A three-gene signature based on MYC, BCL-2 and NFKBIA improves risk stratification in diffuse large B-cell lymphoma

Enrico Derenzini,1,2 Saveria Mazzara,3 Federica Melle,3 Giovanna Motta,3 Marco Fabbri,3 Riccardo Bruna,3 Claudio Agostinelli,4 Alessandra Cesano,5 Chiara Antonia Corsini,5 Ning Chen,5 Simona Righi,4 Elena Sabattini,4 Annalisa Chiappella,7 Angelica Calleri,3 Stefano Fiori,3 Valentina Tabanelli,3 Antonello Cabras,6 Giancarlo Pruneri,8 Umberto Vitolo,9 Alessandro Massimo Gianni,3 Alessandro Rambaldi,10 Paolo Corradini,7 Pier Luigi Zinzani,11 Corrado Tarella1,2 and Stefano Pileri3

1Onco-Hematology Division, IEO European Institute of Oncology IRCCS, Milan and European Institute of Oncology IRCCS, Milan, Italy; 2Department of Health Sciences, University of Milan, Milan, Italy; 3Division of Diagnostic Hematopathology, IEO European Institute of Oncology IRCCS, Milan and European Institute of Oncology IRCCS, Milan, Italy; 4Hematopathology Unit, Department of Experimental, Diagnostic, and Specialty Medicine (DIMES), Bologna University School of Medicine, Bologna, Italy; 5NanoString Technologies Inc, Seattle, WA, USA; 6Laboratory of Hematology-Oncology, IEO European Institute of Oncology IRCCS, Milan and European Institute of Oncology IRCCS, Milan, Italy; 7Division of Hematology, Fondazione IRCCS Istituto Nazionale dei Tumori, University of Milan, Milan, Italy; 8Department of Pathology, Fondazione IRCCS Istituto Nazionale dei Tumori di Milano, Milan, Italy; 9Multidisciplinary Oncology Outpatient Clinic, Candido Cancer Institute, FPO-IRCCS, Candido, Italy; 10Hematology and Bone Marrow Transplant Unit, ASST-Papa Giovanni XXIII, Bergamo, Italy and 11Institute of Hematology and Medical Oncology “L. e A. Seragnoli”, Department of Experimental, Diagnostic, and Specialty Medicine (DIMES), Bologna University School of Medicine, Bologna, Italy

ABSTRACT

Recent randomized trials focused on gene expression-based determination of the cell of origin in diffuse large B-cell lymphoma could not show significant improvements by adding novel agents to standard chemoimmunotherapy. The aim of this study was the identification of a gene signature able to refine current prognostication algorithms and applicable to clinical practice. Here we used a targeted gene expression profiling panel combining the Lymph2Cx signature for cell of origin classification with additional targets including MYC, BCL-2 and NFKBIA, in 186 patients from two randomized trials (discovery cohort) (clinicaltrials.gov. Identifier: NCT00355199 and NCT00499018). Data were validated in three independent series (two large public datasets and a real-life cohort). By integrating the cell of origin, MYC/BCL-2 double expressor status and NFKBIA expression, we defined a three-gene signature combining MYC, BCL-2 and NFKBIA (MBN-signature), which outperformed the MYC/BCL-2 double expressor status in multivariate analysis, and allowed further risk stratification within the germinal center B-cell/unclassified subset. The high-risk (MBN Sig-high) subgroup identified the vast majority of double hit cases and a significant fraction of activated B-cell-derived diffuse large B-cell lymphomas. These results were validated in three independent series including a cohort from the REMoDL-B trial, where, in an exploratory ad hoc analysis, the addition of bortezomib in the MBN Sig-high subgroup provided a progression free survival advantage compared with standard chemoimmunotherapy. These data indicate that a simple three-gene signature based on MYC, BCL-2 and NFKBIA could refine the prognostic stratification in diffuse large B-cell lymphoma, and might be the basis for future precision-therapy approaches.
Introduction

The biologic complexity of diffuse large B-cell lymphoma (DLBCL) was first dissected in the early 2000s by gene expression profiling (GEP) studies, which subdivided DLBCL into two groups based on GEP signatures reminiscent of the respective cell of origin (COO). These studies showed that DLBCL with a gene signature related to activated B lymphocytes (ABC subgroup) had a significantly worse response to anthracycline-based therapies compared to those histogenetically related to germinal center B cells (GCB subtype), and were dependent on nuclear factor κ-B (NF-κB) signaling. Since immunohistochemical algorithms failed to reproduce the results of GEP, the Lymphoma Leukemia Molecular Profiling Project (LLMPP) proposed a targeted GEP (T-GEP) panel (Lymph2Cx) derived from previous studies on fresh/frozen tissue (FFT). This assay was applied on the NanoString platform from formalin-fixed, paraffin embedded (FFPE) tissue from DLBCL patients treated with R-CHOP, identifying three subgroups: GCB, ABC and unclassified, the latter representing about 15% of all cases and prognostically close to the GCB. The reproducibility of this assay was confirmed in several studies. However, recent results from three independent phase III randomized trials based on COO classification were largely negative. Although these unsatisfactory results could be due to several reasons, including unexpected toxicities and suboptimal efficacy of these drugs in vivo, these data also indicate that the clinical development of predictive T-GEP signatures able to complement the COO for precision therapy approaches is an urgent unmet need. Besides the COO, current evidence indicates a negative prognostic value of double MYC and BCL-2 protein overexpression determined by immunohistochemistry (IHC). Furthermore, those DLBCL with concurrent MYC and BCL-2 and/or BCL-6 genomic rearrangements are characterized by an even worse prognosis, being now classified as a separate entity, high-grade B-cell lymphoma (HG-BCL) with double/triple hits (w DH/TH). Recently large genomic studies integrating DNA and RNA sequencing data identified additional DLBCL subgroups beyond the COO and MYC/BCL-2 double expressor (DE) status, based on the mutational landscape, GEP signatures, copy number changes, and differences in outcome. Furthermore, recent studies identified GEP signatures able to define high-risk populations within the GCB/unclassified (GCB/U) subgroup. However, given their complexity, large-scale application of these prognostication algorithms could be difficult in daily clinical practice. The aim of this study was the implementation of a simple T-GEP panel able to complement and improve COO-based prognostic stratification for routine clinical application. We designed a panel of genes corresponding to those of the Lymph2Cx assay for COO determination plus additional candidates selected because of their potential prognostic and/or therapeutic interest including MYC, BCL-2 and central nodes of NF-κB, Janus kinase (JAK)/signal transducer and activator of transcription (STAT), and phosphatidylinositol-3-kinase (PI3K) signaling. This panel of genes was applied to 186 DLBCL enrolled in two recently reported large Italian trials (DLCL04 and R-HDS0305; clinicaltrials.gov. Identifier: NCT00355199 and NCT00499018). We found that a three-gene signature based on MYC, BCL-2 and NFKBIA (MBN signature), identified a significant fraction of ABC cases and a subgroup of GCB/U cases (roughly 30%) enriched in HG-BCL w/DH, at increased risk of treatment failure. These data were validated in a real-life cohort in vitro, where the addition of bortezomib to chemoimmunotherapy provided a significant advantage for high-risk patients identified by the MBN signature.

Methods

Study design

Patients considered in this study had been enrolled in two prospective randomized phase III clinical trials investigating the role of first line autologous stem cell transplant (ASCT) consolidation in intermediate/high-risk DLBCL. Only cases of DLBCL not-otherwise specified (NOS) (including those originally diagnosed as DLBCL and nowadays included in the HG-BCL provisional category) were selected for the present study (Figure 1). Patients’ characteristics and study algorithm are summarized in Table 1 and Figure 1. Results were validated in three independent cohorts, (two in silico validation datasets and one “real-life” cohort): a dataset from Sha and coworkers (n=928 patients: 469 treated with R-CHOP and 459 with R-CHOP plus bortezomib [RB-CHOP]); a public dataset from Lenz et al. (n=233 patients treated with R-CHOP); a “real-life” cohort including 102 consecutive DLBCL-NOS cases with available FFPE tissue, treated with R-CHOP/R-CHOP-like regimens in Bologna (S.Orsola-Malpighi Hospital), and in Milan (European Institute of Oncology) from 2007 to 2018.

This study was approved by the Institutional Review Boards and Ethics Committees of the participating centers, in accordance with the Declaration of Helsinki.

Procedures

Gene expression was measured on the NanoString nCounter Analysis System (NanoString Technologies, Seattle, USA). The T-GEP panel contains 26 genes: 15 genes for COO subtyping; five housekeeping genes (UBXN4, ISY1, R3HDM1, WDR55, TRIM16); and six additional genes (MYC, BCL-2, STAT3, NFKBIA, PTEN, PIK3CA). Besides MYC and BCL-2, the additional genes were selected based on their known functions in key pathways involved in DLBCL lymphomagenesis and potential druggability.

Statistical analysis

Survival data were analyzed retrospectively. We used Kaplan-Meier method for overall survival (OS) and progression-free survival (PFS) analyses. Multivariate and univariate analyses were constructed with the Cox proportional hazards regression model. A P-value ≤ 0.05 was considered statistically significant. Recursive Partitioning Analysis (RPA) was applied to classify patients into more homogenous prognostic groups based on survival. All analyses were performed using R 3.5.0 software. Correlations and differences in patient characteristics were analyzed with the χ² and Fisher’s exact test.

Development of the three-gene prognostic signature (MBN signature)

An expression ratio-based test was developed by selecting those genes significantly deregulated in the high risk subgroups identified by the RPA shown in Figure 2A and whose normalized mRNA levels were significantly associated with OS. We defined high and low MYC and BCL-2 expressors based on the median normalized MYC and BCL-2 mRNA levels. The high-risk groups
included the ABC and double expressor GCB/unclassified (GCB/U) DLBCL (hereafter defined as DEXP_mRNA); the low risk group was constituted by the non-DEXP_mRNA GCB/U subset. Since the expression levels of MYC and BCL-2 on one hand and NFKBIA on the other hand had opposing patterns being inversely associated with OS (with higher MYC/BCL-2 and lower NFKBIA levels associated with worse outcome), we combined the expression levels of the three genes in a synthetic predictor called MBN-Signature (MBN-Sig) and defined as:

\[
MBN-Sig = \frac{(MYC + BCL-2)}{NFKBIA}
\]

Detailed information on study cohorts (Online Supplementary Table S1), T-GEP procedures with list of genes and target sequences, fluorescence in situ hybridization (FISH), IHC methods and antibodies (Online Supplementary Table S2), and random forest (RF) classifier are described in the Online Supplementary Appendix.

**Results**

**Univariate analyses and a decision-tree classification model integrating the cell of origin and MYC/BCL-2 status**

Given their established clinical relevance, we first investigated the prognostic significance of T-GEP-based COO classification and MYC/BCL-2 status in the R-HDS0305 and DLCL04 trials34,35 (discovery cohort). Patient’s characteristics are summarized in Table 1. In line with previous findings11,12 COO classification by T-GEP clearly outperformed the immunohistochemical Hans algorithm for survival prediction and retained its prognostic significance in the presence or absence of ASCT consolidation (Online Supplementary Figure S1A to D). In order to investigate the prognostic impact of concurrent overexpression of MYC and BCL-2, we defined high and low expressors based on

![Diagram](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE10846)
the median normalized MYC and BCL-2 mRNA levels, which correlated well with the respective protein levels assessed by IHC (Online Supplementary Figure S2A). MYC/BCL-2 mRNA double expressors (defined as DEXP_mRNA) patients showed a worse outcome compared to non-DEXP_mRNA cases (Online Supplementary Figure S2B). Although DEXP_mRNA cases were more prevalent in the ABC compared to the GCB/unclassified (GCB/U) subgroup (Online Supplementary Table S3), the prognostic relevance of the MYC/BCL-2 DEXP_mRNA status was particularly evident in the GCB/U subset (Online Supplementary Figure S2C to F). Focusing the analysis on the additional genes (STAT3, NFKBIA, PTEN, PIK3CA), which were selected based on their biologic relevance in potentially druggable pathways, only NFKBIA and STAT3 mRNA levels were significantly associated with patient’s outcome, with low STAT3 and low NFKBIA expression predicting worse prognosis (Online Supplementary Figure S2C to F).
In univariate analyses only the age adjusted International Prognostic Index (aaIPI) score (intermediate-high vs. high), the COO classification, MYC/BCL-2-DE status, NFKBIA and STAT3 levels determined by T-GEP, were significantly associated with OS (Table 2). As observed in the original studies, first-line ASCT consolidation was not associated with patient’s outcome.

In line with the data presented above (Online Supplementary Figure S1 and S2), a recursive partitioning analysis integrating the COO with MYC/BCL-2 status identified three main patient subgroups: two high risk subsets with similar outcome (ABC [n=40] and MYC/BCL-2 DEXP_mRNA GCB/U [n=27]) and a low-risk subgroup including non-DEXP mRNA GCB/U DLBCL, (n=119) (Figure 2A). Evaluating the relative expression of the additional genes included in the panel across the three groups identified by the recursive partitioning analysis (non-DEXP_mRNA GCB/U, DEXP_mRNA GCB/U and ABC DLBCL patients) (Figure 2B), we found that only MYC, BCL-2 and NFKBIA were significantly deregulated in both the high-risk ABC and MYC/BCL-2 DEXP_mRNA GCB/U subgroups, which were characterized by similarly increased MYC and BCL-2 and lower NFKBIA mRNA levels compared to the low risk non-DEMP mRNA GCB/U subset. The NFKBIA gene, a frequent target of deletions and mutations in DLBCL, encodes for the Ikβ-α protein, which is a central node of the NF-kB pathway and inhibits nuclear translocation and activity of the NF-kB transcription factors. STAT3 levels were similar in the high risk ABC and low risk non-DE GCB/U cases being significantly downregulated only in the DEXP_mRNA GCB/U subset. PIK3CA and PTEN levels did not vary significantly across different groups (Figure 2B).

Development of a three-gene prognostic signature combining MYC, BCL-2 and NFKBIA

In an effort to build a GEP signature aimed at refining current prognostication algorithms and suitable for clinical practice, we considered only those genes whose expression was significantly associated with OS and differentially represented in both high risk (ABC and DEXP_mRNA GCB/U) versus low risk (non-DEXP mRNA GCB/U) patient subsets. Using these criteria, we constructed a prognostic signature considering three genes (MYC, BCL-2 and NFKBIA), which combines the MYC/BCL-2 DEXP_mRNA status with NFKBIA expression (hereafter called MBN signature, see methods). Besides MYC and BCL-2, (defining the DEXP_mRNA status), NFKBIA emerged as the best survival predictor by gene ranking according to the predictive power (univariate z score) (Online Supplementary Figure S4). With this strategy, patients were divided in two risk categories characterized by different outcome: low risk patients (MBN-Sig low) had a very favorable prognosis (91% 5-year OS; 84% 5-year PFS), whereas high-risk patients (MBN-Sig high) had a significantly worse prognosis (64% 5-year OS; 59% 5-year PFS) (Figure 3A; Online Supplementary Figure S5A). Importantly the MBN signature retained its significance and outperformed the MYC/BCL-2 DEXP_mRNA status in multivariate analysis (Figure 3B; Online Supplementary Table S4). In fact, only the COO, the aaIPI score and the MBN signature were significantly associated with outcome in multivariate analyses (Figure 3B). These findings were confirmed in silico in a large independent validation cohort of 469 patients treated with R-CHOP (88% 5-year OS and 78% PFS for MBN-Sig low vs. 72% OS and 57% PFS for MBN-Sig high patients) (Figure 3C and D; Online Supplementary Figure S5B; Online Supplementary Table S5). The prognostic value of the MBN signature was further tested in a publicly available data set including 235 patients (from Lenz et al. 2008) treated with R-CHOP/R-CHOP-like regimens and in a real-life cohort (n=102 patients) with similar results (Online Supplementary Figure S6A and B). The MBN signature was able to identify a significant fraction of ABC-derived cases and about a third of GCB/U cases (Figure 3A and C; Online Supplementary Figure S6C and D).

**Table 1. Patients characteristics.**

<table>
<thead>
<tr>
<th>Trial name</th>
<th>RHD50305</th>
<th>DLCL04</th>
<th>P*</th>
<th>RHD50305 + DLCL04</th>
</tr>
</thead>
<tbody>
<tr>
<td>N° of patients</td>
<td>87</td>
<td>99</td>
<td>–</td>
<td>186</td>
</tr>
<tr>
<td>Immuno-CHT alone</td>
<td>49 (56%)</td>
<td>56 (57%)</td>
<td>–</td>
<td>105 (56%)</td>
</tr>
<tr>
<td>Immuno-CHT + ASCT</td>
<td>38 (44%)</td>
<td>43 (43%)</td>
<td>–</td>
<td>81 (44%)</td>
</tr>
<tr>
<td>Median age, y (range)</td>
<td>53 (21-85)</td>
<td>52 (18-85)</td>
<td>–</td>
<td>52 (18-85)</td>
</tr>
<tr>
<td>COO NanoString</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABC</td>
<td>15 (17%)</td>
<td>25 (25%)</td>
<td>ns</td>
<td>48 (21%)</td>
</tr>
<tr>
<td>GCB</td>
<td>53 (61%)</td>
<td>58 (59%)</td>
<td>ns</td>
<td>111 (60%)</td>
</tr>
<tr>
<td>Unclassified</td>
<td>19 (22%)</td>
<td>16 (16%)</td>
<td>–</td>
<td>35 (19%)</td>
</tr>
<tr>
<td>COO Hans IHC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-GCB</td>
<td>58 (67%)</td>
<td>60 (61%)</td>
<td>–</td>
<td>118 (63%)</td>
</tr>
<tr>
<td>GCB-like</td>
<td>29 (33%)</td>
<td>39 (39%)</td>
<td>ns</td>
<td>68 (37%)</td>
</tr>
<tr>
<td>Stage (Ann Arbor)</td>
<td>II-IV</td>
<td>III-IV</td>
<td>–</td>
<td>II-IV</td>
</tr>
<tr>
<td>aaIPI score</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low-Low Intermediate (0-1)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Intermediate-high (2)</td>
<td>55 (63%)</td>
<td>81 (82%)</td>
<td>ns</td>
<td>136 (73%)</td>
</tr>
<tr>
<td>High (3)</td>
<td>32 (37%)</td>
<td>18 (18%)</td>
<td>0.005</td>
<td>50 (27%)</td>
</tr>
</tbody>
</table>

*Two-sided Fisher’s exact test. N°: number; Immuno-CHT: immunochemothetry; ASCT: autologous stem cell transplantation; y: years; COO: cell of origin; IHC: immunohistochemistry; aaIPI: age adjusted international prognostic index; ns: not significant; GCB: germinal center B cells; ABC: activated B-cells; IHC: immunohistochemistry.
Figure 3. Legend on next page.
Real life applicability of the MBN signature

In order to provide a risk stratification tool applicable to routine clinical practice in a prospective manner, we constructed an RF model with the expression of genes characterizing the MBN signature. First, the classifier was trained on the discovery cohort splitting it into training (80%) and test (20%) dataset; in this case, the accuracy of the three-gene model was 95% in the training and 94% in test set. In order to confirm the reliability of this three-gene model, we further tested it in an independent dataset (validation set) consisting of the real-life cohort (n=102 cases). Of note, these cases were profiled with the same T-GEP panel and methods used in the discovery cohort, mitigating batch effects phenomena. As result, the three-gene model accurately classified 85% (87 of 102) cases as either MBN-Sig high or MBN-Sig low subgroups (Figure 4A). As reported in Figure 4B, the model effectively identified MBN-high and low categories with sensitivity (SE) and specificity (SP) of 94% and 76% respectively. Receiver operating characteristic (ROC) curve analysis revealed that the area under the curve (AUC) was 0.94 in the validation set (Figure 4C). Furthermore, this strategy produced a very efficient survival prediction, which as expected showed a worse outcome for the MBN-high subset (Figure 4D) and mirrored the OS curve based on the median MBN value depicted in the Online Supplementary Figure S6B.

Correlation of the MBN signature with fluorescence in situ hybridization status and clinical variables

Focusing the analyses on our discovery cohort of 186 patients (DCL04 and R-HDS0505 trials),24,25 we observed that the MBN signature significantly stratified the prognosis of GCB/U patients (Figure 5A). Since the MBN signature effectively stratified GCB/U DLBCL patients, we investigated correlations between the MBN-signature, FISH status and clinical variables in our discovery cohort. As shown in Figure 5B, we observed a significantly higher frequency of MYC and BCL-2 re-arrangements in the MBN-Sig high subgroup compared to the GCB/U MBN-Sig low subset. According to these observations, there was a significant enrichment of HG-BCL w/DH in the MBN-Sig high subgroup compared to the MBN-Sig low subset (Figure 5B; Online Supplementary Figure S7A). No differences in the number of cases with missing FISH analyses were observed between groups (data not shown). In line with the literature,23,25,27 all these cases, except one, were GCB-derived (data not shown). As previously shown in Figure 3, ABC-derived DLBCL were significantly more represented in the MBN-Sig high subgroup (Figure 5B; Online Supplementary Figure S7A). Finally, no significant differences in the aIPI score (intermediate high vs. high) were observed between groups (Figure 5B; Online Supplementary Figure S7A). These findings were validated in silico in the larger cohort from Sha et al.27 (Figure 5C and D). As observed in the discovery cohort, the MBN signature stratified the prognosis of GCB/U patients (Figure 5C) and identified the vast majority of DH cases. Again ABC-derived DLBCL were more highly represented in the MBN-Sig high subgroup (Figure 5D; Online Supplementary Figure S7B). In this study, the application of a gene expression classifier identified a molecular high grade (MHG) subgroup strongly enriched in DH lymphomas and comprising 9% of the total patient population.27 In order to evaluate how our MBN signature performed in the same patient population, we compared the MBN signature with the MHG signature and with the FISH status (Figure 5C). Notably the MBN-high subgroup was significantly enriched in MHG cases, identifying 76% of MHG DLBCL and the vast majority of DH (Figure 5D; Online Supplementary Figure S7B). Also in this cohort there were no differences in IPI score between groups (Figure 5D; Online Supplementary Figure S7B).

Rationale for a precision therapy approach in MBN-Sig high-risk diffuse large B-cell lymphoma patients

Since the MBN-Sig high subgroup is characterized by relatively higher MYC and BCL-2 expression and lower NFKBIA levels indicative of constitutive NF-κB activity, we next investigated the effect of differential therapeutic strategies in this high-risk patient subset. We first analyzed the impact of ASCT versus standard chemotherapy in the discovery cohort. ASCT consolid-
The aberrant activation of NF-κB observed in lymphoma is associated with decreased abundance of IkB-α (which is encoded by the NFKBIA gene). Since bortezomib is known to increase IkB-α levels by blocking its ubiquitination and therefore inhibiting NF-κB activity, we next examined the Sha dataset performing an exploratory ad hoc analysis to investigate the impact of the addition of bortezomib to standard R-CHOP (RB-CHOP) in the MBN-Sig high subset (characterized by decreased NFKBIA levels). Interestingly, RB-CHOP determined a significant PFS advantage in the MBN-Sig high population.
Figure 5. Legend on following page.
Figure 5. The MBN signature identifies prognostically distinct subgroups including activated B cells and a fraction of germinal center B cells/unclassified diffuse large B-cell lymphoma (DLBCL) enriched in double hit DLBCL cases, providing opportunities for precision therapies. (A) Overall survival (OS) of the germinal center B cells/unclassified germinal center B cells subset in the discovery cohort (n=146 patients) according to the integration of the targeted gene expression profiling (T-GEP) panel (LympH2Cx) with the MBN signature (MBN-Sig), distinguishing two risk categories according to the MBN-Sig. (B) Heatmap representing the three informative genes of the MBN-Sig shown as rows and double hit DLBCL tissue samples shown as columns, in the SHA’s cohort (n=469 patients treated with R-CHOP). (C) OS of the GCB/U subset in the SHA’s validation cohort (n=340 patients) according to the MBN-Sig, showing superimposable results compared to Figure 4C. (D) Heatmap representing the three informative genes of the MBN-Sig shown as rows and DLBCL tissue samples shown as columns, in the SHA’s cohort (n=469 patients treated with R-CHOP). (E) Progression-free survival (PFS) of patients treated with R-CHOP versus RB-CHOP in the MBN-Sig high subgroup (Full SHA’s cohort, n=928 patients; MBN-Sig high n=464 patients). (F) OS of patients treated with R-CHOP versus RB-CHOP in the MBN-Sig high subgroup (Full SHA’s cohort, n=928 patients; MBN-Sig high n=464 patients). In all panels the P-value was calculated with the log rank test. NA: not available; DH: double hit; DE: double expressor (based on DEEXP_mRNA status); aaIPI: age adjusted international prognostic index; COO: cell of origin; IPI: international prognostic index; MHG: molecular high grade. RB-CHOP: R-CHOP plus Bortezomib.

Discussion

In this study we applied a customized T-GEP panel (including the LympH2Cx signature for COO classification and additional genes of potential prognostic and therapeutic interest) to two randomized trials (n=186 patients) performed in the Rituximab era. The aim of this study was the integration of the COO with additional GEP-based variables, and the identification of a gene signature applicable to routine clinical practice, able to refine current prognostication algorithms. The genes of the T-GEP panel were selected considering the relevance of the respective signaling pathways in B-cell lymphomagenic, but more importantly based on their potential druggability.

Our study confirmed the prognostic value of GEP-based COO determination, which clearly outperformed the IHC-based Hans algorithm (the ABC DLBCL subgroups having a significantly inferior OS in all case series evaluated here) (Online Supplementary Figure S1). The COO retained its prognostic value in patients undergoing ASCT consolidation, suggesting that therapy intensification is not able to overcome the negative prognostic value of the COO. A recursive partitioning analysis integrating COO with MYC/BCL-2 DEEXP mRNA status identified three main subgroups (a low risk non-DEXP mRNA GCB/U subset and two high-risk groups including DEXP mRNA GCB/U and ABC-DLBCLs) (Figure 2A). The observation of lower NFKBIA levels in the ABC and DEXP mRNA GCB/U subgroups (overexpressing MYC and BCL-2 to a similar extent) (Figure 2B) suggests that, despite known biologic differences, these DLBCL subsets could share similar oncogenic dependencies on MYC, BCL-2 and the NF-kB pathway (being NFKBIA a negative regulator of NF-kB signaling). This observations prompted us to design a three-gene prognostic signature integrating MYC, BCL-2 and NFKBIA, which we called the MBN signature. The signature was first tested in our discovery cohort of 186 patients, identifying two subgroups characterized by different outcome (Figure 3), and was then applied to three independent datasets (469 patients treated with R-CHOP in the SHA cohort, 27 253 patients from the Lenz cohort, 36 and 102 patients treated in real-life clinical practice with R-CHOP/R-CHOP-like regimens) confirming its high prognostic significance (total number of tested cases 990). Since the discovery cohort had some unique characteristics (such as lack of low aa-IPI cases, a relatively low fraction of ABC cases and no uniform first-line treatment), the extensive validation performed in three additional cohorts treated with R-CHOP/R-CHOP-like regimens confirms that the key findings of the present study are indeed applicable to an unselected DLBCL population. Importantly, the MBN signature defined a high-risk group including a significant fraction of ABC cases (in line with data shown in the Online Supplementary Table S3 demonstrating a higher incidence of MYC/BCL-2 DEEXP mRNA and low NFKBIA expressers in the ABC subgroup), and about 80% of GCB/U cases (Figure 3). Therefore the MBN signature could potentially identify an increased proportion of patients at high risk of treatment failure, confirming standard risk stratifications (COO or DE status). The MBN signature was an independent prognostic predictor, outperforming the MYC/BCL-2-DEEXP mRNA status in multivariate analyses (Figure 3), thus confirming the added value of the third gene (NFKBIA) for prognostic stratification. The possible clinical applicability of the MBN signature was tested in the real-life cohort using an RF prediction model built on the discovery cohort, providing a reliable tool for prospective risk stratification (Figure 4). Importantly, the integration of the MBN signature with the COO allowed the identification of two risk categories within the GCB/U subset. These findings, which were validated in independent cohorts, could have immediate implications (Figure 5A and C; Online Supplementary Figure S6A to D). Two recently published studies confirmed the heterogeneity of the GCB subgroup and identified gene signatures allowing better risk stratification of this patient subset. 26,27 These signatures were able to identify a proportion of HG-BCL with DH/TH and a further group lacking MYC/BCL-2 re-arrangements but characterized by similar clinical features. However, the fact that these signatures are composed by several genes encompassing multiple pathways, could make their successful translation to clinical practice and precision therapy approaches quite challenging.

Our data are in line with these findings confirming that the GCB/U DLBCL subset represents indeed a rather heterogeneous disease category. The MBN signature could identify the majority of tumors with high-grade molecular features (HG-BCL with DH/TH) in the discovery cohort and SHA’s cohort (Figure 5B and D; Online Supplementary Figure S7A and B). Moreover, by applying the MBN signature to the SHA validation cohort we observed that the MBN-Sig high subgroup was significantly enriched in MGH DLBCL cases (Figure 5D; Online Supplementary Figure S7B). Taken together, these data indicate that a simple three-gene signature could efficiently identify high risk GCB/U DLBCL cases. Furthermore, the MBN-Signature is based on potentially druggable targets or pathways. For example, NFKBIA (encoding for IkB-α) could be targeted by proteasome inhibitors 43-45 and by bromodomain and extraterminal protein (BET) inhibitors, which are able to downregulate MYC while increasing IkB-α levels. 36-38 Our
analysis on the impact of bortezomib in the MBN-high subgroup of the Sha cohort 27 from the REMoDL-B trial seems to confirm a potential druggability of the MBN signature: in fact treatment with RB-CHOP (R-CHOP plus bortezomib) was associated with a significantly prolonged PFS which translated in increased OS rates in the MBN-high subgroup, as compared to standard R-CHOP (Figure 5E and F). Proteasome inhibitors, BTK inhibitors and selective BCL-2 inhibitors could be the basis forrationally-designed combinations for the MBN-Sig high DLBCL subgroup. Alternative strategies to target NF-kB include lenalidomide and B-cell receptor signaling inhibitors, all of which are under clinical investigation in DLBCL. Three COO-based phase III trials testing R-CHOP + Bortezomib (Phoenix trial 19) or Lenalidomide (ROBUST trial 18) or bortezomib (REMoDL-B trial 19) did not meet their primary endpoints. Although several factors concurred to these negative results, the development of alternative and druggable molecular signatures represents an unmet need and could be of primary importance for the design of future precision medicine clinical trials.

The results of our study indicate that a simple and cost-effective three-gene assay (MBN signature) could refine current prognostic stratification algorithms providing the rationale for the implementation of precision medicine trials in the MBN-Sig high subset.

Disclosures

ACE and NC are employees and shareholders of NanoString technology; ED has received research funding from TC-Therapeutics, ADC-Therapeutics, Takeda and sits on the Advisory Board for Gilead; ES has received support from Novartis and EusaPharma for educational events; ACab sits on the Advisory Boards with Celgene, Gilead-Kite, Janssen, Igene, Takeda and has received honoraria for lectures from Celgene, Gilead-Kite, Janssen, Roche, Servier; UV has a consulting or advisory role for Celgene, Gilead and Janssen and is part of the speakers’ bureau with Roche, Celgene, Janssen, Gilead Sciences, Abbvie, Sandoz; AR sits on the National or International Advisory Boards for Gilead, Amgen, Novartis, Pfizer, Celgene, Italfarmaco, Sanofi-Aventis, Astellas, Roche, Omeros and has sponsored symposia for Amgen, Novartis, Celgene, Roche; PC has received honoraria for Advisory Board participation or as a lecturer from AbbVie, Amgen, Celgene, Daichii Sankyo, Gilead, Incyte, Janssen, Kite, KiovaKirin, Novartis, Roche, Sanofi, Servier, Takeda; PLZ has received honoraria for speakers’ bureau or Advisory Boards for Verastem, CellTrion, Gilead, Janssen-Cilag, BMS, Servier, Sandaz, MSD, Immune Design, Celgene, Portola, Roche, EusaPharma, Kyowa Kirin, Sanofi; CT sits on the Advisory Board for ADC-Therapeutics; SP sits on the Advisory Boards for Celgene, NanoString, Roche; SM, FM, GM, MF, RB, CA, CC, SR, ACA, SF, VT, ACab, GP, AMG have no conflicts of interest to disclose.

Contributions

ED and SP designed the study, interpreted the data and wrote the manuscript; SM and MF performed bioinformatics and statistical analyses; and SH helped with manuscript writing; FM and CM performed T-GEP experiments; VT, SF, CA and ACA performed immunohistochemistry; SP, ES, VT, SF and CA evaluated immunohistochemistry data; CC and SR performed FISH analyses; ACe and NC helped designing T-GEP experiments and helped with data interpretation; RB helped with data collection; ACab, GP, CT, AMG, PLZ, AR, PC, UV and ACab helped with data collection and interpretation. All authors critically reviewed the draft and approved the manuscript.

Acknowledgements

The authors wish to thank Pier Luigi Antoniotti, Sebastiano Spagnolo, Marco Giaffreda and Virginia Maltoni for technical assistance.

Funding

This study was funded by the AIRC 5x1000 grant to SP (n. 21498) and Italian Ministry of Health with Ricerca Corrente.

References


