantieri L, Ferracin M, et al. MiR-221 controls CDKN1C/p57 and CDKN1B/p27 expression in human hepatocellular carcinoma. *Oncogene.* 2008;27(43):5651-5661.
52. Petrocca F, Visone R, Onelli MR, et al. E2F1-regulated microRNAs impair TGFbeta-dependent cell-cycle arrest and apoptosis in gastric cancer. *Cancer Cell.* 2008;13(3):272-286.
53. Dunleavy K, Pittaluga S, Czuczman MS, et al. Differential efficacy of bortezomib plus chemotherapy within molecular subtypes of diffuse large B-cell lymphoma. *Blood.* 2009;113(24):6069-6076.
54. Barrans S, Crouch S, Smith A, et al. Rearrangement of MYC is associated with poor prognosis in patients with diffuse large B-cell lymphoma treated in the era of rituximab. *J Clin Oncol.* 2010; 28(20):3360-3365.
55. Savage KJ, Johnson NA, Ben-Neriah S, et al. MYC gene rearrangements are associated with a poor prognosis in diffuse large B-cell lymphoma patients treated with R-CHOP chemotherapy. *Blood.* 2009;114(17):3533-3537.
56. O'Donnell KA, Wentzel EA, Zeller KI, Dang CV, Mendell JT. c-Myc-regulated microRNAs modulate E2F1 expression. *Nature.* 2005;435(7043): 839-843.
57. Sander S, Bullinger L, Klapproth K, et al. MYC stimulates EZH2 expression by repression of its negative regulator miR-26a. *Blood.* 2008;112(10): 4202-4212.
58. Chang TC, Yu D, Lee YS, et al. Widespread microRNA repression by Myc contributes to tumorigenesis. *Nat Genet.* 2008;40(1):43-50.
59. Montes-Moreno S, Martinez N, Saez A, et al. MicroRNA expression in DLBCL: identification of unique miRNAs signatures related to clinical outcome prediction in RCHOP treated patients. *Mod Pathol.* 2009;22(supp 1):278A(1262).

Human Pathology (2011) **xx**, xxx–xxx

ELSEVIER

Human PATHOLOGY

www.elsevier.com/locate/humpath

Original contribution

Glioma-associated oncogene homologue 3, a hedgehog transcription factor, is highly expressed in Hodgkin and Reed-Sternberg cells of classical Hodgkin lymphoma☆,☆☆

Wesley O. Greaves MD[a,1], Ji Eun Kim MD[b,1], Rajesh R. Singh PhD[a], Elias Drakos MD, PhD[a], Kranthi Kunkalla MS[a], Beatriz Sánchez-Espiridión MS[c], Juan F. Garcia MD[d], L. Jeffrey Medeiros MD[a], Francisco Vega MD, PhD[a,*]

[a]*Department of Hematopathology, The University of Texas M.D. Anderson Cancer Center, Houston, TX, 77030, USA*
[b]*Department of Pathology, Seoul National University Boramae Hospital, 156-707 Seoul, South Korea*
[c]*Spanish National Cancer Research Centre (CNIO), Madrid, 28029 Spain*
[d]*Centro Oncológico MD Anderson International España, Madrid, 28033 Spain*

Received 23 September 2010; revised 22 December 2010; accepted 23 December 2010

Keywords:
GLI3;
Hedgehog signaling;
Classical Hodgkin lymphoma;
Immunohistochemistry

Summary The hedgehog signaling pathway has been shown to play a pathogenic role in diffuse large B-cell lymphoma and anaplastic large cell lymphoma, but has not been assessed in classical Hodgkin lymphoma. Glioma-associated oncogene homologues 1, 2, and 3 are transcriptional effectors of the hedgehog pathway. In this study, we first used real-time quantitative polymerase chain reaction to investigate the expressions of *GLI1*, *GLI2*, and *GLI3* in 3 classical Hodgkin lymphoma cell lines. *GLI1* and *GLI2* were variably expressed, but *GLI3* was highly expressed in all cell lines. We then used immunohistochemistry to assess glioma-associated oncogene homologues 1, 2, and 3 in 39 classical Hodgkin lymphoma patient samples. Glioma-associated oncogene homologues 1 and 2 were weakly to variably expressed in a subset of classical Hodgkin lymphoma patient samples. In contrast, glioma-associated oncogene homologue 3 showed strong, uniform nuclear expression in virtually all Hodgkin/Reed-Stenberg cells. We then performed an immunohistochemical survey of glioma-associated oncogene homologue 3 expression in 13 cases of nodular lymphocyte predominant Hodgkin lymphoma and 218 non–Hodgkin lymphomas. Most other lymphoma types showed variable or no expression of glioma-associated oncogene homologue 3, with a minor subset of cases of nodular lymphocyte predominant Hodgkin lymphoma, ALK-positive and ALK-negative anaplastic large cell lymphoma, and B-cell lymphoma, unclassifiable with features intermediate between diffuse large B-cell lymphoma and classical Hodgkin lymphoma showing a glioma-associated oncogene homologue 3 staining pattern indistinguish-

☆ This research was supported by funds from the University Cancer Foundation at The University of Texas M.D. Anderson Cancer Center, from a translational grant of The Leukemia & Lymphoma Society, and the K08 Physician-Scientist Award 1 K08 CA143151-01 (NIH). This paper was presented in part at the United States and Canadian Academy of Pathology (USCAP) Annual Meeting, Washington, March 2010.

☆☆ Disclosures/conflicts of interest: None to declare.

* Corresponding author.
E-mail address: fvegava@mdanderson.org (F. Vega).

[1] Wesley Greaves and Ji Eun Kim contributed equally to this work.

doi:10.1016/j.humpath.2010.12.023

able from classical Hodgkin lymphoma. Our data provide a rationale to further investigate the biologic significance of glioma-associated oncogene homologue 3 in classical Hodgkin lymphoma biology.

1. Introduction

Hodgkin lymphoma (HL) represents approximately 30% of all lymphomas and can be divided into classical HL (CHL) and nodular lymphocyte predominant (NLPHL) types, representing 95% and 5% of cases, respectively [1]. CHL can be further divided into 4 histologic types: nodular sclerosis, lymphocyte-rich, lymphocyte-depleted, and mixed cellularity. The pathogenesis of CHL is not well defined, and recent studies have shown that constitutive activation of multiple signaling pathways contributes to the survival and proliferation of Hodgkin/Reed-Sternberg (HRS) cells. In particular, constitutive activation of the nuclear factor κB (NFκB), PI3K/AKT, and JAK/STAT pathways have been implicated [2]. HRS cells also commonly express Bcl-2 protein, suggesting that evasion of apoptotic mechanisms plays a role in HRS cell survival [3,4]. In addition, there is evidence that the microenvironment is essential for survival of HRS cells. It is difficult to grow HRS cells in culture or to generate cell lines, and it is rare to find circulating HRS in peripheral blood. When CHL does disseminate to nonlymphoid organs, HRS cells are accompanied by a typical microenvironment [2]. It seems that the composition of the microenvironment is regulated by a range of cytokines and chemokines produced by HRS cells and the microenvironment [2].

The Hedgehog (HH) protein family is a group of secreted signaling molecules critical for normal mammalian development and homeostasis of adult renewable tissues [5,6]. Recent studies have shown that secreted HH ligands protect nonneoplastic germinal center B cells as well as of low-grade B-cell lymphoma cells from apoptosis [7,8]. The HH signaling pathway is incompletely understood. Briefly, the HH ligands, Sonic, Indian, and Desert, interact with a membrane-bound receptor complex composed of 2 proteins, patched (PTCH) and smoothened (SMO) [9]. PTCH is the HH ligand-binding subunit, and SMO is the signal transduction component. In the absence of HH ligands, PTCH inhibits SMO. Once HH ligands bind to PTCH, this inhibition is released, allowing SMO to transduce the HH signal mediated by the transcription factors glioma-associated oncogene homologues (GLI) 1, 2 and 3 [10]. Although both GLI2 and GLI3 have transcriptional activation and repression properties, GLI1 is a strong positive regulator of HH transcriptional targets and is, itself, a transcriptional target of HH [11]. GLI3 is a bifunctional transcriptional regulator that, in the absence of HH, is proteolytically cleaved to form the truncated GLI3 protein, which acts as a repressor of target gene transcription (GLI3R) [12]. On the other hand, in the presence of HH signals, the full-length form of GLI3 acts a transcriptional activator (GLI3A). The functions of GLI3 in the hematopoietic system are not well characterized.

Inappropriate activation of the HH pathway has been shown in many cancers [13,14] but has not been assessed in CHL. In this study, we investigated the expression of HH transcription factors in a series of CHL patient samples. We found that GLI3 is highly expressed in CHL and shows a distinctive immunohistochemical expression pattern in HRS cells. Expression of GLI3 protein by HRS cells provides a rationale for further investigation of the biologic role of GLI3 in CHL. Furthermore, as GLI3 is expressed by thymic stroma and is described to regulate T-cell differentiation, GLI3 may be involved in regulating the composition of the microenvironment in CHL.

2. Materials and methods

2.1. Cell lines

The CHL-derived cell lines used included KMH2 and HDMYZ (purchased from DSMZ, Braunschweig, Germany) and L428. We also used a DLBCL-derived cell line, BJAB (purchased from American Type Culture Collection [ATCC], Manassas, VA). All cell lines were maintained at 37°C in RPMI 1640 (ATCC) supplemented with 10% heat-inactivated fetal bovine serum (Sigma, St Louis, MO) in a humidified atmosphere containing 5% CO_2.

2.2. Real-time quantitative reverse transcriptase polymerase chain reaction

Total messenger RNA (mRNA) from cell lines was extracted using RNAeasy Minikit (Qiagen, Valencia, CA) as

Table 1 Antibodies used for immunohistochemistry

Antibody	Clone	Manufacturer	Dilution	Incubation
HH	H-160	Santa Cruz Biotechnology, Santa Cruz, CA	1:150	RT, 1 h
GLI1	H-300	Santa Cruz	1:200	RT, 1 h
GLI2	NA	Abcam, Cambridge, MA	1:200	RT, 1.5 h
GLI3	H-280	Santa Cruz	1:150	RT, 1 h
CD20	L26	Dako	1:1000	RT, 1 h
CD30	Ab36749	Abcam	1:150	RT, 1.5 h

Abbreviations: NA, not available; RT, room temperature.

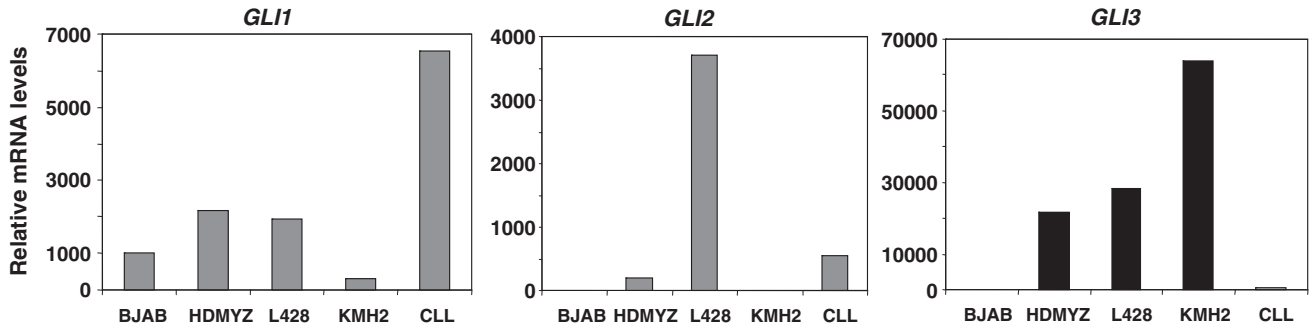


Fig. 1 RT-qPCR analysis for *GLI* molecules in CHL. RT-qPCR analysis in CHL cells shows the presence of mRNAs for *GLI1*, *GLI2*, and *GLI3*. Overall, there is variable expression of *GLI1* and *GLI2* mRNAs in the 3 CHL cell lines tested. However, consistently elevated expression of *GLI3* mRNA is seen in all 3 CHL cell lines. BJAB, a DLBCL cell line, and 1 CLL patient samples are used as control references.

per the manufacturer's protocol. After DNAse digestion, first-strand complement DNA was synthesized from total RNA using random primers and SuperScript II Reverse Transcriptase (Invitrogen, Carlsbad, CA). Gene expression of *GLI3* was analyzed by quantitative reverse transcriptase polymerase chain reaction (RT-qPCR) assay using respective primers (Applied Biosystems, Foster City, CA) and Taqman probes. Reactions were performed using the ABI PRISM 7500HT Sequence Detection system (Applied Biosystems), and the expression of each gene was measured in duplicate and then normalized with 18s ribosomal RNA endogenous gene. Relative quantification of gene expression levels was assessed by using the delta cycle threshold (CT) method. BJAB cells were used as a negative reference control.

2.3. Patients samples and immunohistochemistry

The study group included 33 cases of nodular sclerosis and 6 cases of mixed cellularity CHL. Either routine histologic sections or tissue microarrays were used. We then performed a survey of 13 cases of NLPHL and 218 non-HLs as follows: 54 DLBCL, not otherwise specified (NOS); 28 follicular lymphoma (FL); 21 anaplastic lymphoma kinase (ALK)-positive anaplastic large cell lymphoma (ALCL); 16 ALK-negative ALCL; 16 chronic lymphocytic lymphoma/small lymphocytic lymphoma (CLL/SLL); 13 B-cell lymphoma, unclassifiable, with features intermediate between DLBCL and CHL (BCLU); 12 T-cell/histiocyte-rich large B-cell lymphoma (TCHRLBCL); 12 Burkitt lymphoma (BL); 7 mantle cell lymphoma (MCL); 7 primary mediastinal large B-cell lymphoma (PMLBCL); 7 marginal zone B-cell lymphomas (MZL); 10 natural killer/T-cell lymphoma (NKTCL); 9 peripheral T-cell lymphoma, NOS (PTCL, NOS); and 6 angioimmunoblastic T-cell lymphoma (AILT). The cases of FL and DLBCL in this series include those reported in a previous study plus additional cases of DLBCL [15]. All cases were diagnosed at The University of Texas MD Anderson Cancer Center between 1987 and 2009, except for 9 previously reported gray-zone lymphomas provided by the Centro Oncológico MD Anderson International España (Madrid, Spain). The clinicopathologic features of the BCLU cases used in this study were published previously by 1 of the authors (J.F.G.) [16]. The diagnosis of all the neoplasms was based on morphologic and immunohistological criteria according to the World Health Organization 2008 classification criteria and reviewed by 4 of the authors (L.J.M., F.V.,

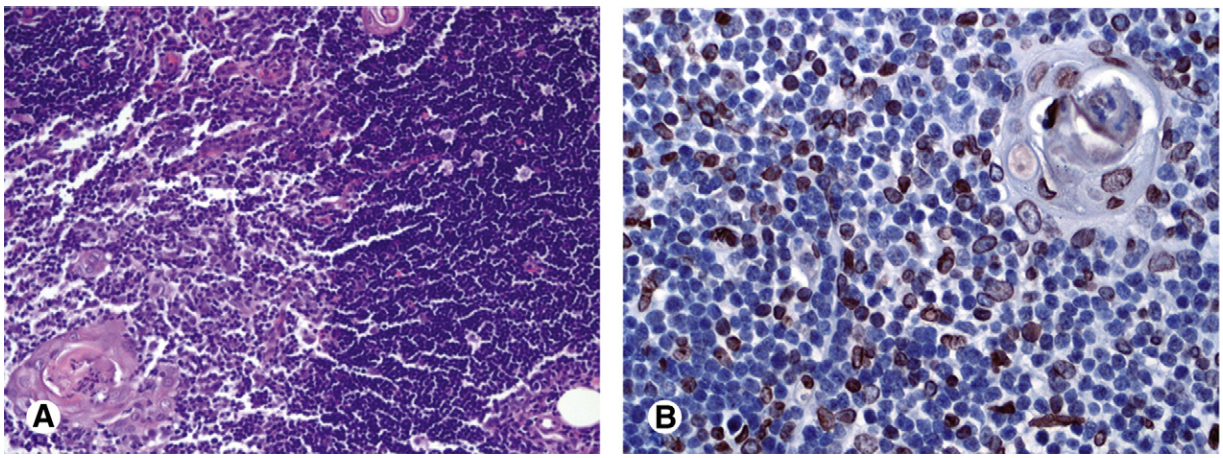


Fig. 2 Immunohistochemical expression of GLI3 in normal adult thymus. A, Benign thymus showing cortex and medulla with stromal cells including epithelial cells, histiocytes, and endothelial cells. B, GLI3 immunostain highlights cortical and medullary stromal cells in the thymus, including Hassall corpuscles. Thymocytes are negative (original magnification ×100-200).

Table 2 Expression of HH signaling proteins by immunohistochemistry in CHL

	No. of cases	Intensity of expression		Total positive	% Positive
		1+	2+		
HH	24	12	8	20	83
GLI1	25	7	0	7	28
GLI2	17	9	0	9	53
GLI3	39	0	39	39	100

J.E.K., and W.O.G.). The diagnosis of PMLBCL was based on clinicopathologic correlation and included only patients with an anterosuperior mediastinal mass, with no disease outside the mediastinum at the time of the diagnosis, and with morphologic and immunophenotypic features consistent with PMLBCL. The diagnosis of ALK-negative ALCL was established for those lymphomas morphologically within the spectrum of ALK-positive ALCL with strong and uniform expression of CD30, but lacking ALK protein expression.

Immunohistochemical methods were performed using routinely processed paraffin-embedded tissue specimens as previously described [17]. Antibodies and staining conditions are listed in Table 1. All tissue sections underwent heat-induced epitope retrieval in pH 6.0 citrate buffer (Dako, Carpinteria, CA). Endogenous peroxidase reaction was blocked by 3% H_2O_2 solution for 10 minutes. To avoid nonspecific binding of primary antibodies, serum-free blocking solution (Dako) was applied for 40 minutes at room temperature. Detection was performed using the LSAB plus-streptavidin-HRP system (Dako). For this study, we analyzed GLI3 protein expression in all 270 cases of lymphoma and 3 thymus specimens. In addition, other HH signaling proteins (HH, GLI1, and GLI2) were assessed in a subset of CHL cases that were available including HH (n = 24), GLI1 (n = 25), and GLI2 (n = 17). The baseline expression of HH signaling proteins in normal lymphoid cells was previously assessed and described [15]. Appropriate external controls were used for each antibody.

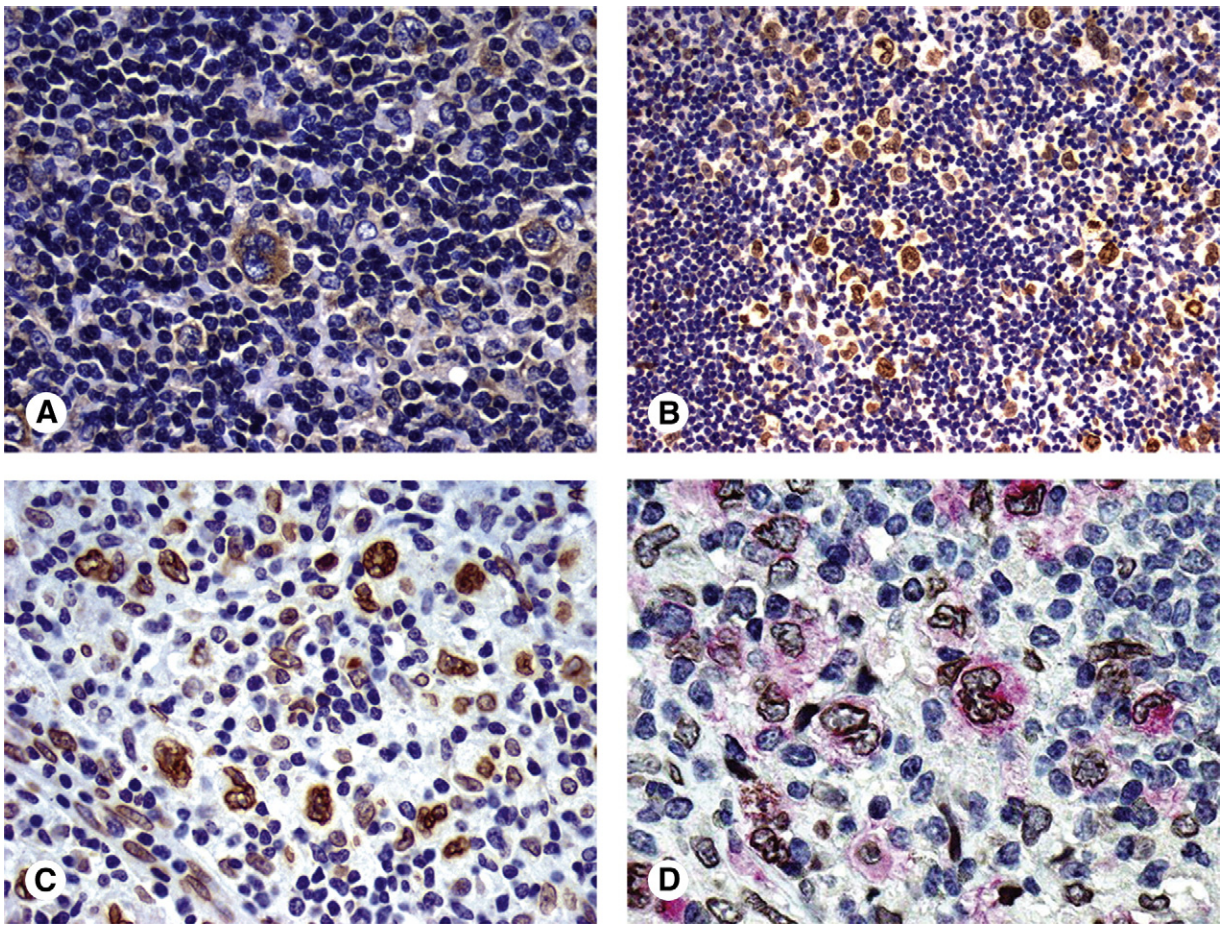


Fig. 3 Immunohistochemical expression of HH ligands and GLI3 in HRS cells. A, Cytoplasmic staining for HH ligands in HRS cells. Lymphocytes are negative for HH ligands (B and C), GLI3 shows characteristic strong nuclear staining with nuclear membrane accentuation in virtually all HRS cells. Note in panel C that the histiocytes and endothelial cells show weaker staining for GLI3 than the HRS cells. The neoplastic cells are easily recognized by their large size, irregular nuclear contours, and, in some cases, multinucleation. D, Double immunostaining with GLI3 (brown chromogen) and CD30 (red) confirms the expression of GLI3 in neoplastic HRS cells (original magnification ×100-1000).

Table 3 Immunohistochemical expression of GLI3 in lymphoid neoplasms

Diagnosis	n	Positive (%)	% Positive tumor cells per case [a], median (range)	Number of cases with % positive tumor cells			
				0%-25%	26%-50%	51%-75%	76%-100%
CHL	39	39 (100)	98 (80-100) [b]	0	0	0	39
NLPHL	13	6 (46)	55 (10-100)	9	1	1	2
CLL/SLL	16	7 (44)	83 (80-90)	9	0	0	7 [c]
FL	28	2 (7)	10 (10)	28	0	0	0
MZL	4	0	0	4	0	0	0
SMZBCL	3	0	0	3	0	0	0
MCL	7	0	0	7	0	0	0
PMLBCL	7	7 (100)	25 (10-40)	6	1	0	0
DLBCL	54	21 (39)	15 (10-50)	51	3	0	0
Gray zone	13	13 (100)	40 (10-100)	5	3	0	5
TCHRLBCL	12	6 (50)	27 (10-80)	9	2	0	1
BL	12	0	0	12	0	0	0
PTCL, NOS	9	1 (11)	25	9	0	0	0
AILT	6	0	0	6	0	0	0
ALCL, ALK negative	16	12 (75)	20 (10-100)	11	0	0	5
ALCL, ALK positive	21	21 (100)	65 (10-100)	5	3	5	8
NKTCL	10	0	0	10	0	0	0
Total	270						

NOTE. Gray zone lymphoma refers to BCLU with features intermediate between DLBCL and CHL.

[a] Only cases that showed positive staining are used to calculate the median.

[b] Three cases of CHL showed <100%, but more than 75%, GLI3-positive tumor cells.

[c] Only CLL/SLL cases with increased large cells showed GLI3 expression.

Cases were considered positive if 10% or more of the tumor cells exhibited cytoplasmic staining for HH or nuclear staining for GLI1, GLI2, and GLI3 as previously reported [18]. Protein expression was scored as negative, weakly positive (1+), and strongly positive (2+) depending on the staining signal intensity. In addition, the percentage of GLI3-positive cells was calculated by manually counting at least 200 tumor cells from representative fields in each case. We performed double immunostaining for GLI3 and CD30 or CD20 in a subset of cases of CHL, NLPHL, and TCHRLBCL and for GLI3 and cytokeratin, CD68, or CD20 in benign thymus using the EnVision G2 System (Dako) according to the manufacturer's recommendations.

2.4. Statistical analysis

All statistical analyses were 2 sided and performed using SPSS 16.0 (SPSS Inc, Chicago, IL) considering P values less than .05 (2-sided) as significant. Comparison of GLI3 expression frequencies among the different lymphoma types was done by Kruskal-Wallis and Mann-Whitney U non-parametric tests.

3. Results

3.1. Real-time qPCR analysis of CHL cell lines

Real-time qPCR analysis showed high expression of *GLI3* in all 3 CHL cell lines assessed. In contrast, *GLI1* and *GLI2* were variably expressed by CHL cell lines (Fig. 1). BJAB cells, a human germinal center–derived DLBCL cell line, and CLL cells isolated from the peripheral blood of a patient were used as control references. Studies in our laboratory have shown that expression of GLI3 is usually not detected in DLBCL cell lines and is frequently detected in circulating CLL cells.

3.2. Baseline expression of GLI3 in normal thymus and reactive lymph nodes as detected by immunohistochemistry

In the thymus, GLI3 was expressed by cortical and medullary stromal cells including epithelial cells, endothelial cells, and histiocytes (Fig. 2). Double immunostaining for CD68 and cytokeratin confirmed the identity of histiocytes and epithelial cells. Thymocytes were negative, and there was no coexpression of GLI3 in CD20-positive B cells (not shown).

In a previous study, we assessed the baseline protein expression of HH ligands, GLI1, GLI2, and GLI3 in benign reactive lymph nodes using immunohistochemistry [15]. Briefly, HH ligands were expressed by follicular dendritic cells and macrophages. GLI1 was not detected in centrocytes or centroblasts, and GLI2 was weakly positive in paracortical T cells and in a subset of centroblasts and centrocytes. GLI3 was expressed by histiocytes, follicular dendritic cells, and endothelial cells but not by lymphoid cells.

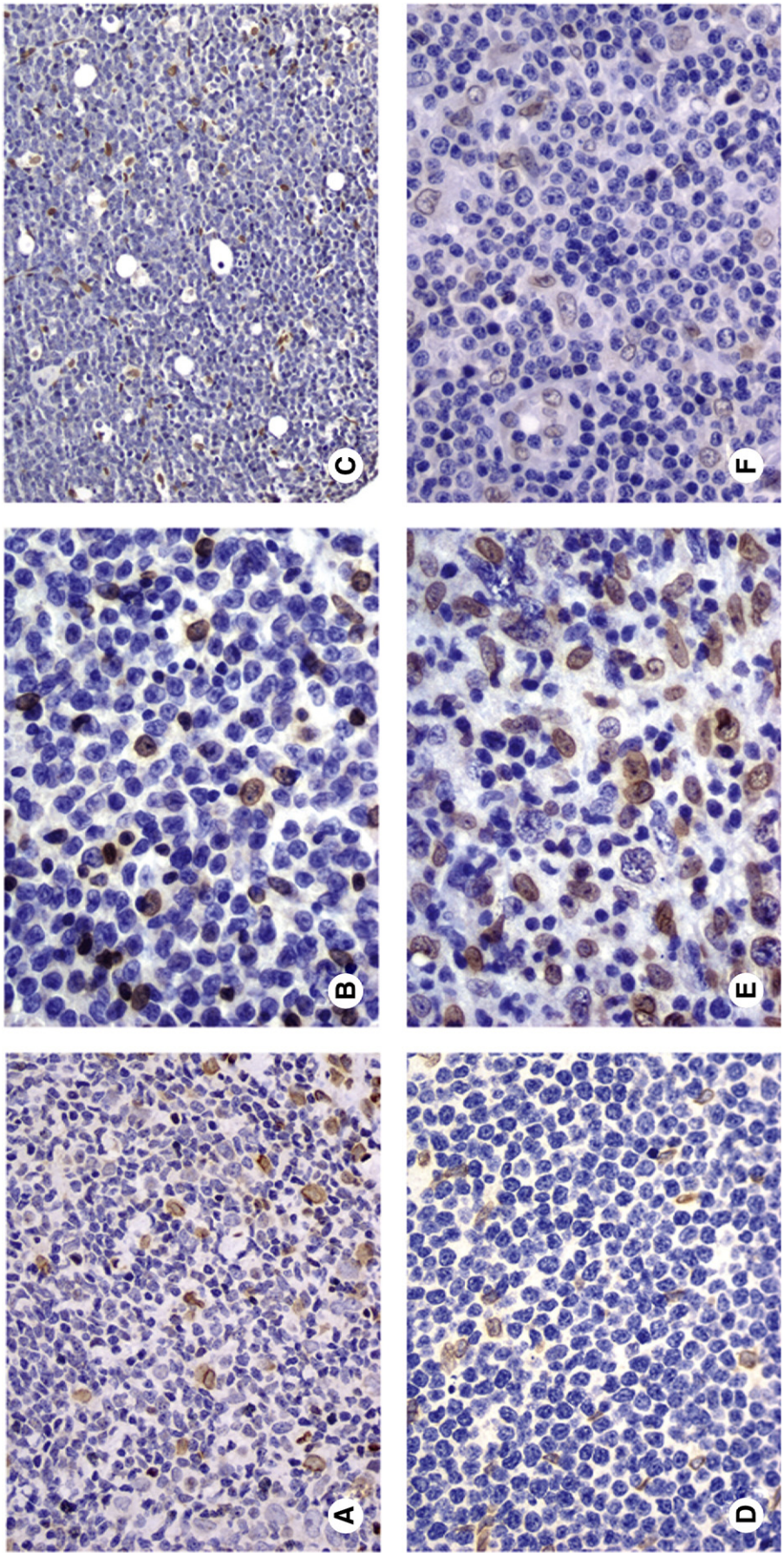


Fig. 4 GLI3 expression is negative in most non-HLs. Examples of non-HLs in which the tumor cells were negative or focally positive for GLI3: FL (A); DLBCL, NOS (B); BL (C); MCL (D); TCHRLBCL (E); and PTCL (F). In these cases, histiocytes, follicular dendritic cells and/or endothelial cells provide an internal positive control because they show positive nuclear staining for GLI3 (original magnification ×100-400).

3.3. GLI3 is characteristically and distinctively expressed in CHL

The expression of the HH signaling proteins in CHL as tested by immunohistochemistry is summarized in Table 2. HH, GLI1, and GLI2 showed variable expression in HRS cells. HH ligands were expressed in 20 (83%) of 24 cases (Fig. 3A). Most positive cases (12 of 20, or 60%) showed weak HH staining, whereas 8 cases showed strong (2+) expression. GLI1 was expressed in 7 (28%) of 25 cases and was weakly positive, and GLI2 was weakly positive in 9 (53%) of 17 cases. In contrast, all 39 CHL tumors showed strong and uniform expression of GLI3 in virtually 100% of the tumor cells (Fig. 3B and C). GLI3 in HRS cells was expressed in a nuclear pattern with nuclear membrane accentuation. Double IHC with CD30 performed in a subset of cases confirmed that the CD30-positive HRS cells were uniformly positive for GLI3 (Fig. 3D).

3.4. Survey of other lymphoma types

We assessed GLI3 expression by immunohistochemistry in a large series of B- and T-cell non-HLs. The findings are summarized in Table 3, and examples of the staining pattern of GLI3 in lymphomas other than CHL are shown in Fig. 4. Our data show that GLI3 is more frequently expressed in CHL tumors compared with all other lymphoma types studied ($P < .0001$).

Four tumor types showed variable GLI3 expression. ALK-negative ALCL, ALK-positive ALCL, BCLU with features intermediate between DLBCL and CHL, and NLPHL showed heterogeneous GLI3 expression, with a variable percentage positive tumor cells with a median of 20%, 65%, 40% and 55%, respectively (Figs. 5A-D). In most of these cases, only a subset of tumor cells were positive for GLI3; however, a subset of cases showed strong nuclear staining of GLI3 in 90% or more of tumor cells in a pattern indistinguishable from that seen in CHL (Fig. 5E-G). These cases included 2 (15%) of 13 NLPHLs, 5 (31%) of 16 ALK-negative ALCLs, 8 (38%) of 21 ALK-positive ALCLs, and 3 (23%) of 13 BCLUs with features intermediate between DLBCL and CHL.

Among B-cell non-HLs, there was no expression of GLI3 in SMZBCL, MZL, MCL, and BL, whereas 2 cases of FL showed positive staining for GLI3 in a subset of centroblasts. Large B-cell lymphomas including all 7 PMLBCL, a subset of DLBCL (NOS), and TCHRLBCL showed relatively low levels of GLI3 expression (median <30% positive tumor cells per case). In contrast with CHL, which showed uniformly strong GLI3 staining in all tumor cells, the intensity of GLI3 staining in the tumor cells of B-cell non-HLs, if positive, was always weak and lacked the distinctive nuclear membrane accentuation seen in CHL. Although the cytologic features as well as the immunostaining pattern of GLI3 allowed the distinction between histiocytes and tumor cells in most cases, we performed double immunostaining with GLI3 and CD20 in some cases of TCHRLBCL with a high content of histiocytes to highlight the tumor cells. These studies confirmed that the tumor cells were predominantly negative for GLI3 (not shown).

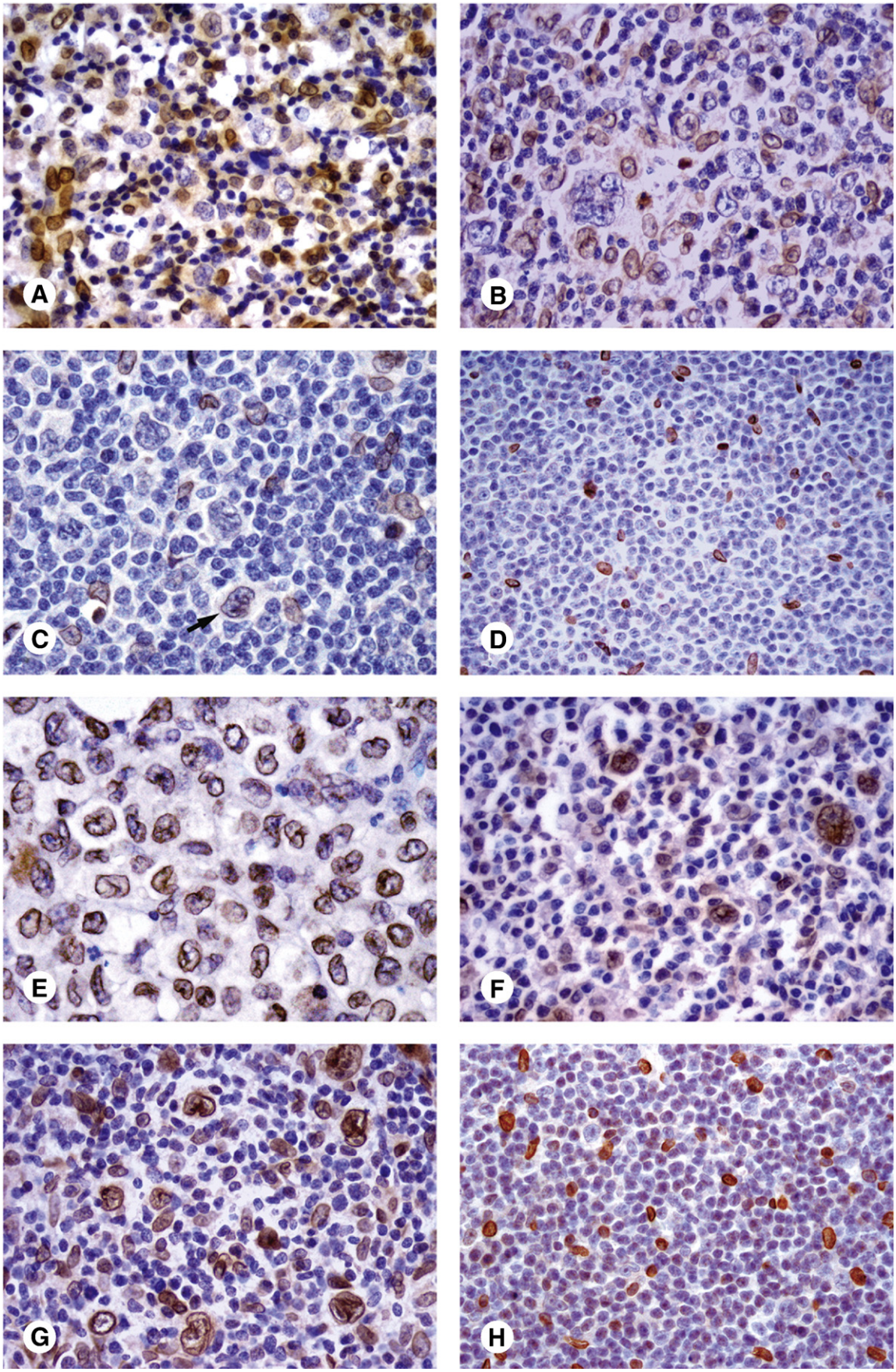
Interestingly, cases of CLL/SLL with significantly increased large cells showed positive GLI3 staining in the large cells (median, 83% positive large tumor cells). These tumor cells showed weak nuclear staining with no nuclear membrane accentuation, a pattern easily distinguishable from that seen in CHL (Fig. 4D). Cases of CLL/SLL without increased large cells did not express GLI3.

Of the T- and natural killer (NK) cell lymphomas, NKTCL, and AILT showed no GLI3 expression, and 1 (11%) of 9 PTCLs, NOS, showed weak nuclear GLI3 expression in a subset of neoplastic cells.

4. Discussion

This is the first study showing high expression of the HH transcription factor GLI3 in HRS cells of CHL. We also show that HRS cells express relatively low levels of other HH-related molecules including HH ligands and the transcription factors GLI1 and GLI2. GLI3 is universally expressed in HRS cells and is either not expressed or significantly less frequently expressed in other lymphomas. However, a small subset of other lymphomas also express GLI3 including ALK-positive ALCL, ALK-negative ALCL, BCLU with features intermediate between DLBCL and CHL, and NLPHL. A subset of tumors in these 4 categories, one third of cases or less, showed a pattern of GLI3 staining indistinguishable from that seen in CHL. The reason for the expression of GLI3 in some of these cases is unclear; however, these findings provide further support for the existence of a subset of lymphomas with intermediate/hybrid features between CHL and non-HL.

Fig. 5 GLI3 is heterogeneously expressed in cases of ALCL, ALK-positive, and ALK-negative, BCLU with features intermediate between DLBCL and CHLs and NLPHL, with most cases showing focal or no expression in the tumor cells. The tumor cells in most cases of ALK-negative ALCL (A), BCLU with features intermediate between DLBCL and CHL (B), NLPHL (C), and CLL/SLL (D) show negative or focally positive staining for GLI3. For those cases with focal positivity, GLI3 was expressed in a subset of the tumor cells with weaker expression than that seen in histiocytes and endothelial cells (black arrow in C). However, a small subset of ALK-negative ALCL (E), BCLU with features intermediate between DLBCL and CHL (F), and NLPHL (G) show a GLI3 expression pattern indistinguishable from CHL, that is, strong nuclear staining with nuclear membrane accentuation in virtually 100% of the tumor cells. The pattern of staining in CLL with increased large cells (H) was characterized by a weak nuclear expression of GLI3, as compared with intrinsic histiocytes and endothelial cells, without nuclear membrane accentuation. Note the stronger expression of GLI3 in histiocytes in comparison with the positive tumor cells.


A
B
C
D
E
F
G
H

Inappropriate activation of the HH signaling pathway has been shown in many cancers including ALK-positive ALCL and DLBCL [18-20]. However, this aberrant activation is mediated by the transcription factors GLI1 and GLI2 and not by GLI3. In fact, GLI3 mainly functions in vivo as a transcriptional repressor. Furthermore, in most systems studied, GLI3 limits HH pathway activation and absence of GLI3 causes up-regulation of HH signaling. Thus, the biologic significance of the expression of GLI3 in HRS is uncertain at this time.

Constitutive activation of the PI3K/AKT and NFκB pathways has been implicated in CHL [2], and crosstalk between these pathways with expression of HH ligands has been described. It has been shown that some growth factors such as platelet-derived growth factor, epidermal growth factor, and insulin-like growth factor 1 increase expression of HH ligands and that this increase is dependent of PI3K and AKT activation [21,22]. We and others have shown that adenoviral delivery of activated AKT up-regulates HH ligand expression and that recombinant HH induces activation of PI3K/AKT [18,22,23]. Similarly, it has been shown that NFκB regulates and contributes to HH ligand expression [24,25]. The mode of action of HH signaling in cancers seems to be complex and may be tumor dependent. Several studies suggest an autocrine action from HH signaling in some tumors, and others favor a paracrine model. In the paracrine model, neoplastic cells secrete HH ligands that may stimulate stromal cells near the tumor (endothelial cells, fibroblasts, and immune cells) to support tumor growth through secreting of growth factors. It is possible that the secretion of HH ligands by HRS contributes in the interactions between tumor cells and surrounding stroma in HL. However, the relationship between PI3K/AKT and NFkB signaling pathways with expression of GLI3 is uncertain at this time.

One of the historic challenges to a more complete understanding of the pathobiology of CHL is the lineage infidelity frequently demonstrated by HRS cells. For instance, for many years, CHL was thought to be derived from histiocytic/dendritic cells because they frequently express histiocytic/dendritic cell markers like fascin [26]. Other lineage-associated markers variably expressed by HRS cells include inhibitor of DNA binding 2 (ID2) (dendritic cell marker); CD3 and CD4 (T-cell markers) and CD15; and colony-stimulating factor 1 receptor (myeloid markers) [27-29]. Some authors have shown that the unique phenotype of HRS cells results from disruption or imbalances in the B-cell lineage transcriptional program and inhibition of the function of B-cell–determining transcription factor E protein 2A (E2A) [28,30]. Currently, HRS cells are considered to be of germinal center B-cell origin because they undergo B-cell receptor gene rearrangement and express B-cell specific transcription factors like paired box 5 (PAX-5) [2,31]. We suggest that the expression of GLI3 in HRS cells is another example of their lineage infidelity given that GLI3 is normally expressed by stromal cells in the lymph node (including follicular dendritic cells, histiocytes and endothelial cells) and thymus (histiocytes, endothelial cells, and epithelial cells), but not by mature or immature lymphocytes.

In addition, CHL is characterized by a rich inflammatory milieu including T and B cells admixed with variable numbers of plasma cells, macrophages, eosinophils, and mast cells [32,33]. Regulatory T cells comprise a considerable fraction of infiltrating T cells in CHL [34,35]. Tanijiri and colleagues [36] showed that the coculture of CHL cell lines with naïve $CD4^+$ T cells promotes the bidirectional differentiation of naïve T cells toward $CD25^+$ $Foxp3^+$ T cells and $CD4^+$ cytotoxic T cells without the cooperation of any other factor and without requiring additional cell types. This would explain, at least in part, the predominance of regulatory T cells in CHL. Furthermore, Hager-Theodorides and colleagues [37,38] recently showed that GLI3 in thymic stromal cells regulates T-cell selection and differentiation. Whether GLI3 plays a regulatory role in the T-cell population in CHL remains to be studied, but we think that it is an interesting hypothesis, and our findings provide a rationale for further investigation of this potential regulatory role.

From a practical point of view, our findings are also of particular interest because we found that a lack of GLI3 staining is a strong negative predictor against the diagnosis of CHL. Thus, GLI3 can be used in conjunction with the currently available extensive list of immunohistochemical markers to help distinguish CHL from possible mimics.

In summary, we have shown that GLI3 is strongly expressed by virtually 100% of the HRS cells of CHL. These data provide a rationale for further investigation of the biologic significance of GLI3 in CHL tumor biology, in particular, a potential regulatory role in the composition of the CHL microenvironment.

References

[1] Swerdlow SH, Campo E, Harris NL, et al., editors. WHO classification of tumours of haematopoietic and lymphoid tissues. Lyon: IARC Press; 2008.
[2] Kuppers R. The biology of Hodgkin's lymphoma. Nat Rev Cancer 2009;9:15-27.
[3] Schmitz R, Stanelle J, Hansmann ML, Kuppers R. Pathogenesis of classical and lymphocyte-predominant Hodgkin lymphoma. Annu Rev Pathol 2009;4:151-74.
[4] Rassidakis GZ, Medeiros LJ, Vassilakopoulos TP, et al. BCL-2 expression in Hodgkin and Reed-Sternberg cells of classical Hodgkin disease predicts a poorer prognosis in patients treated with ABVD or equivalent regimens. Blood 2002;100:3935-41.
[5] McMahon AP, Ingham PW, Tabin CJ. Developmental roles and clinical significance of hedgehog signaling. Curr Top Dev Biol 2003;53:1-114.
[6] Ingham PW, Placzek M. Orchestrating ontogenesis: variations on a theme by sonic hedgehog. Nat Rev Genet 2006;7:841-50.
[7] Sacedon R, Diez B, Nunez V, et al. Sonic hedgehog is produced by follicular dendritic cells and protects germinal center B cells from apoptosis. J Immunol 2005;174:1456-61.

[8] Dierks C, Grbic J, Zirlik K, et al. Essential role of stromally induced hedgehog signaling in B-cell malignancies. Nat Med 2007;13: 944-51.
[9] Johnson RL, Scott MP. New players and puzzles in the Hedgehog signaling pathway. Curr Opin Genet Dev 1998;8:450-6.
[10] Murone M, Rosenthal A, de Sauvage FJ. Hedgehog signal transduction: from flies to vertebrates. Exp Cell Res 1999;253: 25-33.
[11] Eichberger T, Sander V, Schnidar H, et al. Overlapping and distinct transcriptional regulator properties of the GLI1 and GLI2 oncogenes. Genomics 2006;87:616-32.
[12] Wang B, Fallon JF, Beachy PA. Hedgehog-regulated processing of Gli3 produces an anterior/posterior repressor gradient in the developing vertebrate limb. Cell 2000;100:423-34.
[13] Ruiz i Altaba A, Sanchez P, Dahmane N. Gli and hedgehog in cancer: tumours, embryos and stem cells. Nat Rev Cancer 2002;2:361-72.
[14] Pasca di Magliano M, Hebrok M. Hedgehog signalling in cancer formation and maintenance. Nat Rev Cancer 2003;3:903-11.
[15] Kim JE, Singh RR, Cho-Vega JH, et al. Sonic hedgehog signaling proteins and ATP-binding cassette G2 are aberrantly expressed in diffuse large B-cell lymphoma. Mod Pathol 2009;22:1312-20.
[16] Garcia JF, Mollejo M, Fraga M, et al. Large B-cell lymphoma with Hodgkin's features. Histopathology 2005;47:101-10.
[17] Vega F, Medeiros LJ, Leventaki V, et al. Activation of mammalian target of rapamycin signaling pathway contributes to tumor cell survival in anaplastic lymphoma kinase-positive anaplastic large cell lymphoma. Cancer Res 2006;66:6589-97.
[18] Singh RR, Cho-Vega JH, Davuluri Y, et al. Sonic hedgehog signaling pathway is activated in ALK-positive anaplastic large cell lymphoma. Cancer Res 2009;69:2550-8.
[19] Taipale J, Beachy PA. The Hedgehog and Wnt signalling pathways in cancer. Nature 2001;411:349-54.
[20] Singh RR, Kim JE, Davuluri Y, et al. Hedgehog signaling pathway is activated in diffuse large B-cell lymphoma and contributes to tumor cell survival and proliferation. Leukemia 2010;24:1025-36.
[21] Riobo NA, Lu K, Ai X, Haines GM, Emerson Jr CP. Phosphoinositide 3-kinase and Akt are essential for sonic Hedgehog signaling. Proc Natl Acad Sci U S A 2006;103:4505-10.
[22] Yang L, Wang Y, Mao H, et al. Sonic hedgehog is an autocrine viability factor for myofibroblastic hepatic stellate cells. J Hepatol 2008;48:98-106.
[23] Kanda S, Kanetake H, Miyata Y. HGF-induced capillary morphogenesis of endothelial cells is regulated by Src. Biochem Biophys Res Commun 2006;344:617-22.
[24] Kasperczyk H, Baumann B, Debatin KM, Fulda S. Characterization of sonic hedgehog as a novel NF-kappaB target gene that promotes NF-kappaB–mediated apoptosis resistance and tumor growth in vivo. FASEB J 2009;23:21-33.
[25] Nakashima H, Nakamura M, Yamaguchi H, et al. Nuclear factor-kappaB contributes to hedgehog signaling pathway activation through sonic hedgehog induction in pancreatic cancer. Cancer Res 2006;66: 7041-9.
[26] Uehira K, Amakawa R, Ito T, et al. A Hodgkin's disease cell line, KM-H2, shows biphenotypic features of dendritic cells and B cells. Int J Hematol 2001;73:236-44.
[27] Renne C, Martin-Subero JI, Eickernjager M, et al. Aberrant expression of ID2, a suppressor of B-cell–specific gene expression, in Hodgkin's lymphoma. Am J Pathol 2006;169:655-64.
[28] Mathas S, Janz M, Hummel F, et al. Intrinsic inhibition of transcription factor E2A by HLH proteins ABF-1 and Id2 mediates reprogramming of neoplastic B cells in Hodgkin lymphoma. Nat Immunol 2006;7: 207-15.
[29] Paietta E, Racevskis J, Stanley ER, Andreeff M, Papenhausen P, Wiernik PH. Expression of the macrophage growth factor, CSF-1 and its receptor c-fms by a Hodgkin's disease–derived cell line and its variants. Cancer Res 1990;50:2049-55.
[30] Fraser CR, Wang W, Gomez M, et al. Transformation of chronic lymphocytic leukemia/small lymphocytic lymphoma to interdigitating dendritic cell sarcoma: evidence for transdifferentiation of the lymphoma clone. Am J Clin Pathol 2009;132:928-39.
[31] Kuppers R, Rajewsky K, Zhao M, et al. Hodgkin disease: Hodgkin and Reed-Sternberg cells picked from histological sections show clonal immunoglobulin gene rearrangements and appear to be derived from B cells at various stages of development. Proc Natl Acad Sci U S A 1994;91:10962-6.
[32] Aldinucci D, Gloghini A, Pinto A, De Filippi R, Carbone A. The classical Hodgkin's lymphoma microenvironment and its role in promoting tumour growth and immune escape. J Pathol 2010;1221: 248-63.
[33] Xu C, de Vries R, Visser L, et al. Expression of CD1d and presence of invariant NKT cells in classical Hodgkin lymphoma. Am J Hematol 2010;85:539-41.
[34] Marshall NA, Christie LE, Munro LR, et al. Immunosuppressive regulatory T cells are abundant in the reactive lymphocytes of Hodgkin lymphoma. Blood 2004;103:1755-62.
[35] Alvaro T, Lejeune M, Salvado MT, et al. Outcome in Hodgkin's lymphoma can be predicted from the presence of accompanying cytotoxic and regulatory T cells. Clin Cancer Res 2005;11:1467-73.
[36] Tanijiri T, Shimizu T, Uehira K, et al. Hodgkin's reed-sternberg cell line (KM-H2) promotes a bidirectional differentiation of CD4+CD25+ Foxp3+ T cells and CD4+ cytotoxic T lymphocytes from CD4+ naive T cells. J Leukoc Biol 2007;82:576-84.
[37] Hager-Theodorides AL, Furmanski AL, Ross SE, Outram SV, Rowbotham NJ, Crompton T. The Gli3 transcription factor expressed in the thymus stroma controls thymocyte negative selection via Hedgehog-dependent and -independent mechanisms. J Immunol 2009;183:3023-32.
[38] Hager-Theodorides AL, Dessens JT, Outram SV, Crompton T. The transcription factor Gli3 regulates differentiation of fetal CD4- CD8- double-negative thymocytes. Blood 2005;106:1296-304.

Ann Hematol (2011) 90:939–946
DOI 10.1007/s00277-011-1179-2

ORIGINAL ARTICLE

Differential expression of JAK2 and Src kinase genes in response to hydroxyurea treatment in polycythemia vera and essential thrombocythemia

Enriqueta Albizua · Miguel Gallardo · Santiago Barrio · Inmaculada Rapado · Ana Jimenez · Rosa Ayala · Daniel Rueda · Beatriz Sanchez-Espiridion · Eulalia Puigdecanet · Blanca Espinet · Lourdes Florensa · Carles Besses · Joaquin Martinez-Lopez

Received: 10 December 2010 / Accepted: 31 January 2011 / Published online: 18 February 2011

Abstract This study investigates the differential gene expression profile of JAK2^{V617F}-positive myeloproliferative neoplasm (MPN) patients, with and without response to hydroxyurea (HU) treatment. Twenty-one polycythemia vera, 28 essential thrombocythemia, eight secondary erythrocytosis, and 30 controls were studied. Thirty-four genes were overexpressed in patients who did not respond to HU. Of these, some participate in proliferative pathways: *MAPK*, *AKT*, Src kinase (SFK), and JAK2 pathway. JAK2 allele burden was similar between groups of responders and nonresponder. A molecular fingerprint distinguishes JAK2^{V617F}-positive MPN patients without response to HU treatment, with overexpression of *JAK2*, *MAPK14*, *PIK3CA*, and *SFK* genes.

Keywords Polycythemia vera · Essential thrombocythemia · Gene expression profile · RT-PCR · Response to hydroxyurea · *JAK2* · Src family kinase (SFK)

Electronic supplementary material The online version of this article (doi:10.1007/s00277-011-1179-2) contains supplementary material, which is available to authorized users.

E. Albizua · M. Gallardo · S. Barrio · I. Rapado · A. Jimenez · R. Ayala · D. Rueda · J. Martinez-Lopez (✉)
Servicio de Hematología, Hospital Universitario 12 de Octubre, Avenida Córdoba s/n,
28041 Madrid, Spain
e-mail: jmartinezlo@yahoo.es

E. Albizua
Servicio de Hematología, Hospital Virgen de la Salud, Toledo, Spain

B. Sanchez-Espiridion
Spanish National Cancer Research Centre (CNIO), Madrid, Spain

E. Puigdecanet · B. Espinet · L. Florensa
Laboratoris de Citologia Hematológica/Citogenética Molecular. Servei de Patologia, GRETNHE, IMIM-Hospital del Mar, Barcelona, Spain

C. Besses
Servicio de Hematología, IMIM-Hospital del Mar, Barcelona, Spain

Introduction

The discovery of the JAK2^{V617F} point mutation has become a hallmark in the diagnosis of the classic chronic Philadelphia-negative myeloproliferative neoplasms (MPN) namely, polycythemia vera (PV), essential thrombocythemia (ET), and primary myelofibrosis (PMF) [1]. As a result, WHO diagnostic criteria were revised in 2007, and *JAK2* mutations and *MPL* mutations were included as major criteria for the diagnosis of PV, ET, and PMF [2]. Moreover, the JAK2^{V617F} allele burden has been included in response criteria for PV and ET by a recent European Leukaemia Net (ELN) consensus conference [3]. Despite the fact that JAK2^{V617F} has facilitated the diagnosis of MPN and improved our knowledge on their pathogenesis, there are still a number of unanswered questions [4].

Although formerly believed to be a single-step process, recent evidence has shown that the JAK2^{V617F} mutation is neither the sole, nor sufficient molecular event, responsible for MPN. It may also not be the initial one, and other somatic mutations may precede JAK2^{V617F} [5–7]. There is evidence of clonal heterogeneity and presence of different mutations in the same patient [8, 9].

The current first line of treatment for PV and ET patients with a high risk of thrombosis is still hydroxyurea (HU) [10]. The cytoreductive activity of HU is based on its anti-

metabolic role as a cell cycle-specific analog of urea, with activity in the S-phase. HU inhibits the enzyme, ribonucleotide reductase, which converts ribonucleotides to deoxiribonucleotides (critical precursors for de novo DNA biosynthesis and repair). Up to 15–25% of patients show resistance or intolerance to HU [11, 12]. The mechanism of resistance is believed to be the increased expression of ribonucleotide reductase due to gene amplification, increased transcription, and post-transcriptional mechanisms [13]. Despite this knowledge, there is still no clear understanding of the underlying molecular mechanism and consequences of HU activity.

We aimed to analyze the differential gene expression profile of $JAK2^{V617F}$-positive MPN patients with or without response to HU treatment, to determine possible mechanisms for a poor response. As a consequence, we hoped to identify alternative routes for targeted therapy. In addition, we aimed to correlate these results with clinical and laboratory features.

Design and methods

Patients

Forty-nine patients, diagnosed with MPN (21 PV and 28 ET) with the $JAK2^{V617F}$ mutation were included in the study, as well as eight patients with secondary erythrocytosis (SE) to hypoxemia, and 30 healthy control donors. An MPN diagnosis was established on the basis of World Health Organization criteria (WHO) 2001/2008 [2], or the Polycythemia Vera Southern Study Group (according to the standard criteria at the moment of diagnosis). Twenty MPN patients were studied at diagnosis (as control population) and 29 after treatment.

Clinical and laboratory features at diagnosis and follow-up features of the $JAK2^{V617F}$ PV and ET patients were assessed and are shown in Table 1.

Response to treatment with HU in PV and ET was defined according to update recommendations [10] and to clinico-hematological response criteria for PV and ET recently published by the ELN consensus conference [3].

The study was approved by the local ethics committee and written informed consent was obtained from all patients, according to the Declaration of Helsinki.

Nucleic acids extraction and $JAK2^{V617F}$ quantification

Peripheral venous blood was collected in EDTA and immediately processed. Granulocytes were isolated by Ficoll-Paque (Sigma-Aldrich, Saint Louis, Mi USA) density gradient centrifugation, as previously described. Erythrocytes were eliminated by lysis with a commercial buffer (Red Blood Cell Lysis Buffer, from Roche Applied Sciences, Mannheim, Germany). DNA was extracted using the Maxwell 16 SEV automated extraction system (Promega co., Madison, WI, USA). mRNA was extracted from 1×10^6 cells, with TRIzol (R) Reagent (Invitrogen Life Technologies, Inc., Paisley, UK) according to the manufacturer's instructions, with more

Table 1 The clinical and laboratory features of $JAK2^{V617F}$-positive PV and ET patients at diagnosis and follow-up

Clinical and laboratory variables	PV with treatment	ET with treatment	PV at diagnosis	ET at diagnosis
Patients	10	16	11	12
Gender (M/F)	5/5	6/10	6/5	2/10
Age at diagnosis[a]	64.5 (37–77)	65 (20–92)	57 (21–74)	56 (39–87)
Splenomegaly at diagnosis	0/10	0/16	1/11	1/12
Hepatomegaly at diagnosis	2/10	2/16	1/11	0/12
Hemoglobin (g/100 mL) at diagnosis[a]	20.05 (17.5–22.5)	15.2 (12.5–16.4)	15.5 (14.2–21.3)	13.9 (11.9–19.5)
Hematocrit at diagnosis[a]	58 (51.1–66.4)	45.9 (38.5–49.4)	46.3 (42.7–61)	43 (39.2–59)
WBC ($\times 10^9$/L) at diagnosis[a]	11.1 (5.91–14.7)	9.61 (5.96–19.2)	8.44 (5.2–14.34)	10.2 (6.9–14.8)
Platelets ($\times 10^9$/L) at diagnosis[a]	436.5 (305–1328)	821 (552–11580)	582 (215–819)	655 (538–996)
Increased LDH at diagnosis	4/10	3/16	1/11	1/2
Thrombotic events at diagnosis	0/10	3/16	2/11	1/12
Hemorrhagic events at diagnosis	0/10	2/16	0/11	0/12
Thrombotic events after diagnosis	4/10	3/16	–	–
Hemorrhagic events after diagnosis	0/10	0/16	–	–
JAK2V617F allele burden[a]	20.53% (1.64–100)	6.95% (0.67–29.52)	36.84% (2.32–128.87)	23.45% (2.32–36.09)
Disease duration in months	60.5 (1–122)	118 (4–216)	0	0
Treatment duration in months[a]	60.5 (1–122)	118 (4–216)	0	0
Response to HU treatment	4/10	7/16	–	–

PV polycythemia vera, *ET* essential thrombocythemia, *HU* hydroxyurea, *M* male, *F* female, *WBC* white blood cells, *LDH* lactate dehydrogenase, *NA* not applicable

[a] Median value (range) is reported

than 90% granulocytes. Lymphocyte contamination was assessed in five samples by flow cytometry, and was less than 2% of the total cell count. DNA and RNA quantity and quality were assessed on a ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA) for low-density array experiments, and by nanoelectrophoresis using the Nano lab-on-a-chip assay for total eukaryotic RNA (Bioanalyzer, Agilent Technologies, Palo Alto, CA, USA) for microarrays. Retrotranscription was performed with a MMLV-modified enzyme, according to manufacturer's instructions (Applied Biosystems, Palo Alto, CA, USA). Mutational screening for $JAK2^{V617F}$ was performed with real-time (RT) PCR using DNA from whole peripheral blood from the study sample [14, 15].

Microarray analyses

As screening proceeding, 14 samples from JAK2 ^{V617F}-positive patients diagnosed with MPN (3 PV and 11 ET) were included in the study, as well as 20 samples from healthy control donors (in four pools from five samples); 11 MPN were under treatment and 3 were under no treatment. Microarray expression profiles were obtained using the Whole Human Genome Oligo Microarray Kit (Agilent Technologies, Inc., Santa Clara, CA, USA). Hybridization was performed with Stratagene Universal Human Reference RNA, and a mix of tumor cell lines (Agilent Technologies Stratagene, Inc., Santa Clara, CA, USA). Briefly, for each sample, 2 μg of total RNA was mixed with 2 μL of a 5,000-fold dilution of Agilent's Two-Color Spike-in RNA control. The mixture was amplified using the low-input RNA amplification kit (Agilent Technologies Inc., Santa Clara, CA, USA). Following amplification and labeling with Cy3, each sample was assessed on the Nanodrop ND-1000 (NanoDrop Technologies, Wilmington, DE, USA) to measure yield and specific activity. The amplified and labeled samples were hybridized in a rotating oven to Agilent 44K human whole genome microarrays, according to the manufacturer's instructions. The images were scanned with a G2565BA Microarray Scanner System (Agilent Technologies Inc., Santa Clara, CA, USA). The data were processed and normalized with the use of Feature Extraction (v.9.0) software. Each time point was additionally normalized to time zero. Genes with more than 10% missing values were discarded. Normalized log2 ratios were scaled between arrays to make all data comparable, and processed again with pre-processor of GEPAS 3.1 (http://gepas3.bioinfo.cipf.es/) to purify the data.

The student's *t* test was employed as statistical tool in the gene expression pattern analysis suite (GEPAS 3.1), Pomelo II (http://pomelo2.bioinfo.cnio.es/) and gene set enrichment analysis (GSEA) programs, used to interpret gene expression data and to determine statistical significance of differential gene expression, as described [16].

The supervised classification of healthy control donors, ET and PV samples into categories was performed.

Low-density quantitative array (LDA)

To reassess microarray data, 40 $JAK2^{V617F}$ -positive patients, diagnosed with MPN were included in the study: 20 MPN patients receiving HU treatment (10 PV and 10 ET) and 20 MPN newly diagnosed patients (10 PV and 10 ET), used as control population to assay the influence of phenotype. Eight SE to hypoxemia as well as ten healthy control donors (in two pools from five samples) were also included in the study. Related to the 20 MPN patients with HU treatment, 9 of them showed response to HU (four PV and five ET) and 11 were non-responders (six PV and five ET). Among the MPN patients with HU treatment, two PV and three ET patients had been previously included in the microarray assay

Eighty-four genes were selected for their differential expression from several different dicothomous analyses in the microarray study. Another nine genes, though without differential expression in microarray analysis, were selected for their key role in erythropoiesis, thrombopoiesis, or signaling pathways in MPN; and three housekeeping genes (*GADPH*, *18 s rRNA*, and *ABL1*) were also included (Supplementary Table 1: Supplementary Data). All these genes were assayed by quantitative PCR (using the 96 TaqMan Low Density Arrays platform LDA). Besides, three individual TaqMan assays were performed with three additional genes of interest according to other MPN studies, in the same samples of patients.

Custom-made TaqMan low-density arrays with a setup of 96 different genes were applied for relative mRNA quantification. cDNA samples were subjected to RT-PCR in duplicate in an ABI PRISM 7900HT Sequence Detection System (Applied Biosystems, Palo Alto, CA, USA). Relative expression of each gene was quantified by the comparative cycle threshold (Ct) method, using the Abelson murine leukemia viral oncogene homolog 1 (*ABL1*) as an endogenous control. Gene expression quantification was achieved using the comparative Ct method for relative quantification, in which the amount of target is expressed as $2^{-\Delta Ct}$ [15].

Statistical analyses

Statistical analyses of gene expression results in LDA, and of the clinical and biological data and their relationship with the gene expression data were performed using the SPSS15 software (SPSS Inc., Chicago, Illinois, USA). Differences between non-normally distributed group data were analyzed using the non-parametric Wilcoxon analysis. Statistical significance was considered when *P* value was less than 0.05.

Results

Microarray results based on biological and functional pathways analyses

Microarray analysis, used as screening test, showed 84 genes with significant differential expression by GSEA (conditions required in GSEA analysis/core/enrichment/YES), related to biological and functional pathways of GSEA tool (significant different expression by GSEA was considered when P value was <0.05 and/or the false discovery rate (FDR) was <0.25).

Genes were selected from different analyses performed in a dichotomic manner: ET versus PV, ET versus controls, PV versus controls, and MPN with response to HU against those without response to HU.

Table 2 Differential gene expression of $JAK2^{V617F}$-positive polycythemia vera (PV) and essential thrombocythemia (ET) patients

Genes	ET vs. PV			ET vs. PV (with HU treatment)			ET vs. PV (at diagnosis)			Response to HU treatment		
	P value	Ratio		*P* value	Ratio		*P* value	Ratio		*P* value	Ratio	
Number of patients	20/20	ET 20	PV 20	10/10	ET 10	PV 10	10/10	ET 10	PV 10	11/9	Yes 11	No 9
ACSL	0.044[a]	4.072	4.609	0.000[a]	3.611	5.602	0.598	4.534	3.617	0.017[a]	4.235	5.060
ANXA1	0.040[a]	2.282	2.756	0.000[a]	1.953	3.275	0.975	2.610	2.209	0.020[a]	2.320	2.973
BTK	0.0032[a]	0.467	1.636	0.000[a]	−0.707	1.807	0.729	1.642	1.466	0.016[a]	0.082	1.121
CASP8	0.003[a]	2.838	3.459	0.000[a]	2.210	3.775	1.000	3.466	3.144	0.039[a]	2.667	3.391
CD44	0.008[a]	2.422	3.229	0.000[a]	1.796	3.961	0.424	3.049	2.497	0.005[a]	2.348	3.527
CXCL1	0.054	2.359	3.156	0.000[a]	1.214	3.168	0.940	3.504	3.144	0.014[a]	1.713	2.775
EDN1	0.218	−1.292	−0.679	0.002[a]	−1.990	−0.433	0.598	−0.593	−0.926	0.009[a]	−1.818	−0.471
F2R	0.02[a]	−2.409	−2.308	0.091	−2.404	−1.632	0.452	−2.414	−3.019	0.019[a]	−2.420	−1.526
FCER1G	0.0021[a]	3.511	4.682	0.000[a]	2.494	5.415	0.497	4.527	3.950	0.004[a]	3.295	4.760
FRMD4B	0.001[a]	1.148	2.579	0.000[a]	0.104	2.931	0.845	2.191	2.228	0.001[a]	0.735	2.473
GAPDH	0.011[a]	2.464	3.167	0.000[a]	1.719	4.164	0.577	3.208	2.170	0.037[a]	2.520	3.457
GSK3B	0.094	1.550	2.191	0.000[a]	0.567	2.650	0.220	2.533	1.733	0.024[a]	1.176	2.137
HMGB2	0.014[a]	3.188	3.732	0.000[a]	2.185	4.375	0.869	4.190	3.089	0.014[a]	2.754	3.924
IGF1R	0.052	2.417	2.703	0.000[a]	1.715	3.538	0.798	3.120	1.867	0.026[a]	2.259	3.076
IKZF1	0.022[a]	1.102	2.187	0.000[a]	0.503	2.190	0.752	1.700	2.184	0.024[a]	0.970	1.807
ITGAM	0.065	2.718	3.384	0.000[a]	1.905	3.912	0.283	3.532	2.857	0.004[a]	2.400	3.530
JAK2	0.002[a]	1.527	2.862	0.000[a]	0.444	2.995	0.892	2.611	2.729	0.008[a]	1.096	2.482
LITAF	0.002[a]	4.366	5.142	0.000[a]	4.146	6.122	0.497	4.586	4.161	0.016[a]	4.693	5.673
LRMP	0.007[a]	2.496	3.532	0.000[a]	1.558	3.740	0.752	3.434	3.324	0.001[a]	2.134	3.280
LYN	0.002[a]	4.535	5.686	0.000[a]	0.648	1.931	0.616	5.321	5.450	0.017[a]	4.386	5.372
MAP2K1	0.053	1.070	1.683	0.001[a]	1.931	0.647	0.598	1.493	1.435	0.045[a]	0.972	1.677
MAPK14	0.006[a]	3.606	4.314	0.000[a]	3.200	5.414	0.460	4.012	3.215	0.007[a]	3.834	4.885
MMP14	0.000[a]	−4.422	−2.794	0.001[a]	−4.430	−2.759	0.003[a]	−4.414	−2.829	0.106	−3.995	−3.105
PECAM1	0.021	2.641	3.562	0.000[a]	1.676	3.782	0.869	3.605	3.342	0.030[a]	2.414	3.114
PIK3CA	0.015[a]	−0.490	0.217	0.000[a]	−1.424	0.434	0.752	0.445	0.000	0.020[a]	−0.881	−0.024
PLA2G4A	0.082	−1.904	−1.446	0.002[a]	−2.609	−1.301	0.798	−1.199	−1.590	0.022[a]	−2.289	−1.547
RAF1	0.005[a]	2.039	2.954	0.000[a]	1.352	3.082	0.845	2.726	2.826	0.003[a]	1.811	2.713
S100A4	0.059	3.779	4.334	0.000[a]	3.072	4.845	0.478	4.486	3.823	0.037[a]	3.679	4.301
SEC31B	0.114	−0.260	0.194	0.002[a]	−0.463	0.750	0.346	−0.058	−0.391	0.026[a]	−0.184	0.544
SELL	0.005[a]	3.681	4.666	0.000[a]	2.765	5.148	0.892	4.596	4.184	0.014[a]	3.527	4.482
SELP	0.010[a]	−1.000	0.040	0.002[a]	−1.635	0.166	0.641	−0.365	−0.086	0.019[a]	−1.067	−0.327
SKAP2	0.008[a]	2.230	3.254	0.000[a]	1.330	3.675	0.845	3.131	2.832	0.013[a]	2.053	3.052
SPON1	0.469	−5.401	−5.314	0.030[a]	−5.398	−4.389	0.459	−5.403	−6.239	0.027[a]	−5.269	−4.435
TAP1	0.072	2.677	3.522	0.000[a]	2.441	3.923	0.684	2.913	3.121	0.042[a]	2.921	3.500
YWHAH	0.025[a]	1.650	2.436	0.000[a]	1.188	2.937	0.660	2.137	1.880	0.034[a]	1.672	2.538

[a] Results of the low-density array (LDA) analysis ($P<0.05$) and results of Wilcoxon analyses between clinical and laboratory features relative to the gene expression profile are shown

Low-density array analysis

Gene expression profile and response to HU treatment

A set of 34 genes were found to be differentially expressed in MPN patients who either responded or showed no response to HU treatment, and were significantly overexpressed in those patients with no response (Table 2; Fig. 1). Interestingly, the following genes were significantly overexpressed in patients who showed no response to HU treatment: *JAK2* (ER=2.482 vs. 1.096; *P*=0.008), *IGF1R* (ER=3.076 vs. 2.259; *P*=0.026), *MAPK14* (ER=4.885 vs. 3.834; *P*=0.007), *FCRE1G* (ER=4.760 vs. 3.295; *P*=0.004), *LYN* (ER=5.372 vs. 4.386; *P*=0.017), *BTK* (ER=1.121 vs. 0.082; *P*=0.016), and *PIK3CA* (ER=4.826 vs. 3.731; *P*=0.011) (Fig. 1).

JAK2 was overexpressed in patients who did not respond to treatment; and in patients who did respond to treatment, *JAK2* expression was lowered to equivalent levels to those of SE patients and healthy controls. A similar gene expression pattern was observed for *MAPK14*, *FCRE1G*, *LYN*, *BTK*, and *PIK3CA* (Fig. 1).

These differences were not due to disease duration because when we compared patients with a short disease duration (below 2 years) and patients with a long disease duration (over 2 years) no differences in gene expression profile were found (Supplementary Table 2, Supplementary Data).

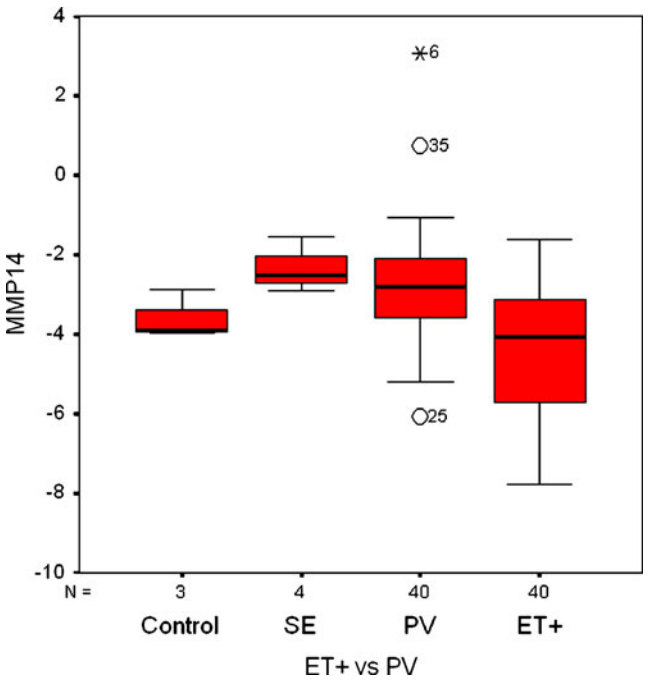


Fig. 1 *MMP14* gene is differentially expressed in JAK2 ^{V617F} positive polycythemia vera (PV) and essential thrombocythemia (ET) patients at diagnosis. Box plot of *MMP14* expression ratios: *Y*-axis: *MMP14* ratios; *X*-axis: from left to right: healthy controls, SE, PV and ET patients

The gene expression profile of the control group was similar to that of patients who responded to HU treatment.

JAK2 allele burden quantification

In order to assay if the lowering expression of JAK2 gene was due to a decrease of the JAK2^{V617F} allele burden, it was measured in patients at diagnosis and under HU treatment. When measured after HU treatment, JAK2^{V617F} allele burden diminished in PV, but only in a significant manner in ET (*P*=0.003). However, the allele burden differences between patients who responded, or those who did not respond to HU, were not statistically significant (Fig. 1).

Gene expression profile associated with response is not influenced by phenotypic diversity

To assay whether the different *gene expression profile* between responders to HU treatment and the non-responder group could be due to phenotypical differences, gene expression of patients under HU treatment was compared with a balanced population of patients at diagnosis.

Upon separate analysis of samples from PV and ET patients receiving HU treatment, and those patients at diagnosis, only one gene, *MMP14*, was overexpressed in PV newly diagnosed patients compared to ET at diagnosis. Moreover, this overexpression was maintained when comparison was established between treated and untreated ET patients against PV treated and untreated (ER: expression ratios=−4.442 vs.−2.794; *P*=0.000) (Table 2; Fig. 2).

Correlation between clinical and laboratory features and gene expression profile in PV and ET

Clinical and laboratory variables, referred in Table 1 were analyzed and compared in a dichotomic manner with gene expression profile to search correlates: (1) at diagnosis: gender, age, splenomegaly, hepatomegaly, hemoglobin, hematocrit, white blood cells, platelets, lactate dehydrogenase, thrombotic events, hemorrhagic events, JAK2^{V617F} allele burden; (2) and in the follow up: disease duration, treatment duration, thrombotic events after diagnosis, hemorrhagic events after diagnosis, and response to HU treatment (Supplementary Table 2: Supplementary Data).

Discussion

Microarray analysis looking for a differential gene fingerprint profile between JAK2^{V617F} MPN different groups treated with HU resulted in the identification of 84 genes with significant differential expression. These genes play

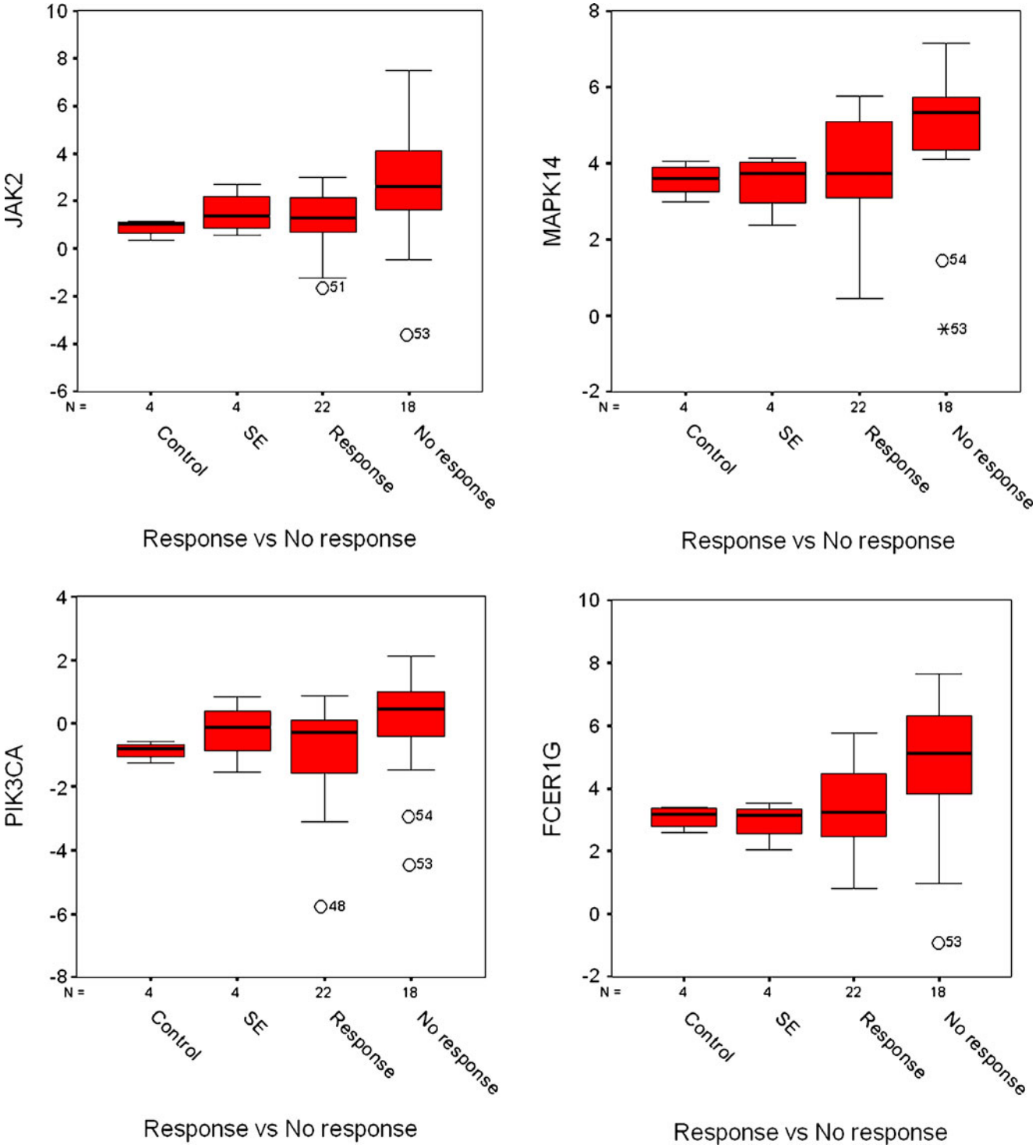


Fig. 2 The differential gene expression profile in polycythemia vera (PV) and essential thrombocythemia (ET) patients relative to their response to hydroxyurea (HU) treatment. Box plot of *JAK2*, *MAPK14*, *PIK3CA* and FCER1G expression ratios: Y axis: ratio; X axis: in every box plot compared groups: healthy controls, SE, PV and ET

key roles as transcription factors (*LYN*, *PIK3CA*), in MAP kinase (MAPK) signaling pathways (*MAPK14*, *RAF1*, *RASA1*, *YWHAH*), as tyrosine kinases (*JAK2*, *BTK*), in development and function of the hematological system (*CD44*, *MAP1B*, *JARID2*, *HMGB2, IKZF1*), and in other signaling and apoptotic pathways (*SKAP2*, *PLAUR*, *LRMP*). Some of these genes have been previously outlined in other microarray analyses or functional studies [17, 18].

When we looked at this gene signature with LDA (a real-time PCR approach), we observed a differential gene expression profile in MPN patients according to response to HU. We found 34 genes that were overexpressed in patients who did not respond to HU treatment. Interestingly, among these 34 overexpressed genes, some participate in four important proliferative pathways: *MAPK14*, *MAP2K1*, *RAF1* (MAPK pathway), *PIK3CA*, *FRMD4B* (AKT pathway), *LYN*, *FCER1G*, *BTK*, *SKAP2* (Src family kinase, SFK), and *JAK2*, *IGFR1* (JAK2 pathway). Gene expression did not change with disease evolution but it did with the response to treatment.

A number of genes in the MAPK pathway were differentially expressed in MPN patients relative to their response to HU. An overexpression of genes implicated in this pathway was shown in the group who did not respond to HU, notably *RAF1*, *MAP2K1*, and *MAPK14*. In the group of patients who did respond to HU treatment, MAPK genes expression was lowered to levels equivalent to those observed in SE patients and controls. The activation and overexpression of *MAPK14* has been described in PV pathogenesis [1, 17, 19]. Overexpression of genes in the MAPK pathway may be involved in the lack of response to HU treatment, although by yet unknown mechanisms, and probably related to *JAK2* gene expression.

The *AKT* gene, with a crucial role in the AKT pathway, has been widely implicated in the pathogenesis of MPN (PV and ET), by overactivation and overexpression [1, 19, 20]. We found that the *PIK3CA* gene, an AKT regulator, was also overexpressed in patients who did not respond to HU treatment. This overexpression may also be involved in the lack of response to treatment, and correlate with *JAK2* overexpression.

The SFK are also suggested to contribute to signaling by EpoR [21], although they seem not to be indispensable for the PV phenotype induced by $JAK2^{V617F}$, as shown by bone marrow reconstitution studies in mice deficient for LYN, FYN, or Fgr kinases [22]. Our results showed that BTK and other SFK members, such as LYN, FCER1G, and SKAP2, were overexpressed in patients who did not respond to HU treatment. Thus, an overexpression of these genes may be related to a lack of response to HU, while decreased levels, similar to those observed in SE patients and controls, may be related to a good response to HU treatment. These findings could be of great interest, since those patients with overexpressed SFK genes may benefit from treatment with SFK inhibitors (such as Dasatinib). In fact, some groups are testing Dasatinib in cell cultures from PV patients with some success [22, 23].

In addition, HU treatment in MPN patients showed a resulting decrease in, not only the $JAK2^{V617F}$ allele burden, but also in response to treatment in *JAK2* gene expression. Interestingly, while *JAK2* was overexpressed in patients who did not respond to treatment, in those patients who responded well to HU, *JAK2* gene expression was lowered to levels equivalent to those of SE patients and healthy controls. We also showed that although the $JAK2^{V617F}$ allele burden diminished after HU treatment, as reported by others [24], there was no statistical difference between the levels of the $JAK2^{V617F}$ allele burden in our different patient groups, with or without a response to HU. However, *JAK2* gene differential expression between the same groups of patients with or without a response to HU treatment was statistically significant. Thus, *JAK2* overexpression may reflect of lack of response to HU, although the mechanism for this is unknown. The criteria for a lack of response to HU treatment could thus be expressed, either by the mutant allele burden, and/or by *JAK2* overexpression. To our knowledge, this is the first study to show these results, and could be useful as an additional parameter of response.

Finally, we found that this different pattern of expression was not due to phenotypic differences, as demonstrated when compared with a control group of patients without treatment.

Only one gene, *MMP14*, remained significant between PV and ET patients at diagnosis. *MMP14* is a matrix metalloprotease involved, among other things, in tumor metastasis. It has an indirect connection with some proliferative pathways, such as the MAPK pathway. Some authors have made reference to the importance of the MAPK pathway in PV pathogenesis, and its connection with JAK2 [1, 17]. On the other hand, MMP14 has also been demonstrated to be overexpressed in PMF, and is modulated by *TIMP2*, which has erythroid-potentiating activity [25, 26]. Thus, the overexpression of *MMP14* in PV, compared with ET, may contribute to the final myeloproliferative phenotype of PV by enhancing the commitment to erythropoiesis via MAPK pathway, either by JAK/STAT, TIMP2, or by alternative routes.

To conclude, our results suggest that a differential molecular fingerprint of genes distinguishes MPN patients who either do or do not respond to HU treatment. *JAK2*, *MAPK14*, *PIK3CA*, and SFK genes are overexpressed in patients who do not respond to HU. In our group of patients, the $JAK2^{V617F}$ allele burden decreased with HU treatment; however, a significant diminution was not observed in response to treatment. On the other hand, patients with no response to treatment overexpressed *JAK2*. These results need to be confirmed with further functional studies.

References

1. James C, Ugo V, Le Couedic JP, Staerk J, Delhommeau F, Lacout C et al (2005) A unique clonal JAK2 mutation leading to

constitutive signalling causes polycythaemia vera. Nature 434 (7037):1144–1148
2. Tefferi A, Thiele J, Orazi A, Kvasnicka HM, Barbui T, Hanson CA et al (2007) Proposals and rationale for revision of the World Health Organization diagnostic criteria for polycythemia vera, essential thrombocythemia, and primary myelofibrosis: recommendations from an ad hoc international expert panel. Blood 8:8
3. Barosi G, Birgegard G, Finazzi G, Griesshammer M, Harrison C, Hasselbalch HC et al (2009) Response criteria for essential thrombocythemia and polycythemia vera: result of a European LeukemiaNet consensus conference. Blood 113(20):4829–4833
4. Goldman JM, Green AR, Holyoake T, Jamieson C, Mesa R, Mughal T et al (2009) Chronic myeloproliferative diseases with and without the Ph chromosome: some unresolved issues. Leukemia 23:1708–1715
5. Levine RL, Belisle C, Wadleigh M, Zahrieh D, Lee S, Chagnon P et al (2006) X-inactivation-based clonality analysis and quantitative JAK2V617F assessment reveal a strong association between clonality and JAK2V617F in PV but not ET/MMM, and identifies a subset of JAK2V617F-negative ET and MMM patients with clonal hematopoiesis. Blood 107(10):4139–4141, Epub 2006 Jan 24
6. Kralovics R, Teo SS, Li S, Theocharides A, Buser AS, Tichelli A et al (2006) Acquisition of the V617F mutation of JAK2 is a late genetic event in a subset of patients with myeloproliferative disorders. Blood 108(4):1377–1380
7. Lambert JR, Everington T, Linch DC, Gale RE (2009) In essential thrombocythemia, multiple JAK2-V617F clones are present in most mutant-positive patients: a new disease paradigm. Blood 114 (14):3018–3023
8. Li S, Kralovics R, De Libero G, Theocharides A, Gisslinger H, Skoda RC (2008) Clonal heterogeneity in polycythemia vera patients with JAK2 exon12 and JAK2-V617F mutations. Blood 111(7):3863–3866
9. Delhommeau F, Dupont S, Della Valle V, James C, Trannoy S, Masse A et al (2009) Mutation in TET2 in myeloid cancers. N Engl J Med 360(22):2289–2301
10. Vannucchi AM, Guglielmelli P, Tefferi A (2009) Advances in understanding and management of myeloproliferative neoplasms. CA Cancer J Clin 59(3):171–191
11. Harrison CN, Campbell PJ, Buck G, Wheatley K, East CL, Bareford D et al (2005) Hydroxyurea compared with anagrelide in high-risk essential thrombocythemia. N Engl J Med 353 (1):33–45
12. Barosi G, Besses C, Birgegard G, Briere J, Cervantes F, Finazzi G et al (2007) A unified definition of clinical resistance/intolerance to hydroxyurea in essential thrombocythemia: results of a consensus process by an international working group. Leukemia 21(2):277–280
13. Akerblom L, Ehrenberg A, Graslund A, Lankinen H, Reichard P, Thelander L (1981) Overproduction of the free radical of ribonucleotide reductase in hydroxyurea-resistant mouse fibroblast 3T6 cells. Proc Natl Acad Sci USA 78(4):2159–2163
14. Rapado I, Albizua E, Ayala R, Hernandez JA, Garcia-Alonso L, Grande S et al (2008) Validity test study of JAK2 V617F and allele burden quantification in the diagnosis of myeloproliferative diseases. Ann Hematol 87(9):741–749
15. Rapado I, Grande S, Albizua E, Ayala R, Hernandez JA, Gallardo M et al (2009) High resolution melting analysis for JAK2 Exon 14 and Exon 12 mutations: a diagnostic tool for myeloproliferative neoplasms. J Mol Diagn 11(2):155–161
16. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA et al (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA 102(43):15545–15550
17. Pellagatti A, Vetrie D, Langford CF, Gama S, Eagleton H, Wainscoat JS et al (2003) Gene expression profiling in polycythemia vera using cDNA microarray technology. Cancer Res 63 (14):3940–3944
18. Kralovics R, Teo SS, Buser AS, Brutsche M, Tiedt R, Tichelli A et al (2005) Altered gene expression in myeloproliferative disorders correlates with activation of signaling by the V617F mutation of Jak2. Blood 106(10):3374–3376
19. Schwemmers S, Will B, Waller CF, Abdulkarim K, Johansson P, Andreasson B et al (2007) JAK2V617F-negative ET patients do not display constitutively active JAK/STAT signaling. Exp Hematol 35(11):1695
20. Shide K, Shimoda HK, Kumano T, Karube K, Kameda T, Takenaka K et al (2008) Development of ET, primary myelofibrosis and PV in mice expressing JAK2 V617F. Leukemia 22 (1):87–95
21. Chin H, Arai A, Wakao H, Kamiyama R, Miyasaka N, Miura O (1998) Lyn physically associates with the erythropoietin receptor and may play a role in activation of the Stat5 pathway. Blood 91 (10):3734–3745
22. Zaleskas VM, Krause DS, Lazarides K, Patel N, Hu Y, Li S et al (2006) Molecular pathogenesis and therapy of polycythemia induced in mice by JAK2 V617F. PLoS ONE 1:e18
23. Wappl M, Jaeger E, Streubel B, Gisslinger H, Schwarzinger I, Valent P et al (2008) Dasatinib inhibits progenitor cell proliferation from polycythaemia vera. Eur J Clin Investig 38(8):578–584
24. Girodon F, Schaeffer C, Cleyrat C, Mounier M, Lafont I, Santos FD et al (2008) Frequent reduction or absence of detection of the JAK2-mutated clone in JAK2V617F-positive patients within the first years of hydroxyurea therapy. Haematologica 93(11):1723–1727
25. Bock O, Neuse J, Hussein K, Brakensiek K, Buesche G, Buhr T et al (2006) Aberrant collagenase expression in chronic idiopathic myelofibrosis is related to the stage of disease but not to the JAK2 mutation status. Am J Pathol 169(2):471–481
26. Stetler-Stevenson WG, Bersch N, Golde DW (1992) Tissue inhibitor of metalloproteinase-2 (TIMP-2) has erythroid-potentiating activity. FEBS Lett 296(2):231–234

HDAC inhibitors induce cell cycle arrest, activate the apoptotic extrinsic pathway and synergize with a novel PIM inhibitor in Hodgkin lymphoma-derived cell lines

Molecular studies have shown that the tumoural component of classical Hodgkin lymphoma (cHL), Hodgkin and Reed-Sternberg (HRS) cells, shows a profound disturbance of the cell-cycle and apoptosis regulation (Küppers, 2009). Although cHL is often curable, there remains a group of patients who do not achieve remission with initial therapy and about 30% patients will die of the disease. Therefore, novel targeted agents need to be developed to reduce treatment-related toxicity and to further patient outcome.

Histone deacetylase inhibitors (HDACi) are new agents with demonstrable preclinical antitumour activity in a wide range of malignancies. However, their exact mechanism of action remains to be fully elucidated.

The rationale for evaluating HDACi in cHL is based on observations implicating epigenetics in the HRS-phenotype, on high levels of HDACs in approximately 80% cHL cases (Gloghini *et al*, 2009), and on the deregulation in cHL of pathways that can be restored by HDACi in other cancer types (Garcia *et al*, 2003).

In order to test drugs therapeutically useful in cHL, we used the Connectivity Map (Cmap) program and a gene expression signature from poor-prognosis cHL-patients, (Sanchez-Aguilera *et al*, 2006). Cmap revealed two HDACi, Trichostatin A and SAHA, to be the most specific drugs able to reverse the cHL-chemoresistance signature, suggesting that HDACi might be very efficient in cHL-treatment (Fig S1A).

To demonstrate HDACi efficiency as antitumoral agents in cHL, four cHL-derived cell lines were treated with two HDACi (suberoylanilide hydroxamic acid (SAHA) and sodium butyrate (SB)). HDACi significantly reduced cell viability in all cHL cell lines in a dose-dependent manner (Fig 1A), with different effects on cell-cycle in each cell line. Both HDACi induced a time-dependent G2/M phase arrest in L428 and KMH2 cell lines, similar to recent findings with SAHA (Buglio *et al*, 2008) and depsipeptide (Hartlapp *et al*, 2009); while HDLM2 and L1236 cell lines did not experience cell-cycle arrest (Figs 1B and S1B). Moreover, a significant and time-dependent increase in apoptosis in all cell lines was observed after HDACi-treatment (Fig 1C).

As L428 and KMH2 seemed to respond similarly to HDACi (G2/M arrest), while the response of HDLM2 and L1236 was different (no cell-cycle arrest), L428 was used as the sensitive cell line and HDLM2 as the resistant cell line in further experiments.

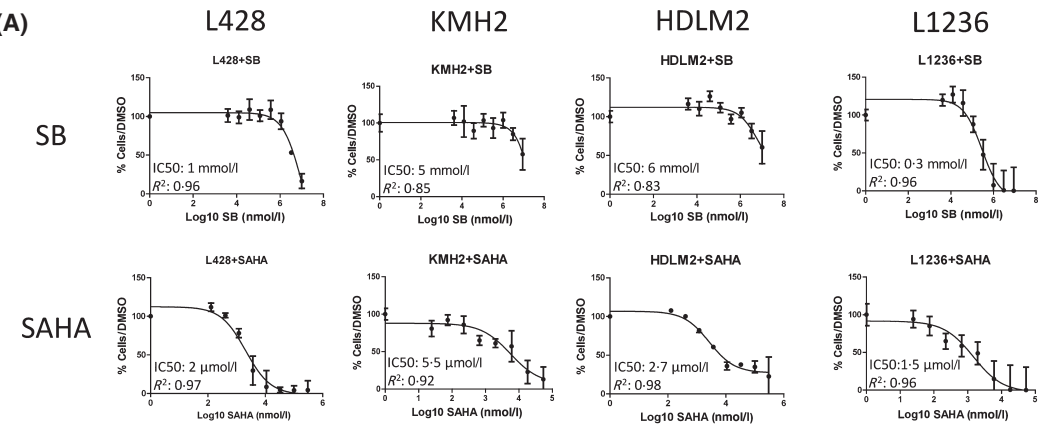
To ensure that HDACi were able to target HDACs we confirmed histones H3 and H4 hyperacetylation upon HDACi-treatment (Fig S2A).

The differential sensitivity observed in cHL-cell lines after HDACi-treatment could be explained by differences in basal status of acetylated histones and gene expression profile (Fig S2A,B), because the HDLM2 cell line, with a poor-prognosis molecular profile compared with L428, and a inactivating *TP53* mutation (Feuerborn *et al*, 2006), is more resistant to HDACi-apoptosis. Moreover, overexpression of antiapoptotic proteins, such as Bcl2, is known to be involved in resistance mechanisms to HDACi.

Since histone-acetylation is a well-known mechanism of gene expression regulation, we evaluated genes that showed modified expression after HDACi-treatment. Data obtained from three independent hybridizations for each time point were analysed by a *t*-test. This approach revealed deregulation of <5% of the whole genome in each HDACi-treated cHL-cell line (false discovery rate [FDR]<0·1) in agreement with most of studies performed in HDACi-treated cancer cell lines. Gene Set Enrichment Analysis (GSEA) identified several pathways altered by HDACi (Fig 1D). The *t*-test found 1946 differen-

Fig 1. Effect of HDACi on cHL-derived cell lines. L428, KMH2, HDLM2 and L1236 cell lines were treated with SB or SAHA at several doses and times, and (A) cell viability, (B) cell cycle, (C) apoptosis, and (D–E) gene expression changes were determined. (A) Cell viability was measured by ATP intracellular content and calculated as the percentage of DMSO-treated cells; 50% inhibitory concentration (IC50) values for 48 h are shown. (B) L428 and KMH2 cells were arrested at G2/M phase and underwent cell death in a time-dependent manner. HDLM2 and L1236 cells were not arrested, but cell death was induced after long treatments. Numbers indicate the percentage of HDACi-treated cells minus percentage of DMSO-treated cells (mean ± SEM) (*indicates statistically significant differences from the control, $P < 0{\cdot}05$). (C) Apoptosis was calculated as the percentage of non-viable cells (AnnexinV+/Propidium Iodide- and AnnexinV+/Propidium Iodide+ cells) in HDACi-treated cells minus the percentage of non-viable cells in DMSO-treated cells (mean ± SEM) (*indicates statistically significant differences from the control, $P < 0{\cdot}05$). (D) Altered pathways in L428 and HDLM2 cell lines after treatment with 1 mmol/l SB and 2 μmol/l SAHA at the indicated times (FDR < 0·25). Red, green and white indicate upregulation, downregulation and no significant deregulation, respectively. Intensity of colours indicates the degree of significance (dark: FDR < 0·05; intermediate: 0·05 < FDR < 0·1; light: 0·1 < FDR < 0·25). (E) Hierarchical clustering with all 1946 significantly deregulated genes (FDR < 0·1) in L428 and HDLM2 cell lines after treatment with 1 mmol/l SB and 2 μmol/l SAHA at the indicated times. Data were normalized with respect to untreated control cells. Highlighted genes indicate the mechanism of action of HDACi in cHL cell lines.

(A)



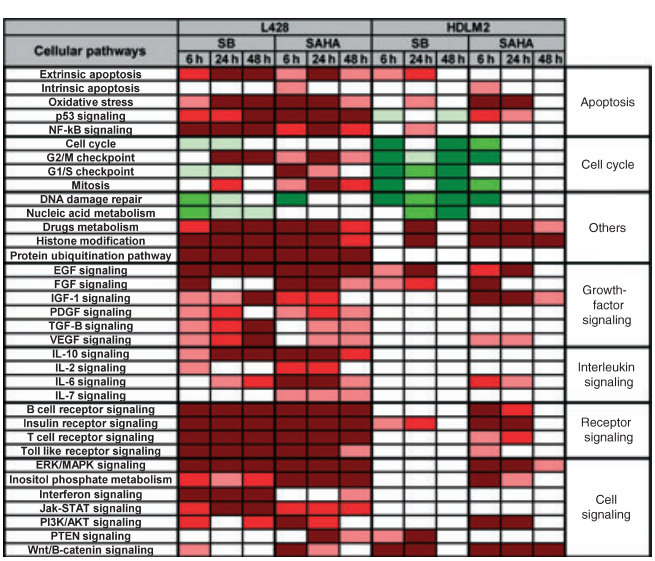
(B) HDACi-promoted cell cycle changes

	SB		SAHA	
	24 h	48 h	24 h	48 h
L428	SubG0: 2·4±0·8 G2/M: 8·9±1·5*	SubG0: 2·6±1·1 G2/M: 7·5±1·7*	SubG0: 3·2±0·7 G2/M: 17·6±2·6*	SubG0: 13·8±3·1* G2/M: 10·3±2·1*
KMH2	SubG0: 2·2±2·6 G2/M: 2·7±2·0	SubG0: 0·7±0·1* G2/M: 9·9±0·1*	SubG0: 5·1±0·5 G2/M: 22·9±0·6*	SubG0: 3·8±0·1* G2/M: 55·6±1·6*
HDLM2	SubG0: 4·0±1·5 G2/M:−2·0±1·4	SubG0: 12·5±2·6* G2/M:−4·3±2·2	SubG0: 5·2±0·7* G2/M:−0·6±0·9	SubG0: 6·1±2·4 G2/M:−0·7±1·6
L1236	SubG0: 21·5±6·3 G2/M:−2·6±0·0	SubG0: 14·3±0·9* G2/M:−4·2±0·5	SubG0: 19·6±3·8* G2/M:−1·1±1·8	SubG0: 48·6±1·3* G2/M:−10·4±0·2*

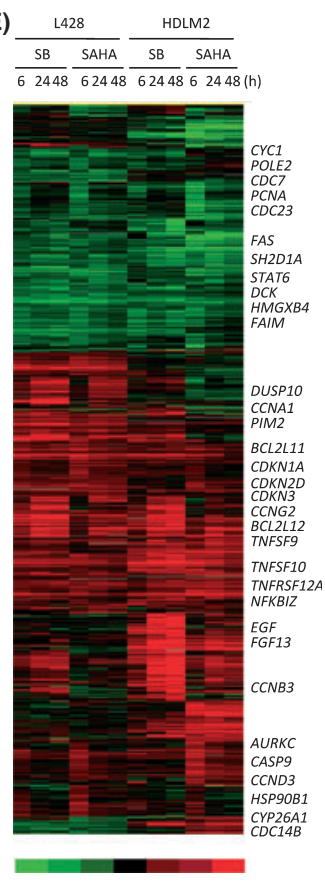
(C) HDACi-induced apoptosis

	SB			SAHA		
	6 h	24 h	48 h	6 h	24 h	48 h
L428	−0·2±0·8	6·1±1·0*	4·3±1·8*	−0·1±0·5	6·8±2·1*	13·6±4·3*
KMH2	1·4±2·4	−6·1±0·3	5·3±0·3	4·2±0·4*	14·5±0·6*	45·5±0·8*
HDLM2	0·2±0·5	4·5±1·6	16·6±1·2*	1·6±0·2	3·0±1·4*	7·2±0·8*
L1236	0·6±1·9	0·8±1·3	10·0±4·7*	1·2±0·1	5·3±0·0*	23·9±3·4*

(D)



(E)



tially expressed genes (FDR < 0·1) from 0 h to at least one subsequent time point in both cHL-cell lines after treatment (Fig 1E). In accordance with the functional analysis, a number of cell cycle (*CDKN1A, CDKN2D)* and apoptosis genes (*TNFSF9, TNFSF10, TNFRSF12A, BCL2L11* and *BCL2L12)* were significantly deregulated. A number of genes associated with chemoresistance in cHL were consistently downregulated by HDACi (Fig 1E). This supports that HDACi could be added to current chemotherapy protocols, given that a number of evidences have indicated the benefits of the combined treatment (Bhalla *et al*, 2009). The complete list of significant genes is available in Table S1.

GSEA analyses identified apoptosis as one of the most important HDACi-modified pathways. To confirm these alterations, we measured the key proteins involved in apoptosis. Consistent with the results of our GSEA, we observed subtle HDACi-effects on Bax levels (intrinsic apoptosis) and a release of Cytochrome-c (Fig S2C). Moreover, there was a decrease in ProCaspase-8 and Fas levels, indicating extrinsic apoptotic pathway activation. Surprisingly and in contrast with many other tumour entities, HRS cells strongly expressed Fas, activating nuclear factor (NF)-kB transcription factor (Mathas *et al*, 2004), and suggesting the importance of death receptor apoptosis signalling in HRS cells. Furthermore, the extrinsic apoptotic pathway, alone or in conjunction with the intrinsic pathway, has already been reported to be induced by SAHA in other cell types (Gillenwater *et al*, 2007).

Unexpectedly, we found a strong upregulation of the proto-oncogene *PIM2* in both HDACi-treated cHL-cell lines (Figs 1E and 2A). Moreover, we have previously observed significant upregulation of *PIM1* (FDR = 0·03) and *PIM2* (FDR = 0·001) in a series of 29 cHL-patients (Fig S3A) (published data from Sanchez-Aguilera *et al*, 2006). These findings prompted us to hypothesize that HDACi effects could benefit from *PIM2* inhibition. We performed cell viability assays with a novel PIM inhibitor (PIMi) and found it to be efficient in four cHL-cell lines (Fig 2B). PIMi efficiency was checked by measuring its target pBAD(Ser112) (Fig S3B).

Finally, we performed experiments with HDACi and PIMi in cHL-cell lines. The combination SB+PIMi was as efficient as each single drug in all cell lines but HDLM2, where the effect was synergistic. However, the combination SAHA+PIMi enhanced the reduction in cell viability in a synergistic manner in all four of the cell lines (Combination Index = 0·19, 0·57, 0·73 and 0·83 for HDLM2, L428, KMH2 and L1236, respectively) (Fig 2C). This finding highlights the importance of SAHA+PIMi combination in cHL therapeutics. Further studies will be needed to fully understand why *PIM2* is upregulated after HDACi-treatment, the underlying mechanisms of this drug combination, and the role of *PIM1* and *PIM2* in cHL.

In conclusion, we have demonstrated that HDACi exert antiproliferative and proapoptotic effects on cHL-cell lines by altering the expression of cell-cycle and apoptosis-genes, with common and different (cell line-dependent) effects. Additionally, HDACi can downregulate the expression of genes associated with chemoresistance in cHL-patients. Importantly, SAHA acts in a highly synergistic manner with a novel

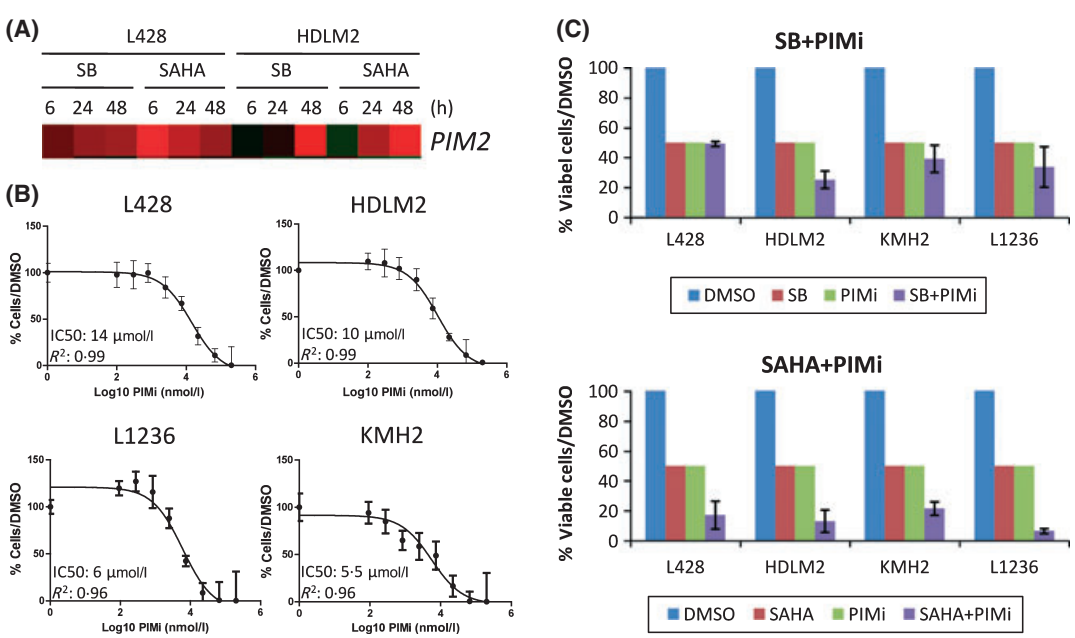


Fig 2. HDACi have a synergistic effect with a PIM inhibitor. (A) After treatment with HDACi, a significant upregulation of *PIM2* was observed in L428 and HDLM2 cell lines. The expression value of *PIM2* was normalized with respect to control untreated cells. (B) Four cHL-derived cell lines were treated with a novel PIMi for 48 h, and IC50 values were calculated by MTT (3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) assay. (C) Cell lines were treated simultaneously with the PIMi and with HDACi (SB and SAHA) for 48 h, using IC50 values of each drug in each cell line, and cell viability was measured by MTT assay. Data are representative of two independent experiments.

PIMi in cHL-cell lines. Collectively, our results provide the proof-
of-principle that justifies the investigation of HDACi in combination with PIM inhibitors for cHL-treatment.

Acknowledgements

The authors would like to thank to Antonia Alvarez (Flow Cytometry and Confocal Microscopy Service, University of the Basque Country) for her technical assistance with the flow cytometer; and to Gonzalo Gomez (Bioinformatics Unit, Spanish National Cancer Research Centre) for his bioinformatic help.

Conflict of interest

The authors declare no competing financial interests.

Financial support

This work was supported by grants from RTICC (RD/06/0020/0048), the Basque Government (GIC07/61-IT-463-07), Ministerio de Sanidad y Consumo (FIS, PI08/1985) y Fundación Mutua Madrileña. EMS is supported by a grant from the Department of Education, Universities and Research of the Basque Government (BFI08·207).

Authorship

EMS performed experiments, analysed and interpreted the data and wrote the manuscript; MSB interpreted the data and revised the manuscript; MER and CGA performed experiments; BSE analysed data; JRB contributed vital reagents for the development of the novel PIM inhibitor; MAP revised the manuscript; AGO designed the research and revised the manuscript; JFG designed the research, interpreted the data and revised the manuscript.

Esperanza Martín-Sánchez[1,2]
Margarita Sánchez-Beato[1]
María Elena Rodríguez[1]
Beatriz Sánchez-Espiridión[1]
Cristina Gómez-Abad[1]
James R. Bischoff[3]
Miguel A. Piris[1]
África García-Orad[2]
Juan F. García[1,4]

[1]Lymphoma Group, Molecular Pathology Programme, Spanish National Cancer Research Centre (CNIO), Madrid, [2]Department of Genetics, Physical Anthropology and Animal Physiology, University of the Basque Country, Bilbao, [3]Experimental Therapeutics Programme, Spanish National Cancer Research Centre (CNIO), and [4]Department of Pathology, MD Anderson International, Madrid, Spain.
E-mail: jfgarcia@mdanderson.es

Supporting information

Additional Supporting information may be found in the online version of this article:

Figure S1. (A) Output from Connectivity Map (Cmap). Top ten pharmaceutical perturbagens identified through the Cmap that reverse cHL-chemoresistance signature (with negative enrichment score) in the panel of cell lines included in the software. Trichostatin A and SAHA, also known as vorinostat, belong to the HDACi family. (B) Cell-cycle changes induced by HDACi in cHL-derived cell lines. HDACi induced a time- and dose-dependent G2/M arrest and cell death in L428 and KMH2 cells, while no cell cycle arrest but cell death after high doses/long treatments was induced in HDLM2 and L1236 cells. Numbers indicate mean ± SEM of at least three independent experiments.

Figure S2. Molecular alterations induced by HDACi in cHL-derived cell lines. (A) Effect of HDACi on histone acetylation: L428 and HDLM2 cell lines were treated with SB (1 mmol/l) and SAHA (2 μmol/l) at the indicated times, and levels of acetylated and total histones H3 and H4 were measured by western blot. Numbers represent the ratio of acetylated/total histones. (B) Basal differentially expressed genes between L428 and HDLM2 cell lines (FDR < 0·05). Highlighted genes could explain the different sensitivity and response to HDACi treatment. (C) Effect of HDACi on key proteins involved in apoptosis: L428 and HDLM2 cell lines were treated with SB (1 mmol/l) and SAHA (2 μmol/l) at the indicated times, and levels of whole-cell lysed proteins were measured by western blot. Numbers represent the ratio of measurements of each protein relative to α-tubulin levels.

Figure S3. (A) *PIM1* and *PIM2* expression levels in cHL. Two sorted centroblasts samples, five lymph nodes and a series of 29 cHL cases (14 with favourable outcome and 15 with unfavourable outcome) were profiled for *PIM1* and *PIM2* expression levels. (B) Levels of pBAD(Ser112) and BAD were measured after treatment with IC50 values of PIMi in cHL-cell lines.

Table S1. Expression values normalized with respect to untreated control cHL cell lines and a brief description extracted from Feature Extraction software of all 1946 significantly deregulated genes identified by a *t*-test (FDR < 0·1) in L428 and HDLM2 cell lines after treatment with 1 mmol/l SB and 2 μmol/l SAHA at the indicated times.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Keywords: HDAC inhibitors, Hodgkin lymphoma, gene expression, extrinsic apoptosis, PIM inhibitors.

First published online 19 October 2010
doi:10.1111/j.1365-2141.2010.08401.x

References

Bhalla, S., Balasubramanian, S., David, K., Sirisawad, M., Buggy, J., Mauro, L., Prachand, S., Miller, R., Gordon, L.I. & Evens, A.M. (2009) PCI-24781 induces caspase and reactive oxygen species-dependent apoptosis through NF-kappaB mechanisms and is synergistic with bortezomib in lymphoma cells. *Clinical Cancer Research*, **15,** 3354–3365.

Buglio, D., Georgakis, G.V., Hanabuchi, S., Arima, K., Khaskhely, N.M., Liu, Y.J. & Younes, A. (2008) Vorinostat inhibits STAT6-mediated TH2 cytokine and TARC production and induces cell death in Hodgkin lymphoma cell lines. *Blood*, **112,** 1424–1433.

Feuerborn, A., Moritz, C., Von Bonin, F., Dobbelstein, M., Trumper, L., Sturzenhofecker, B. & Kube, D. (2006) Dysfunctional p53 deletion mutants in cell lines derived from Hodgkin's lymphoma. *Leukemia and Lymphoma*, **47,** 1932–1940.

Garcia, J.F., Camacho, F.I., Morente, M., Fraga, M., Montalban, C., Alvaro, T., Bellas, C., Castano, A., Diez, A., Flores, T., Martin, C., Martinez, M.A., Mazorra, F., Menarguez, J., Mestre, M.J., Mollejo, M., Saez, A.I., Sanchez, L. & Piris, M.A. (2003) Hodgkin and Reed-Sternberg cells harbor alterations in the major tumor suppressor pathways and cell-cycle checkpoints: analyses using tissue microarrays. *Blood*, **101,** 681–689.

Gillenwater, A.M., Zhong, M. & Lotan, R. (2007) Histone deacetylase inhibitor suberoylanilide hydroxamic acid induces apoptosis through both mitochondrial and Fas (Cd95) signaling in head and neck squamous carcinoma cells. *Molecular Cancer Therapeutics*, **6,** 2967–2975.

Gloghini, A., Buglio, D., Khaskhely, N.M., Georgakis, G., Orlowski, R.Z., Neelapu, S.S., Carbone, A. & Younes, A. (2009) Expression of histone deacetylases in lymphoma: implication for the development of selective inhibitors. *British Journal of Haematology*, **147,** 515–525.

Hartlapp, I., Pallasch, C., Weibert, G., Kemkers, A., Hummel, M. & Re, D. (2009) Depsipeptide induces cell death in Hodgkin lymphoma-derived cell lines. *Leukemia Research*, **33,** 929–936.

Küppers, R. (2009) The biology of Hodgkin's lymphoma. *Nature Reviews Cancer*, **9,** 15–27.

Mathas, S., Lietz, A., Anagnostopoulos, I., Hummel, F., Wiesner, B., Janz, M., Jundt, F., Hirsch, B., Johrens-Leder, K., Vornlocher, H.P., Bommert, K., Stein, H. & Dorken, B. (2004) c-FLIP mediates resistance of Hodgkin/Reed-Sternberg cells to death receptor-induced apoptosis. *The Journal of Experimental Medicine*, **199,** 1041–1052.

Sanchez-Aguilera, A., Montalban, C., de la Cueva, P., Sanchez-Verde, L., Morente, M.M., Garcia-Cosio, M., Garcia-Larana, J., Bellas, C., Provencio, M., Romagosa, V., de Sevilla, A.F., Menarguez, J., Sabin, P., Mestre, M.J., Mendez, M., Fresno, M.F., Nicolas, C., Piris, M.A. & Garcia, J.F. (2006) Tumor microenvironment and mitotic checkpoint are key factors in the outcome of classic Hodgkin lymphoma. *Blood*, **108,** 662–668.

CHIC cells: a novel ALK+ cell line derived from a relapsed anaplastic large cell lymphoma

Since the initial description of anaplastic large cell lymphoma (ALCL) as a proliferation of large CD30+ lymphoid cells, the morphological spectrum of ALCL positive for anaplastic lymphoma kinase (ALCL, ALK+) has expanded, and beyond the common pattern most frequently encountered, several variants have been identified, including the lymphohistiocytic and the small cell patterns (Swerdlow *et al*, 2008). Recent data suggest that tumours with variant morphology may carry a worse prognosis than the common pattern, with a tendency to more frequent and early relapses (Lamant *et al*, 2007).

Only ten ALK+ ALCL cell lines are currently available, and most were derived from tumours demonstrating the common type morphology (Drexler & MacLeod, 2004). We have established a novel cell line (CHIC) from the cerebrospinal fluid of a 32-year-old man with relapsing/refractory ALK+ ALCL whose initial tumour exhibited lymphohistiocytic features (Table I, Fig 1A–C).

The cell line was derived with no feeder layer or addition of growth factors and, in contrast to others, did not require prior passage into immunodeficient animals. CHIC cells proliferated in permanent uninterrupted suspension culture for more than 12 months as single non-attached cells or in loose clumps. They were frozen and stored in liquid nitrogen for 1 year, and recovered after thawing without any loss of cellular features or cell viability.

CHIC cells are large and pleomorphic, morphologically and immunophenotypically similar to the patient's malignancy (Table I, Fig 1D–E), showing positivity for CD30 and EMA, with nuclear and cytoplasmic expression of p-ALK. The cells are negative for T-cell associated markers, express CD43 and have a cytotoxic phenotype (perforin+, granzyme B+, TIA-1+). Expanded immunophenotyping demonstrated the expression of classical activation markers, such as CD25, CD26, CD45RO and HLA-DR. *In situ* hybridization confirmed the absence of Epstein-Barr virus (EBV). Cytogenetics revealed a complex karyotype including the t(2;5)(p23;q35) translocation, inducing *NPM1-ALK* fusion. The same *TRG* gene rearrangement was evidenced in CHIC cells, in xenografts (see below) and in the patient's tumour, demonstrating their T-cell origin and their derivation from one common clone.

Subcutaneous xenografts of CHIC cells were tumourigenic to non-obese diabetic severe combined immunodeficiency (NOD/SCID) mice. The tumours growing at the inoculation sites were histologically and immunophenotypically similar to the patient's lymphoma and the cell line (Table I, Fig 1F–H). Autopsy showed no evidence of metastases. This xenograft mouse model was optimized by the use of Matrigel, enhancing engraftment of CHIC cells and allowing the development of tumours in a reasonable time frame (62% of mice had developed tumours at 2 weeks after injection and 81% at 8 weeks).