It's ALL in the diagnosis

The molecular diagnosis of human cancer will hasten the development of treatments tailored to the abnormalities present in each patient's tumor cells. Recent gene expression profiling studies of pediatric acute lymphoblastic leukemia (ALL) suggest that the molecular diagnosis of these diseases is right around the corner.

Children with acute lymphoblastic leukemia (ALL) are treated with combination chemotherapy with an expected cure rate of 80%. This success stands as one of the crowning achievements of modern cancer therapy. Unfortunately, ALL is a common disease, and much more work needs to be done to achieve cures for all children. Furthermore, a small proportion of children treated for ALL develop secondary acute myelogenous leukemia (AML), which can be a fatal event. Thus, a major goal in ALL is to accurately diagnose patients who are at risk for either relapse or secondary leukemias, and to develop alternative therapies for these patients. This challenge has been undertaken by three recent gene expression profiling studies of pediatric ALL (Armstrong et al., 2001; Ferrando et al., 2002; Yeoh et al., 2002 [this issue of *Cancer Cell*]).

Recurrent chromosomal translocations and imbalances are the defining molecular features of ALL, and these oncogenic events identify clinically distinct subgroups of patients. In B cell ALL (B-ALL), translocations often fuse two genes, resulting in the translation of a fusion oncoprotein (Ferrando and Look, 2000). The t(12;21)/TEL-AML1 and t(1;19)/E2A-PBX1 translocations define patients with relatively good treatment outcome, especially following dose intensified chemotherapy. Patients with t(9;22)/BCR-ABL or t(4;11)/MLL-AF4 translocations have a relatively poor prognosis, but hematopoietic stem cell transplantation from a HLA-matched sibling donor has been shown to be beneficial for patients with BCR-ABL translocations (Arico et al., 2000). A hyperdiploid chromosomal content (>50 chromosomes) defines a favorable prognostic group in B-ALL, but the molecular correlates of this chromosomal imbalance have been obscure. T cell ALL (T-ALL) is characterized by recurrent, but infrequent, chromosomal translocations that result in overexpression of transcription factors that are presumed to deregulate early thymocyte development, resulting in leukemia (Ferrando and Look, 2000). As in B-ALL, these translocations define patient subsets with different clinical outcomes (Ferrando and Look, 2000).

A primary goal of the recent gene expression profiling studies was, therefore, to identify which genes are expressed in association with each chromosomal abnormality. Interestingly, this was an exceedingly easy task. Each chromosomal abnormality in B-ALL had a characteristic gene expression signature composed of a large number of genes. Indeed, the leukemia subtypes were found to be as distinct in gene expression as are various types of epithelial cancer (Ramaswamy et al., 2001; Su et al., 2001). The differentially expressed genes could be fashioned into predictors of the leukemia subtypes that functioned with extremely high accuracy. Importantly, in the largest study (Yeoh et al., 2002), the leukemia subtype predictor was developed on a training set of ALL cases and was tested on an independent set of cases in which it performed with 96% accuracy. Given the large number of statistical associations that are made during the analysis of gene expression profiling data, the use of an independent validation set is an exceedingly important analytical method that avoids the danger of overfitting statistical models to the data.

As in most large-scale gene expression profiling studies, unexpected obser-

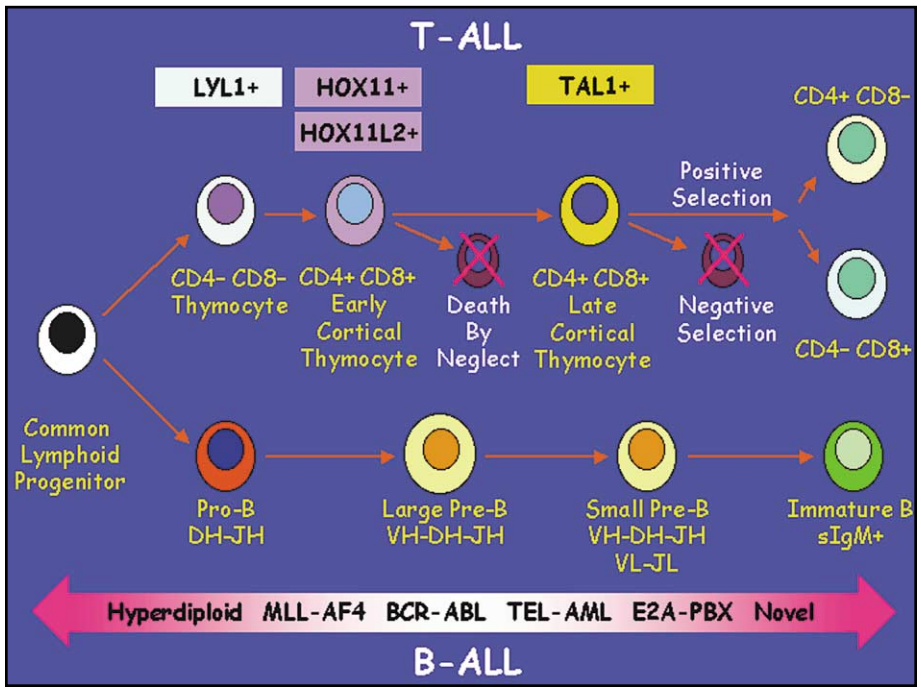

Figure 1. Molecular subtypes of acute lymphoblastic leukemia (ALL)

B-cell ALLs (B-ALL) are derived from precursor B cells in the bone marrow. The variety of chromosomal translocations and abnormalities of chromosome number that cause malignant transformation are shown. Gene expression profiles of B-ALLs define gene expression signatures of each of these oncogenic events. T cell ALLs (T-ALLs) are derived from precursor T cells in the thymus. Infrequent but recurrent translocations lead to the overexpression of the transcription factors LYL1, HOX11, HOX11L2, and TAL1. Overexpression of these genes in the absence of translocation (indicated by the + sign) occurs in more cases of T-ALL. Molecular subtypes of T-ALL defined by oncogene overexpression are characterized by gene expression signatures that suggest that the T-ALLs are blocked at discrete stages in normal thymocyte differentiation, as indicated.

vations were plentiful. 70% of the genes that were found to be preferentially expressed in the hyperdiploid leukemia subtype are encoded on chromosomes 21 and X (Yeoh et al., 2002). This unusual result is not apparently explained by numerical abnormalities of these chromosomes, but instead may indicate a regulatory mechanism that operates at the level of whole chromosomes. Another unexpected observation was that 14 B-ALL cases that lacked the recurrent B-ALL chromosomal abnormalities nonetheless shared a large gene expression signature that distinguished them from the other leukemia subtypes (Yeoh et al., 2002). These cases constitute a novel leukemia subtype for which the oncogenic event is currently unknown. Finally, 20% of the B-ALL cases could not be sorted into any of the leukemia subtypes by gene expression (Yeoh et al., 2002). Since some chromosomal translocations in B-ALL are exceedingly infrequent, it may be necessary to profile many more ALLs, possibly thousands, to define rare leukemia subtypes to which these orphan cases belong.

In contrast to the high frequency of recurrent translocations in B-ALL, only 30% of T-ALL cases have chromosomal translocations (Ferrando and Look, 2000). Gene expression profiling of T-ALL provided a fascinating explanation for this disparity; namely, that the oncogenes HOX11, TAL1, and LYL1 that are involved in T-ALL translocations can also be overexpressed by other means in cases lacking translocations (Ferrando et al., 2002). The overexpression of each T-ALL oncogene was associated with a characteristic gene expression signature, and these signature genes were used to organize all T-ALL cases into clusters with related gene expression. In this way, some cases that did not overexpress a particular oncogene were nevertheless grouped with cases that did, based on similar expression of other genes. Intriguingly, the HOX11 cluster included some cases that overexpressed a highly related gene, HOX11L2, instead of HOX11. Importantly, other oncogenic events, such as deletion of the p16/ARF locus, were found to be selectively present in certain T-ALL subtypes defined in this manner, suggesting that these subtypes represent pathogenetically distinct diseases.

Deregulation of oncogenes appears to be central to the molecular diagnosis of ALL, but what is the mechanism by which these oncogenes give rise to distinct leukemia subtypes? The clearest answer to this question comes from the study of T-ALL in which the leukemia subtypes defined by oncogene overexpression corresponded to different stages in normal thymocyte differentiation (Ferrando et al., 2002), presumably indicating that the oncogenes act by blocking differentiation (Figure 1). The cell of origin of a tumor can be important in determining the clinical outcome of patients, as shown for diffuse large B cell lymphoma (Alizadeh et al., 2000), since stages of normal differentiation vary in proliferative potential and propensity for apoptosis. Indeed, the T-ALL subtypes were found to define clinically distinct patient subsets (Ferrando et al., 2002). The association of the B-ALL subtypes with normal B cell differentiation is less clear (Yeoh et al., 2002), but those cases with the MLL-AF4 translocation most resembled very early B cell precursors that coexpress both B cell and myeloid markers (Armstrong et al., 2001). Gene expression profiling of well-defined subsets of human B cell precursors will be required to carefully assess the relationship between B-ALL subtypes and normal B cell development.

From the clinical standpoint, DNA microarrays represent a viable alternative to conventional karyotyping and fluorescence in situ hybridization for the diagnosis of leukemia subtypes, especially since gene expression profiles can identify cases with translocations that are cytogenetically invisible (Yeoh et al., 2002). In B-ALL, gene expression profiles may be able not only to classify patients into leukemic subtypes, but also to predict treatment failure (Yeoh et al., 2002). Interestingly, gene expression-based predictors of relapse only worked among cases that shared a common chromosomal abnormality (Yeoh et al., 2002). This result suggests that the leukemia subtypes differ profoundly in oncogenic mechanisms and utilization of signaling pathways, and therefore have distinct mechanisms to resist chemotherapy. An intriguing, but puzzling, result was that development of secondary AML could also be predicted from gene expression in the B-ALL blasts, even though the AML and B-ALL cells are thought to arise from different precursor cells (Yeoh et al., 2002). Conceivably, a subset of B-ALL patients could have an underlying molecular lesion in a hematopoietic stem cell that predisposes them to secondary AML and is reflected in the gene expression profiles of their B-ALL blasts.

An important caveat is that the predictors of patient survival presented in these gene expression profiling papers will need to be validated in an independent set of cases. Ideally, DNA microarray technology will be included in any new prospective clinical trial in pediatric ALL and, for that matter, in all cancer trials. It will be interesting to compare the gene expression profiles of adult and pediatric ALLs, since the same translocations are recurrent in adult ALL, yet adults fail therapy more often. Now that the molecular diagnosis of ALL is well underway, the burden will shift toward identifying novel therapies that target the specific regulatory pathways active in each leukemia subtype, leading, we hope, to the eradication of this disease.

Louis M. Staudt

Metabolism Branch
Center for Cancer Research
National Cancer Institute
9000 Rockville Pike
Bethesda, Maryland 20892
E-mail: lstaudt@mail.nih.gov

Selected reading

Alizadeh, A.A., Eisen, M.B., Davis, R.E., Ma, C., Lossos, I.S., Rosenwald, A., Boldrick, J.C., Sabet, H., Tran, T., Yu, X., et al. (2000). Nature *403*, 503–511.

Arico, M., Valsecchi, M.G., Camitta, B., Schrappe, M., Chessells, J., Baruchel, A., Gaynon, P., Silverman, L., Janka-Schaub, G., Kamps, W., et al. (2000). N. Engl. J. Med. *342*, 998–1006.

Armstrong, S.A., Staunton, J.E., Silverman, L.B., Pieters, R., den Boer, M.L., Minden, M.D., Sallan, S.E., Lander, E.S., Golub, T.R., and Korsmeyer, S.J. (2001). Nat. Genet. *30*, 41–47.

Ferrando, A.A., and Look, A.T. (2000). Semin. Hematol. *37*, 381–395.

Ferrando, A.A., Neuberg, D.S., Staunton, J., Mignon, L.L., Huard, C., Raimondi, S.C., Behm, F.G., Pui, C.-H., Downing, J.R., Gilliland, D.G., et al. (2002). Cancer Cell *1*, 75–87.

Ramaswamy, S., Tamayo, P., Rifkin, R., Mukherjee, S., Yeang, C.H., Angelo, M., Ladd, C., Reich, M., Latulippe, E., Mesirov, J.P., et al. (2001). Proc. Natl. Acad. Sci. USA *98*, 15149–15154.

Su, A.I., Welsh, J.B., Sapinoso, L.M., Kern, S.G., Dimitrov, P., Lapp, H., Schultz, P.G., Powell, S.M., Moskaluk, C.A., Frierson, H.F., Jr., and Hampton, G.M. (2001). Cancer Res. *61*, 7388–7393.

Yeoh, E.-J., Ross, M.E., Shurtleff, S.A., Williams, W.K., Patel, D., Mahfouz, R., Behm, F.G., Raimondi, S.C., Relling, M.V., Patel, A., et al. (2002). Cancer Cell *1*, this issue, 133–143.