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

Beatriz Sánchez-Espiridión, Abel Sánchez-Aguilera, Carlos Montalbán, Carmen Martín, Rafael Martínez, Joaquín González-Carrero, Concepción Poderos, Carmen Bellas, Manuel F. Fresno, Cesar Morante, María J. Mestre, Miguel Mendez, Francisco Mazorra, Eulogio Conde, Ángel Castaño, Pedro Sánchez-Godoy, José F. Tomás, Manolo M. Morente, Miguel A. Piris, and Juan F. García 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 G2-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 TH2 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...
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.

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 TLDAs 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 tumoral 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 (Bio; Santa Cruz Biotechnology); anti-CCNH (Cell Signaling Technology); and anti-HIST1H and anti-HIST2A (Upstate Biotechnology). These markers were chosen as representative of the different signatures and for the availability of reliable antibodies for paraffin-embedded tissue.
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 and other public sources available through the Molecular Signature Database 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.

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 linear 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: 18 www.biocarta.com
19 http://www.broad.mit.edu/gsea/msigdb/index.jsp
A Landatenth of that volume (5 mL 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 TLDAAs included HMBS, GUSB, TRP, 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 χ² test.

---

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 positive cHL samples for each marker expressed by the HRS cells of the microenvironment.
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 ($\chi^2$ 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 \( \leq 0.15 \)). 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 |

<table>
<thead>
<tr>
<th>Gene-sets enriched in TDB</th>
<th>Normalized enrichment score</th>
<th>Nominal probability</th>
<th>Genes included in TLD assays</th>
</tr>
</thead>
<tbody>
<tr>
<td>G2-M pathway</td>
<td>-1.67</td>
<td>0.016</td>
<td>AURKA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>BUB1B</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>BUB3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CCNH</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CDC2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CDC3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CDC6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CCNA2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CCNE2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CCNH</td>
</tr>
<tr>
<td>Chaperone pathway</td>
<td>-1.24</td>
<td>0.269</td>
<td>DNAJA2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>HSPA4</td>
</tr>
<tr>
<td>Drug resistance pathway</td>
<td>-1.23</td>
<td>0.204</td>
<td>DCK</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MLH1</td>
</tr>
<tr>
<td>Mitogen-activated protein kinase pathway</td>
<td>-0.83</td>
<td>0.736</td>
<td>GRB2</td>
</tr>
<tr>
<td>Mitochondrial pathway</td>
<td>-0.77</td>
<td>0.883</td>
<td>BCL2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CYCS</td>
</tr>
<tr>
<td>Gene-sets enriched in MDB</td>
<td>-1.50</td>
<td>0.063</td>
<td>CD3D</td>
</tr>
<tr>
<td>T-cell pathway</td>
<td></td>
<td></td>
<td>Sh2D1A</td>
</tr>
<tr>
<td>Monocyte/macrophage</td>
<td>-1.39</td>
<td>0.161</td>
<td>ALDH1A1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>LYZ</td>
</tr>
</tbody>
</table>

**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 TLD assay.
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 (G2-M and G1-S) that were significantly enriched at nominal P values <0.05, as indicated by the receiver operating characteristic curve, in agreement with previous results (24). Also, G1 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 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 (χ² = 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
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 TDA 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

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>Gene name</th>
<th>Area under receiver operating characteristic curve</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCCIP</td>
<td>BRCA2 and CDKN1A interacting protein</td>
<td>0.756</td>
<td>0.008</td>
</tr>
<tr>
<td>CASP3</td>
<td>Caspase-3, apoptosis-related cysteine peptidase</td>
<td>0.765</td>
<td>0.006</td>
</tr>
<tr>
<td>CCNE2</td>
<td>Cyclin E2</td>
<td>0.656</td>
<td>0.051</td>
</tr>
<tr>
<td>CSEL1</td>
<td>CSE1 chromosome segregation 1-like (yeast)</td>
<td>0.657</td>
<td>0.050</td>
</tr>
<tr>
<td>CTSL</td>
<td>Cathepsin L</td>
<td>0.678</td>
<td>0.046</td>
</tr>
<tr>
<td>CYCS</td>
<td>Cytochrome c, somatic</td>
<td>0.753</td>
<td>0.009</td>
</tr>
<tr>
<td>DCK</td>
<td>Deoxycytidine kinase</td>
<td>0.747</td>
<td>0.010</td>
</tr>
<tr>
<td>DNAJ1A2</td>
<td>DNAJ (Hsp40) homologue, subfamily A, member 2</td>
<td>0.671</td>
<td>0.049</td>
</tr>
<tr>
<td>HSPA4</td>
<td>Heat shock protein 90 kDa α (cytosolic), class A member 1</td>
<td>0.662</td>
<td>0.050</td>
</tr>
<tr>
<td>ITGA4</td>
<td>Integrin, α4 (antigen CD49D, α4 subunit of VLA-4 receptor)</td>
<td>0.659</td>
<td>0.050</td>
</tr>
<tr>
<td>LYZ</td>
<td>Lysozyme (renal amyloidosis)</td>
<td>0.674</td>
<td>0.047</td>
</tr>
<tr>
<td>RSN</td>
<td>Restin (Reed-Sternberg cell-expressed</td>
<td>0.706</td>
<td>0.033</td>
</tr>
<tr>
<td></td>
<td>intermediate filament-associated protein</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TYMS</td>
<td>Thymidylate synthetase</td>
<td>0.750</td>
<td>0.010</td>
</tr>
</tbody>
</table>
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 G2-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 also probably 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 (Fundoación Hospital Alcorcón); M.J. Mestre, R. Quibé, 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. Provenzio (Hospital Puerta de Hierro); A. Castaño and P. Sánchez-Goody (Hospital Severo Ochoa); C. Martín and R. Martínez (Hospital Clínico Universitario San Carlos); J. Menárguez, P. Sabin, 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’Onocolgia); M. Mollejo and M.A. Cruz (Hospital Virgen de la Salud); H. Álvarez-Arguelles and M. Llanos (Hospital Universitario Canarias); C. Morané (Hospital Cabués); F. Mazarro 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 Alzare Buylas); 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

9. Hasenclever D, Diehl V. A prognostic score for ad-


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


Updated version

Access the most recent version of this article at:
http://clincancerres.aacrjournals.org/content/15/4/1367

Supplementary Material

Access the most recent supplemental material at:
http://clincancerres.aacrjournals.org/content/suppl/2009/02/16/15.4.1367.DC1

Cited articles

This article cites 57 articles, 29 of which you can access for free at:
http://clincancerres.aacrjournals.org/content/15/4/1367.full#ref-list-1

Citing articles

This article has been cited by 7 HighWire-hosted articles. Access the articles at:
http://clincancerres.aacrjournals.org/content/15/4/1367.full#related-urls

E-mail alerts

Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions

To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions

To request permission to re-use all or part of this article, use this link
http://clincancerres.aacrjournals.org/content/15/4/1367.
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.