Gene expression analysis provides a potential rationale for revising the histological grading of follicular lymphomas
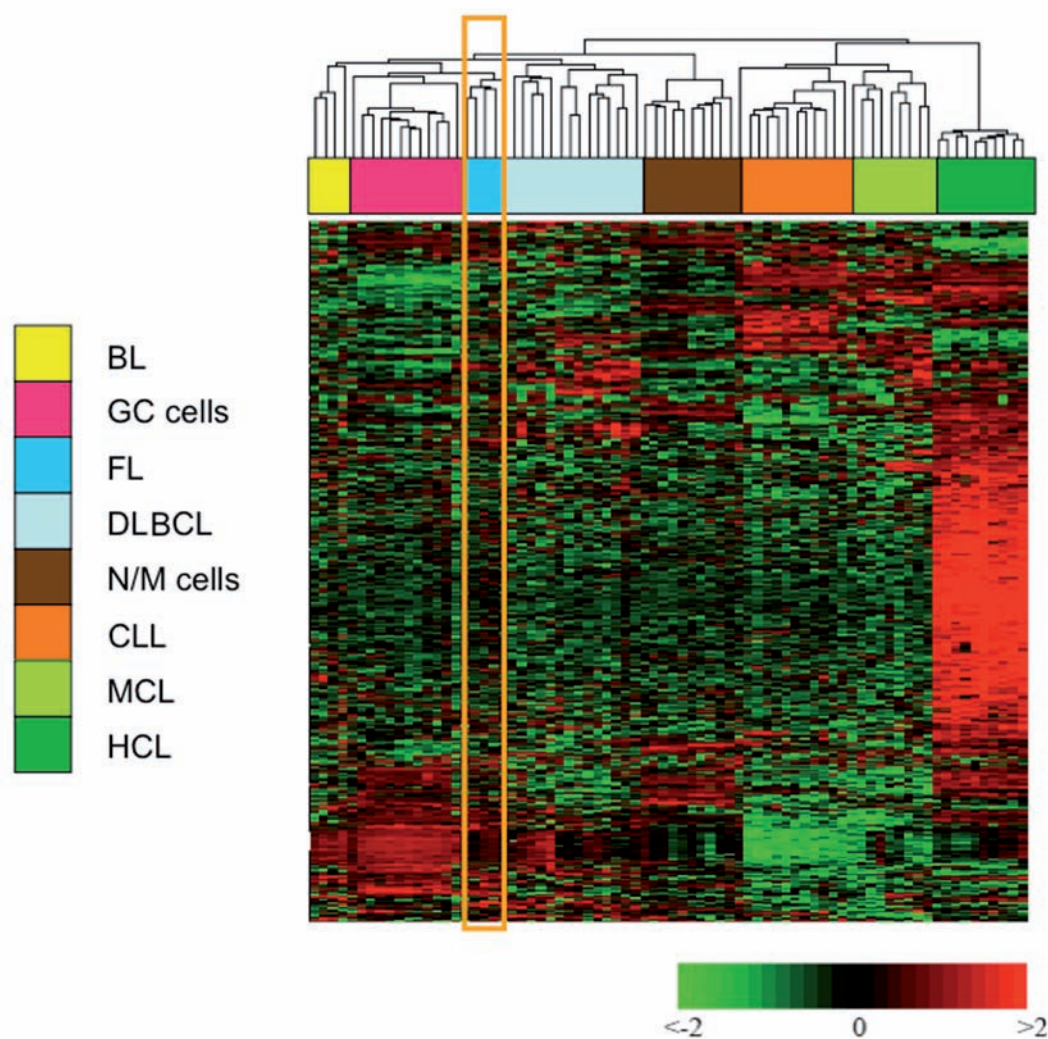
Pier Paolo Piccaluga,[1,2] Andrea Califano,[3] Ulf Klein,[2] Claudio Agostinelli,[2] Beatriz Bellosillo,[4,5] Eva Gimeno,[6] Sergi Serrano,[4,5] Francisco Solè,[4,5] Yonghui Zang,[7] Brunangelo Falini,[5] Pier Luigi Zinzani,[1] and Stefano A. Pileri[1]

[1]Department of Hematology and Oncology "L. and A. Seràgnoli", Hematopathology and Hematology Units, S. Orsola-Malpighi Hospital, University of Bologna, Bologna, Italy; [2]Institute for Cancer Genetics, [3]Center for Computational Biology and Biochemistry, [7]The Herbert Irving Comprehensive Cancer Center, Columbia University, NY, USA; [4]Unitat de Recerca en Neoplàsies Hematològiques-Parc Recerca Biomèdica Barcelona (URNHE-PRBB), Barcelona, Spain; [5]Laboratori de Citogenètica i Biologia Molecular, Departament de Patologia, [6]Clinical Hematology Department, Hospital del Mar, IMAS, Barcelona, Spain and [6]Hematopathology, Insitute of Hematology, Monteluce Hospital, Perugia University, Perugia, Italy
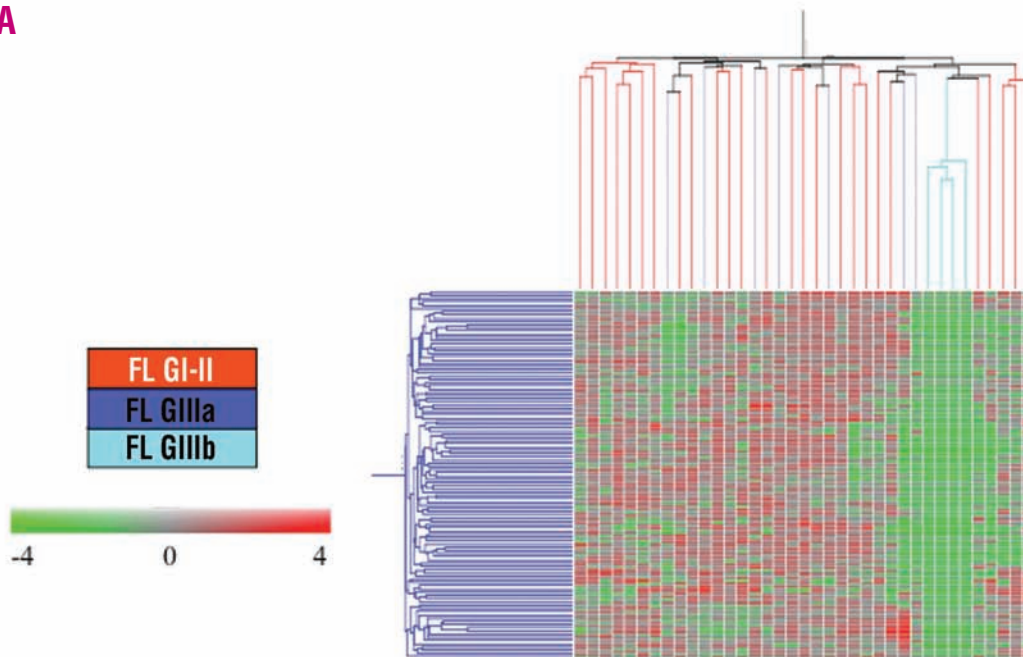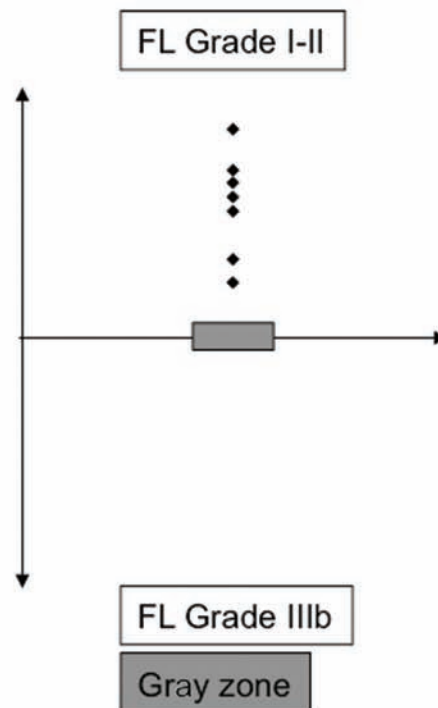
Supplementary Figure S1. Unsupervised hierarchical clustering of gene expression profiles generated from FL, normal B lymphocytes and B-cell derived malignances. The unsupervised analysis was performed on six enriched FL, normal B cells (including naïve cells -N, memory cells – M, centrocytes – CC, and centroblasts – CB) and a representative panel of B-cell tumors that included four cases of Burkitt's lymphoma (BL), 16 cases of diffuse large B cell lymphoma (DLBCL), ten cases of mantle cell lymphoma (MCL), ten cases of hairy cell leukemia (HCL) and ten cases of B-cell chronic lymphocytic leukemia (B-CLL). The dendrograms were generated using a hierarchical clustering algorithm based on the average-linkage method. In the matrix, each column represents a sample and each row represents a gene. The color scale bar shows the relative gene expression changes normalized by the standard deviation (0 is the mean expression level of a given gene). The samples (16 HCL, 6 FL, 4 BL, 16 DLBCL, 10 MCL, and 10 B-CLL) are clustered according to their expression of 332 genes. FL is clearly distinct from other malignancies, being closer to the germinal center- related ones (BL and DLBCL).
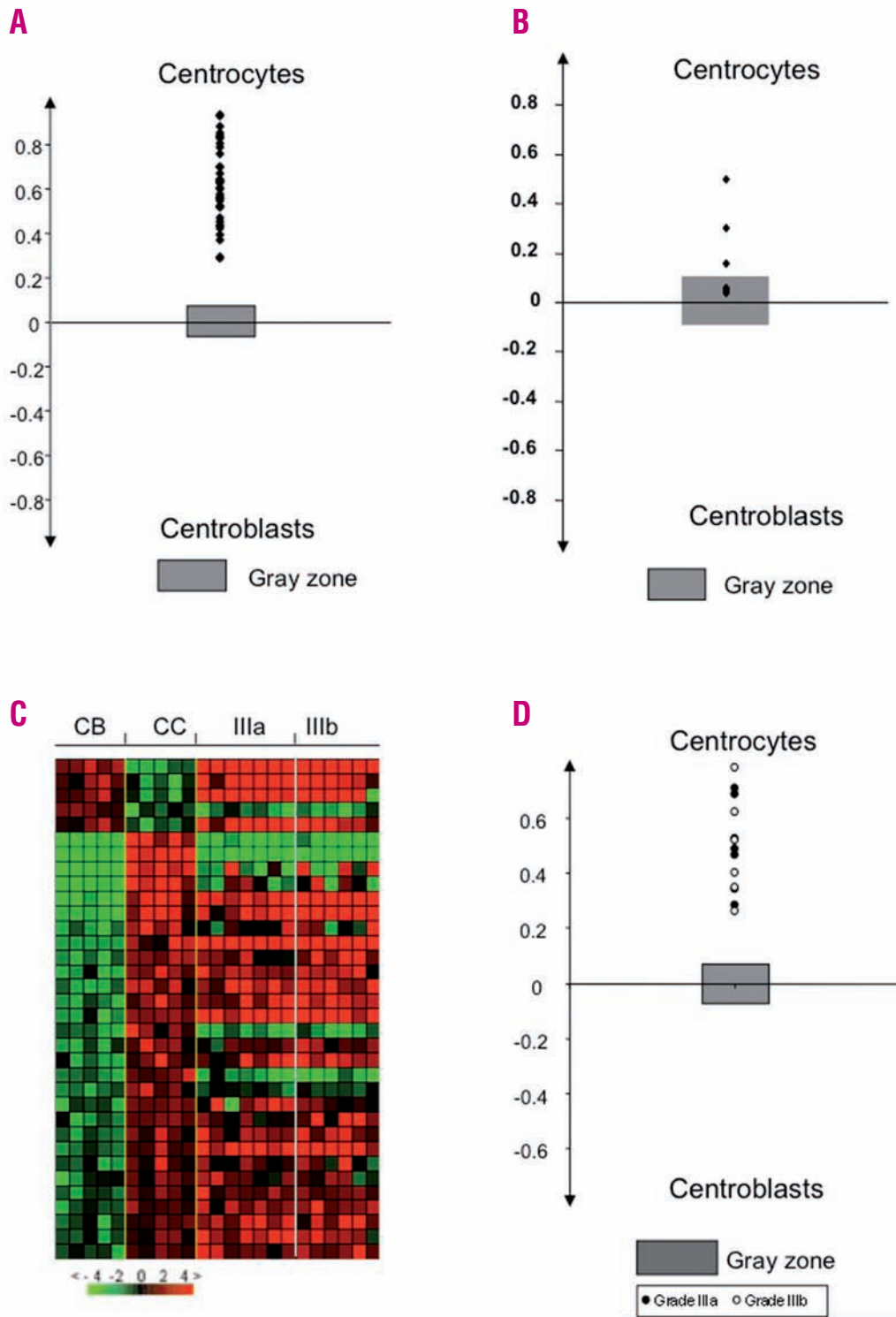
Supplementary Figure S2. Relatedness of the gene expression profile of FL to normal germinal center B cells. Analysis in FL of genes differentially expressed in centrocytes (CC) and centroblasts (CB). These genes were identified by supervised analysis. A cell-type classification is used to measure the relatedness of FL to CC and CB. The gray area marks 95% of confidence: the *p*-value decreases with increasing distance from the x axis. (A) The expression of genes differentially expressed in CC and CB was investigated in purified FL cases represented on the right side of the matrix. (B) FL seems to be more related to CC cells than to CB, when the cell-type classification is adopted. (C) The expression of genes differentially expressed in CC and CB was then specifically investigated in unpurified cases of FL GIIIa and GIIIb (represented on the right side of the matrix). The cell-type classification showed again that FL seems to be more related to CC cells than to CB, irrespectively of the amount of large cells.
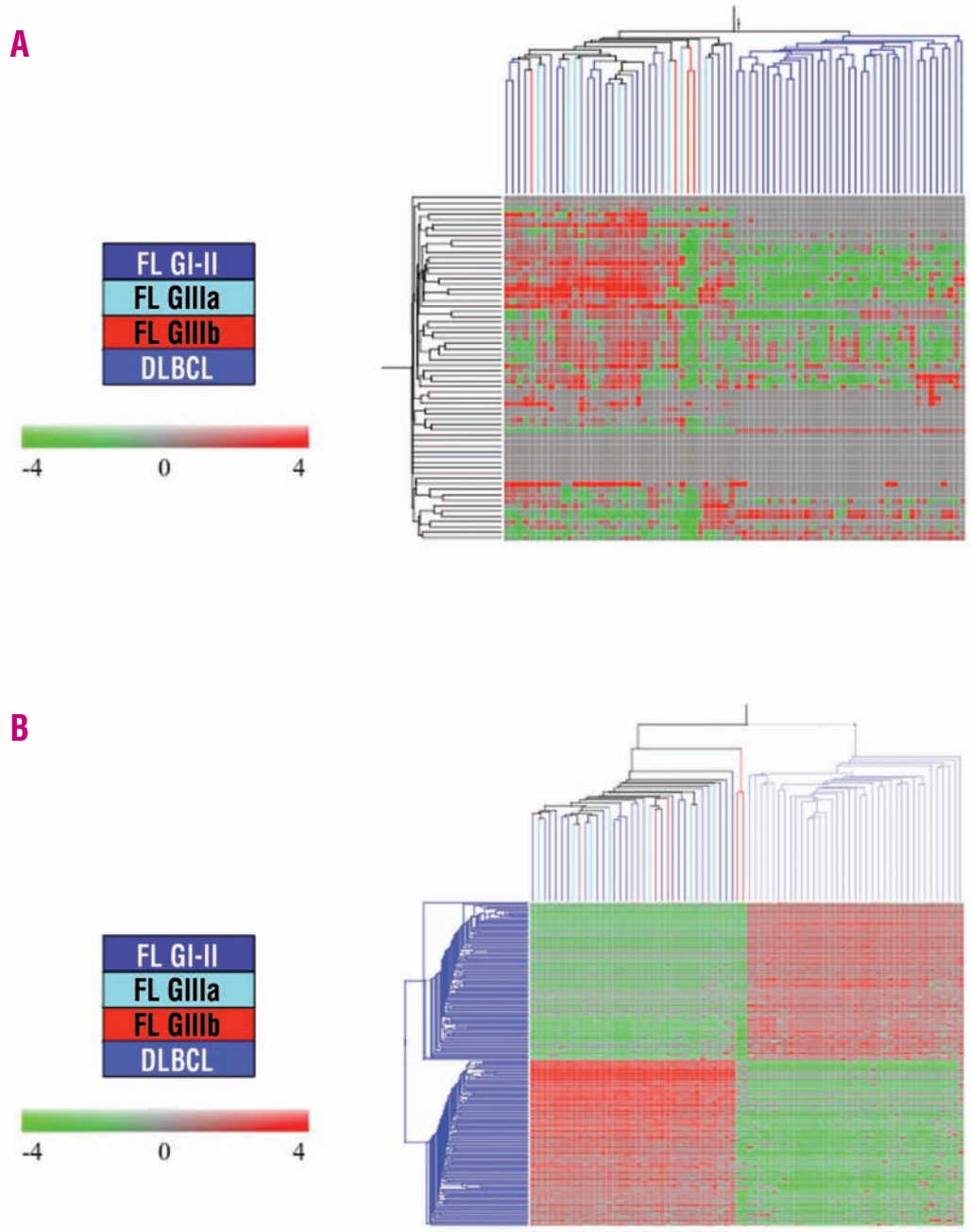
Supplementary Figure S3. FL GIIIb is in fact closer to FL than to DLBCL based on gene expression profile. (A) Hierarchical clustering of 37 FL and 37 GCB-DLBCL cases. The samples are clustered according to the expression of 64 genes that emerged from supervised analysis comparing enriched samples of FL vs DLBCL. FL GIIIb actually cluster together with the other FL rather than with DLBCL. (B) Hierarchical clustering of 37 FL and 37 GCB-DLBCL cases. The samples are clustered according to the expression of 268 genes that emerged from supervised analysis comparing FL GIIIa vs GCB-DLBCL. FL GIIIb actually cluster together with the other FL rather than with DLBCL.

Supplementary Figure S4. FL GIIIa is closer to FL GI-II than to FL GIIIb based on gene expression profile. (A) Hierarchical clustering of 37 FL cases. The samples are clustered according to the expression of 114 genes that emerged from supervised analysis comparing samples of FL GI-II vs FL GIIIb. FL GIIIa actually cluster together with the former group. (B) A cell-type classification was used to measure the relatedness of FL GIIIa to FL GI-II and GIIIb. The gray area marks 95% of confidence: the *p*-value decreases with increasing distance from the x axis. All GIIIa cases appear to be significantly more related to the indolent (GI-II) forms.