

Microarrays in Diffuse Large B-Cell Lymphoma

Chen Lossos
Biochemistry 118Q
Professor Brutlag
December 3, 2009

Diffuse large B-cell lymphoma (DLBCL) is an aggressive malignancy of mature B lymphocytes, with more than 25,000 new occurrences in the United States every year and accounting for nearly 40% of Non-Hodgkins Lymphoma cases. Typical presentation of DLBCL includes rapidly enlarging lymph nodes or tumors in extra-nodal sites. Night sweats, fever and weight loss are observed in approximately 30% of patients. Patients with DLBCL have highly variable clinical courses: although most patients initially respond to chemotherapy, fewer than half of the patients achieve a durable remission (Fauci). Clinical prognostic models, such as the International Prognostic Index (IPI) have been developed to identify DLBCL patients who are unlikely to be cured with standard therapy. The clinical factors which constitutes IPI (age, performance status, stage of disease, number of extranodal sites and serum lactate dehydrogenase levels), however, frequently fail to accurately predict patient survival. This, according to Schipp et al. suggests that there exist underlying biological mechanisms which are responsible for the observed differences of survival of patients with similar IPI scores (Shipp).

Like many cancers, the biological mechanisms underlying DLBCL pathogenesis are markedly intricate, involving relationships between numerous genes, signaling pathways, and regulatory processes. Therefore, to study a single gene to identify disease cause and pathogenesis does not accurately reflect gene expression changes. Instead, a molecular technique capable of evaluating multiple components of the biological processes of tumorigenesis and pathogenesis is necessary to fully understand these processes in DLBCL. The advent of DNA microarray technology has done just this. Microarrays have not only allowed for the development of better diagnostic and treatment techniques of DLBCL, but have also illuminated the pathogenesis of the disease. Most importantly, microarrays have allowed for the subclassification of DLBCL into

two distinct categories with unique clinical outcomes and survival: Germinal Center Diffuse Large B-Cell Lymphoma (GC-DLBCL) and Activated peripheral Blood B-Cell Diffuse Large B-Cell Lymphoma (ABC-DLBCL).

The first microarray study of DLBCL, carried out at Stanford by Alizadeh et al. provided the strongest initial evidence for the subcategorization of DLBCL. This study utilized a specialized cDNA array called Lymphochip, which was constructed by selecting genes expressed in lymphoid cells or which were previously reported to be involved in cancer biology. Using an unsupervised hierarchical clustering method (i.e. microarray data are analyzed without using external information), 42 tumors from patients treated with anthracycline based chemotherapy revealed two distinct subgroups of DLBCL based on gene expression. The expression patterns were characteristic of normal germinal center (GC) B-cells and activated peripheral blood B-cells. In addition to these categories, expression signatures (genes of similar function which cluster together) were observed in genes associated with proliferation and lymph nodes. The strongest evidence for the categorization of DLBCL into the two subtypes, however, came from patients' survival rate. Although the average five-year survival for all patients was 52%, 76% of GC B-like DLBCL patients alive after five years, as compared with only 16% of activated B-like DLBCL patients. Thus, this difference in patient survival based on the GC and ABC classification system provides a possible explanation for the observed differences in survival between patients with similar IPI scores. Thus, the molecular differences between the two patient groups were accompanied by a remarkable divergence in clinical behavior, suggesting that GC B-like DLBCL and activated B cell DLBCL should be regarded as distinct diseases.

Although this data indicated for the need to subclassify DLBCL into GC-DLBCL and ABC-DLBCL, there remained a need to validate this classification in independent sets of patients

or to confirm the findings by other biological parameters. The support needed for these DLBCL subgroups was provided by a study performed by the Lymphoma/Leukemia Molecular Profiling Project Group, which analyzed tumors from 240 DLBCL patients treated with anthracycline-based chemotherapy. Using a similar methodology to Alizadeh et al. (a cDNA Lymphochip and unsupervised clustering), this study confirmed the clustering of patients into GC and ABC-like DLBCL categories. Moreover, the clinical impact of this molecular classification was confirmed, as a much better prognosis was seen in patients with GC-like DLBCL (5-year survival of 60%) compared with ABC (35%) (Rosenwald et al).

The subdivision into GC-DLBCL and ABC-DLBCL were further confirmed by biological presentations of patients. The mutation of the variable heavy (V_H) chain of the Ig gene, frequently rearranged and over-expressed in B cell lymphomas, provided the opportunity to do this. 14 DLBCL patient tumor samples were clustered by hierarchical clustering using the Lymphochip. 7 of the patients occupied a common branch and showed a gene expression profile similar to the GC-DLBCL seen in Alizadeh et al. All 7 of these patients displayed the mutation of the Ig gene. In contrast, none of the 7 patients who exhibited ABC-DLBCL characteristic expression displayed ongoing mutations in the Ig gene. This suggests that GC-DLBCL and ABC-DLBCL originate from different ontogeny lymphocyte precursors (Lossos et al PNAS).

In addition to the mutation of the Ig gene, other findings substantiated the need to subdivide DLBCL into two entities. The t(14;18)(q32;q21) translocation, which involves the BCL-2 gene and amplification of the c-rel locus on chromosome 2p, was detected almost exclusively in Germinal Center-like DLBCL. Furthermore, Bea et al. found that trisomy 3, gains of chromosomal regions 3q and 18q21-q22, and losses of 6q21-22 were associated with ABC-like DLBCL. Further, GC-like DLBCL had frequent gains of 12q12. A parallel analysis yielded

that these DNA amplifications were strongly correlated with impact on genes whose expression is involved chromosomal regions in a subgroup-specific fashion (Huang et al). These findings strengthened the need to subdivide DLBCL into GC-like and ABC-like subtypes.

Despite the advances that unsupervised clustering analysis brought to understanding DLBCL, this method has its limitations. Namely, expression profiling by DNA microarrays using unsupervised clustering is unable to distinguish the impact single genes have in lymphoma tumorigenesis and pathogenesis. The Lymphoma/Leukemia Molecular Profiling Project Group addressed this by examining the correlation between the expression of individual genes and patient outcomes. A model of patients survival, on the basis of the expression of 17 genes (BCL-6, clone 1334260, clone 814622, HLA-DPa, HLA-DQa, HLA-DRa, HLADrb, a-actinin, collagen type III a 1, connective-tissue growth factor, fibronectin 1, KIAA0233, urokinase plasminogen activator, c-myc, E21G3, NPM3, and BMP6) was constructed using supervised microarray clustering (Rosenwald et al). The score from this model, based on the expression of the aforementioned genes, correlated significantly with clinical outcome in DLBCL and contributed significantly to IPI's predictive power. Thus, microarrays have the potential to not only analyze the entire genome, but focus on specific genes implicated in disease.

In an additional study, another group used 6817 genes to describe gene expression in 58 DLBCL tumors. Two groups of patients (those with cured versus refractory disease) were used to create supervised clustering in an effort to develop an outcome predictor model. A set of 13 genes (dystrophin related protein 2, protein kinase C gamma, MINOR, 5-hydroxytryptamine 2B receptor, H731, transducin-like enhancer protein 1, PDE4B, protein kinase C-beta-1, oviductal glycoprotein, zink-finger protein C2H2-150, and 3 expressed sequence tags) were found to accurately predict DLBCL outcome. GC signature defining genes, determined in the

Lymphochip studies, again segregated DLBCL patients into the GC- and ABC-like DLBCL (Shipp et al). However, contrary to other studies, the overall survival of these two groups was not different. Also interesting to note was this study had no genes in common with the model proposed by Lymphoma/Leukemia Molecular Profiling Project Group, suggesting that the parameter to which gene expression data is correlated to influences which genes can be used to predict clinical outcome. Thus, gene expression data have linked specific genes to clinical presentations, illuminating the pathogenesis of disease and implicating specific genes in particular cellular processes.

In an attempt to reconcile the differences achieved by the aforementioned studies and to devise a technically simple method that could be applicable for routine clinical use, Lossos et al evaluated the mRNA expression of genes previously reported to predict survival in DLBCL from 66 patients treated with anthracycline based therapy. The top 6 genes, ranked according to their predictive power, were used to construct a model based on their relative individual contribution into a multivariate analysis (mortality-predictor score= $(-0.0273 \times LMO2) + (-0.2103 \times BCL6) + (-0.1878 \times FNI) + (0.0346 \times CCND2) + (0.1888 \times SCYA3) + (0.5527 \times BCL2)$, where negative coefficients are associated with longer predicted survival). Among the selected genes, LMO2, BCL-6, and FN1 predicted longer survival, whereas CCND2, SCYA3, and BCL-2 predicted shorter survival. Based on the expression of these 6 genes, patients could be subdivided into IPI-independent low, intermediate, and high risk groups with different 5-year survival rates, ranging from 65% in the low-risk to 15% in the high risk subgroups (low risk score less than 0.063; medium risk, from 0.063 to <0.093; and high risk, 0.093 or higher). This model was subsequently validated in the data sets available from previously reported studies, including the two previously discussed

studies, and was shown to be a statistically significant ($P=0.004$) predictor of patient survival (Lossos et al NEJM).

It can be seen that microarrays have been extremely beneficial in studying the genetics of Diffuse Large B-Cell Lymphoma, leading to a greater understanding of the disease and understanding underlying molecular signatures with characteristic expression that suggest disease pathogenesis and genes implicated with the malignancy. The full potential of microarrays, however, has yet to be fully realized in DLBCL. In fact, microarrays have the potential to “(1) identify previously unrecognized disease entities with distinct biologic and clinical features, (2) elucidate the key genetic profiles and lesions that define each of these new nosologic entities, (3) discover new molecular targets for future therapeutic intervention, (4) identify genes that play a potential role in determining prognosis, (5) discover previously unknown genes of major clinical relevance from numerous expressed sequence tags (EST) clones present on the arrays, and (6) identify gene expression signatures correlated with response to specific therapeutic agents” (Lossos et al). This therefore could lead to patient-tailored chemotherapy regimens which could improve patient outcome and improve treatment efficiency. Therefore, there remains to be much research done with DNA microarrays which can lend a better understanding of disease progression and ultimately treatment targets for DLBCL.

Although not many modifications have been made as a result of microarray data on clinical treatment of DLBCL, some advances in patient treatment have been reported. Based on work by Coiffier et al, which showed that addition of CHOP-retuximab to chemotherapy of regimens has led to increased survival (5 years) in 10-15% of patients, Lenz et al analyzed 181 patients treated without R-CHOP and 233 patients treated with the addition of R-CHOP to their chemotherapeutic regimens to determine what molecular signatures corresponded to improved

patient response to chemotherapy. This study found that while patient gene expression clustered into four signatures (GC-B Cells, lymph node, proliferation and MHC associated genes), an optimal survival model for R-CHOP combined the germinal center B-cell and lymph nodes signatures. Thus, based on this model it is possible to predict survival in patients treated with R-CHOP.

The potential for microarrays to improve our current understanding of cancers and the changes in gene expression the disease induces is great. Although there have been many gene expression studies which have helped illuminate many aspects of many diseases, Diffuse Large B-Cell Lymphoma among them, many more studies are needed to fully understand disease pathogenesis and genes implicated in disease. While this may necessitate large sources of funding and man-power (expressed in the form of many hours spent in research laboratories), the benefits of this information could lead to improved and personalized treatments based on the individual gene expression data of patients and must be performed for the benefit of all patients.

Works Cited

- Alizadeh AA, Eisen MB, Davis RE, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 2000;403:503-11.
- Bea S, Zettl A, Wright G, Salaverria I, Jehn P, Moreno V, et al. Diffuse large B-cell lymphoma subgroups have distinct genetic profiles that influence tumor biology and improve gene-expression-based survival prediction. *Blood* 2005;106: 3183 – 3190.
- Fauci, Anthony S., Eugene Braunwald, Dennis L. Kasper, Stephen L. Hauser, Dan L. Longo, J. Larry Jameson, and Joseph Loscalzo, Eds. *Harrison's Principles of Internal Medicine*. 18th Edition. NY: McGraw Hill, 2008.
- Huang JZ, Sanger WG, Greiner TC, Staudt LM, Wiesenburger DD, Pickering DL, et al. The t(14;18) defines a unique subset of diffuse large B-cell lymphoma with a germinal center B-cell gene expression profile. *Blood* 2002;99: 2285 – 2290.
- The International Non-Hodgkin's Lymphoma Prognostic Factors Project. A predictive model for aggressive non-Hodgkin's lymphoma. *N Engl J Med* 1993;329:987-94.
- Lenz G, Wright G, Dave S, Xiao W, Powell J, Zhao H, et al. Stromal Gene Signatures in Large Cell B-Cell Lymphomas. *N Engl J Med*. 2008 Nov 27;359(22):2313-23.
- Lossos IS, Alizadeh AA, Eisen MB, Chan WC, Brown PO, Botstein D, et al. Ongoing immunoglobulin somatic mutation in germinal center B cell-like but not in activated B cell-like diffuse large cell lymphomas. *Proc Natl Acad Sci USA* 2000; 97:10209 – 10213.
- Lossos IS, Czerwinski BA, Alizadeh AA, Wechser MA, Tibshirani R, Botstein D, Levy R. Prediction of survival in diffuse large B-cell lymphoma based on the expression of six genes. *N Engl J Med* 2004;350:1828 – 1837.
- Shipp MA, Ross KN, Tamayo P, et al. Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. *Nat Med* 2002;8:68-74.
- Rosenwald A, Wright G, Chan WC, et al. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N Engl J Med* 2002; 346:1937-47.