Online supplementary Methods

Case Selection

Two hundred and nineteen patients (125M/94F; median age, 61 years) consecutively diagnosed with *de novo* DLBCL between 2002 and 2007 were retrieved from the files of 5 institutions of the *Grup per l'Estudi dels Limfomes de Catalunya i Balears* (GELCAB) with the only criterion of the availability of adequate histological material. All the tumors were classified as DLBCL according to the current WHO classification. Patients with previous indolent lymphoma, immunodeficiency-associated lymphomas, post-transplant lymphoproliferative disorders, intravascular, central nervous system, primary effusion lymphomas and primary mediastinal lymphomas were excluded from the study. Patients with BL were also excluded. Formalin-fixed and paraffin-embedded tissue was available in all cases.

Morphology and immunohistochemistry

The diagnosis of DLBCL was based on the criteria established in the World Health Organization (WHO) classification, and the tumors were classified into the more common morphological variants. Immunohistochemical (IHC) studies were performed with a panel of monoclonal and polyclonal antibodies reactive in paraffin-embedded sections using a peroxidase-labeled detection system, standard antigen retrieval protocols, and automated immunostainers (Dako autostainer, Denmark; Bond-Max, Leica Microsystems, Germany). Tissue microarrays (TMAs) were constructed using a tissue arrayer (MTA I; Beecher Instruments) and included two 1-mm representative cores of each case. Standard methods for tissue fixation (10% buffered formalin) and processing were used in the participating centers. The panel of antibodies included common B and T-cell markers as well as CD10 (clone 56C6), MUM1/IRF4 (clone MUM1p), BCL2 (clone 124), and Ki-67 (clone MIB-1) all from Dako, and BCL6, kindly

provided by Dr. Roncador (Centro Nacional de Investigaciones Oncologicas, Madrid, Spain).

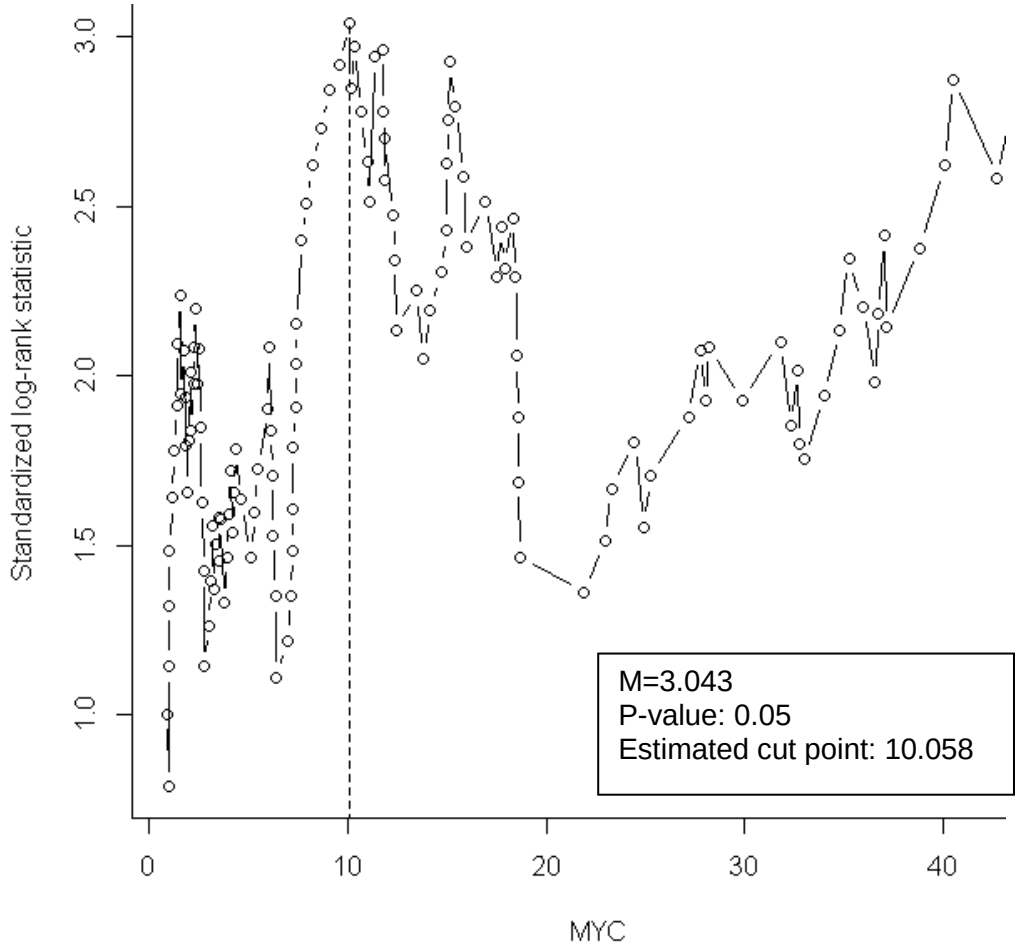
MYC antibody (clone Y69, Epitomics, USA) was used at dilution 1/40, with incubation of the primary antibody for 1 hour and antigen retrieval at pH 6 for 30 minutes. A MYC TMA that included tonsil sections, Burkitt lymphomas and Burkitt cell lines Raji and Namalwa was used as external control. MYC immunostaining was evaluated using a computerized image analysis method. Digitalized images were acquired with the ScanScope CS System at 20x magnification and then quantified using a nuclear algorithm with the Aperio ImageScope software version 9.0.0.1521 (Aperio Technologies, Vista, CA) in TMA sections. Cases with small biopsy samples and cores that dropped off of the TMAs were studied as whole-tissue sections, and a mean number of 7000 cells were evaluated per case (range 850-50000 cells).

We also evaluated MYC immunostaining in a semi-quantitative manner, and the samples were stratified into 5 groups based on the estimated percentage of positive cells: 1 ($\leq$10% positive tumor cells), 2 (11-25%), 3 (26-50%), 4 (51-75%), 5 (>75%). With this approach the distribution of MYC expression was: group 1, 43 (25%); 2, 42 (25%); 3, 37 (22%); 4, 25 (15%), and 5, 23 (13%).

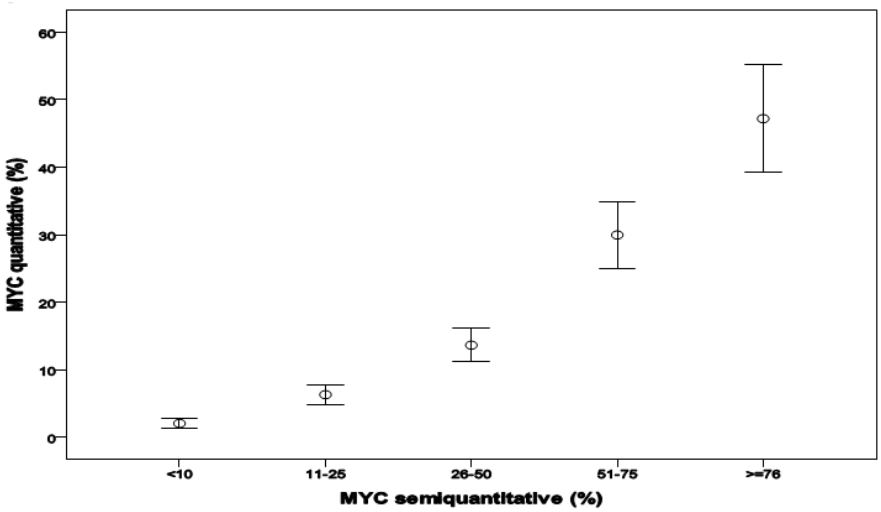
Fluorescence “In Situ” Hybridization (FISH)

FISH was done on 3 to 4 µm thick sections of TMA, using split signal DNA probes from Dako specific for the following loci: 18q21 (*BCL2*), 18q21 (*MALT1*), 3q27 (*BCL6*) and 8q24 (*MYC*). Tonsil sections were used as controls. For each tumor and tonsil sample a minimum of 100 evaluable nuclei were scored. The cut-off value used to detect rearrangements was 3%. Moreover, the mean number of numerical and structural genetic alterations was evaluated in each case, in order to assess the incidence of the genetic events in the tumors.

Supplementary Figure 1. Optimal cut-off of the quantitative MYC assessment for predicting survival (R statistical package, v. 2.8.1, Vienna, Austria; ref. 28)



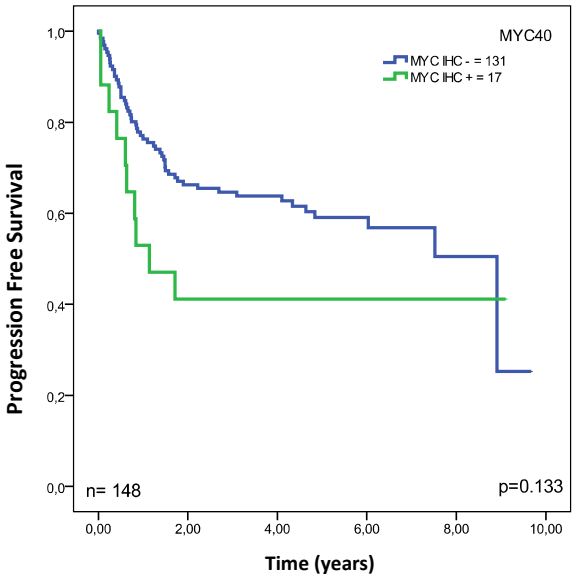
Supplementary Figure 2. Correlation between computerized and semiquantitative methods evaluating of MYC protein expression



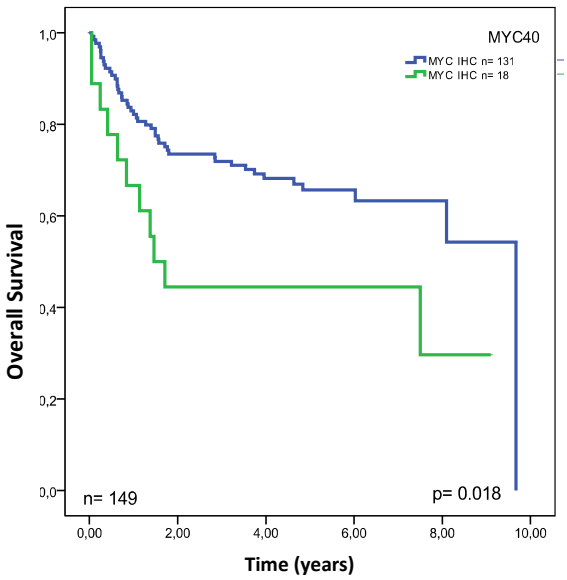
p<0.001

Supplementary Figure 3. Kaplan-Meier analysis of *de novo* DLBCL patients treated with immunochemotherapy with curative intention. A) Progression-free survival (PFS) and B) Overall survival (OS) according to MYC expression using a computerized method and a threshold of 40%; C) Progression-free survival (PFS) and D) Overall survival (OS) according to MYC expression using a manual method with a threshold of 25%

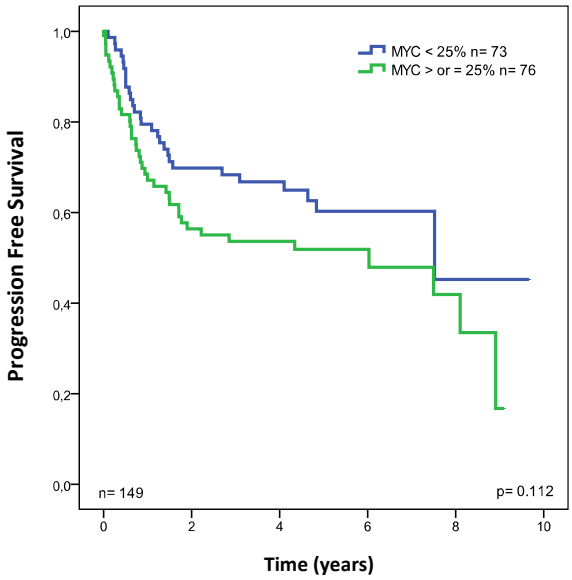
S3A



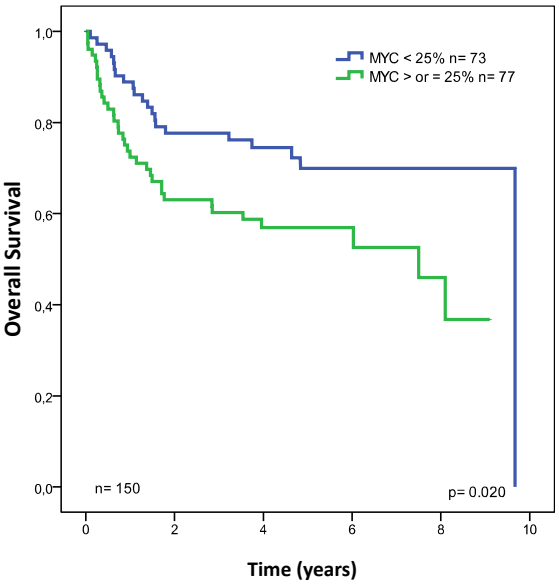
S3B



S3C



S3D



Supplementary Table 1: MYC gene and protein expression in cases with genetic alterations.

Cases	MYC gene	MYC protein
Rearranged		
1	15%	19%
2	79%	30%
3	80%	47%
4	80%	49%
5	90%	51%
6	90%	82%
7	93%	75%
8	> 95%	4%
9	>95%	30%
10	>95%	33%
11	>95%	80%
12	>95%	51%
Amplified		
1	91%	12%
2	97%	Not determined
3	>95%	15%

Supplementary Table 2: Main clinico-pathologic data according to the MYC expression using a cut-off of 40%

	MYC negative (n=144)	MYC positive (n=21)
Immunohistochemistry		
CD10+*	36/140 (26%)	11/20 (55%)
BCL6+	86/138 (62%)	15/18 (83%)
MUM1+	70/132 (53%)	7/19 (37%)
BCL2+	65/120 (54%)	12/16 (75%)
FISH		
BCL2 alterations	48/131 (37%)	8/18 (44%)
BCL6 alterations	55/127 (43%)	7/16 (44%)
MALT1 alterations	28/126 (22%)	2/17 (12%)
Clinical data		
Age ≥ 60	83/145 (57%)	16/21 (76%)
Gender (M:F)	80:67	13:8
Extranodal*	40/123 (33%)	11/18 (61%)
Stage III-IV	87/142 (61%)	16/21 (76%)
Elevated LDH*	68/129 (53%)	17/21 (81%)
IPI-III/IV high*	63/137 (46%)	16/19 (84%)
Complete response	94/140 (67%)	13/19 (68%)
Median OS (years)*	9.67	1.46
Median PFS (years)	8.41	1.14

* $P<0.050$

Supplementary Table 3: Summary of multivariate analyses according to the MYC expression using a cut-off of 10% or 40% obtained by a computerized method, and 25% obtained using a semi-quantitative approach.

	MYC 10% N=141		**MYC 40% N=141**		**MYC 25% N=141**	
	***P* value**	**Relative risk**	***P* value**	**Relative risk**	***P* value**	**Relative risk**
Progression-free survival						
MYC gene status	NS	-	NS	-	NS	-
MYC expression	NS	-	NS	-	NS	-
IPI	<0.001	1.5	0.002	2.0	<0.001	3.9
Overall survival						
MYC gene status	NS	-	NS	-	NS	-
MYC expression	0.023	1.95	NS	-	NS	-
IPI	0.001	1.5	<0.001	2.2	<0.001	4.7