Immunity, Vol. 15, 375–385, September, 2001, Copyright ©2001 by Cell Press

Signatures of the Immune Response

Review

A.L. Shaffer, Andreas Rosenwald, Elaine M. Hurt, Jena M. Giltnane, Lloyd T. Lam, Oxana K. Pickeral, and Louis M. Staudt[1]
Metabolism Branch
National Cancer Institute
National Institutes of Health
9000 Rockville Pike
Building 10, Room 4N114
Bethesda, Maryland 20892

Summary

A compendium of global gene expression measurements from DNA microarray analysis of immune cells identifies gene expression signatures defining various lineages, differentiation stages, and signaling pathways. Germinal center (GC) B cells represent a discrete stage of differentiation with a unique gene expression signature. This includes genes involved in proliferation, as evidenced by high expression of G2/M phase regulators and low expression of ribosomal and metabolic genes that are transcriptional targets of c-*myc*. GC B cells also lack expression of the NF-κB signature genes, which may favor apoptosis. Finally, the transcriptional repression signature of BCL-6 reveals how this factor can prevent terminal differentiation of B cells and cause B cell lymphomas.

Introduction

The light at the end of the reductionist tunnel has turned out to be the headlight of a locomotive called genomics. Now that complete genome sequences are known for many organisms, including ourselves, we are faced with the daunting task of resynthesis. How can we make functional sense of a genome's worth of nucleotide sequence? We need a map. A powerful cartographic tool for exploring the genome is gene expression profiling, a technology that can provide an overview of the genome interacting with its environment. Taking this genomic view of the immune system, one sees a wonderfully integrated set of gene expression programs that allow immune cells to respond to internal developmental cues and external antigenic challenges. Although there are many ways to monitor global gene expression, we will focus on the insights gained using cDNA microarrays (Staudt and Brown, 2000).

The application and analysis of cDNA microarrays have been reviewed previously (Alizadeh and Staudt, 2000; Staudt, 2001). In brief, a substrate, typically a glass slide or silicon wafer, is arrayed with cDNAs or oligonucleotides complementary to transcripts from several thousand genes. We have developed a microarray referred to as the Lymphochip which holds more than 10,000 individual human cDNAs on a single slide (Alizadeh et al., 1999, 2000). The Lymphochip gene set is enriched for genes expressed in lymphocytes, including thousands of named genes of known function as well as genes of unknown function derived from normal and malignantly transformed lymphocyte cDNA libraries. The mRNA expression from two samples is compared on a single microarray by converting each RNA into a cDNA probe that incorporates a different fluorochrome (Cy5 or Cy3). Following hybridization of these probes to a single array, the array is washed and scanned to reveal the hybridization of the probes to each spot on the array. These Cy5 and Cy3 fluorescence measurements are converted into ratios of relative gene expression (Cy5/Cy3) that accurately reflect the true difference in mRNA levels between samples (Schena et al., 1995, 1996; Shaffer et al., 2000; Taniguchi et al., 2001).

A single microarray provides thousands of individual gene expression measurements and, in a short period of time, a scientist can perform a series of microarray experiments that populate a database with many millions of measurements. Though there are many ways to parse these enormous gene expression data sets (Bassett et al., 1999; Sherlock, 2000), we use hierarchical clustering, based on Pearson correlation relationships, as the primary tool for data mining (Eisen et al., 1998). Clustering can group genes with similar expression patterns across all samples, group samples with similar gene expression patterns, or both. This process defines gene expression signatures composed of genes that are coordinately expressed in samples related by some identifiable criterion such as cell type, differentiation state, or signaling response.

Figure 1 presents a compendium of gene expression data assembled from 986 Lymphochip experiments (columns) and representing approximately 2000 genes measured by 3418 array elements (rows). Data are displayed using a color scale to indicate the level of gene expression relative to an appropriate control sample; red denotes higher and green denotes lower expression. Horizontal colored bars indicate the cell type and/or conditions tested, whereas vertical black bars delimit gene expression signatures as described below.

Gene Expression Signatures of Lineage and Developmental Stage

Immunologists have used many criteria to define developmental stages as cells progress from immature precursors to terminally differentiated effectors, including DNA rearrangement status, surface marker expression, and protein secretion. Microarrays provide a new, comprehensive way to define these stages by their characteristic gene expression signatures.

Signatures defining the B and T lymphocyte lineages are readily identified in this gene expression map (Figure 1). The T cell signature contains components of the T cell receptor (TCRα, TCRβ, CD3δ, CD3γ), signaling proteins (LAT, TRIM, SAP), and surface markers (CD5, CD2). An extensive B cell signature is defined by components associated with the B cell receptor (μ heavy chain, CD79A, CD79B, syk, btk, lyn), as well as other cell surface receptors (CD20, CD19, CD22, CD27, CD40, CD45,

[1]Correspondence: lstaudt@mail.nih.gov

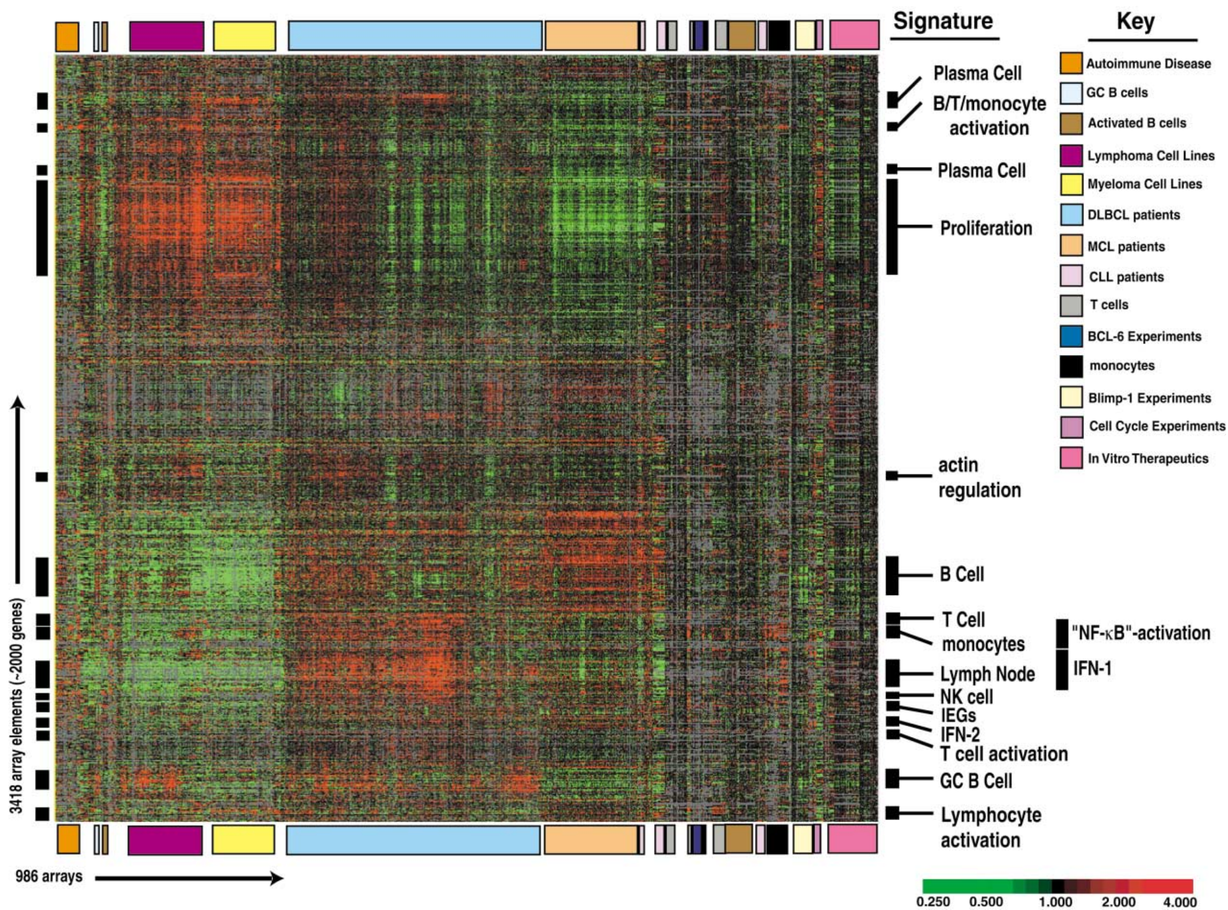

Figure 1. Lymphochip Clustering Analysis Defined Gene Expression Signatures

Gene expression measurements are shown from 986 microarray experiments investigating normal immune cells and disease states (see key at right). 3814 common array elements (approximately 2000 unique genes) were analyzed by hierarchical clustering. Data filters based on spot quality and signal intensity were used to capture the most reliable data. Spots not meeting these criteria were excluded from analysis and appear gray. Relative gene expression levels (ratios) are indicated by the color bar in the lower right. Gene expression signatures are described in detail in the text.

and CD72) and transcription factors (PAX5, OCT-2, Spi-B) important in B cell development.

Microarray analysis also identifies the distinct genetic programs characterizing different developmental stages within a lineage. Two stages of B cell differentiation, plasma cells and germinal center (GC) B cells, have their own signatures (Figure 1). The plasma cell signature is defined by two gene clusters that include two transcriptional regulators of plasmacytic differentiation, blimp-1 and XBP-1, as well as the adhesion molecule syndecan-1 (Messika et al., 1998; Reimold et al., 2001; Turner et al., 1994). Microarray analysis also defines a GC B cell-specific signature which will be discussed in detail below (Figures 1 and 4, see also Alizadeh et al., 2000; Ma and Staudt, 2001).

There is also a prominent signature representing several nonlymphoid lineages that arises from the analysis of lymph node biopsies taken from patients with diffuse large B cell lymphoma (DLBCL) (Figure 1 and Alizadeh et al., 2000). This includes several macrophage (CD18, CD64) and NK cell (NK4, CD49, NKG5) markers in addition to genes expressed in activated monocytes (IL-8, IL-1β, LARC, MCP-1, GRO2). Many genes involved in the organization of secondary lymphoid organs are found in this signature: chemokines (BLC, APRIL), matrix components (fibronectin, collagen), and matrix regulators (matrix metalloproteases [MMPs], tissue inhibitors of metalloproteases [TIMPs]). Whether this signature represents a normal lymph node signature or reflects the reaction of the nonlymphoid lineages to the B cell tumor remains an interesting area to be explored.

Gene Expression Signatures of Cellular Physiology

The global view of gene expression provided by microarrays makes it clear that certain cellular processes, like proliferation, are coordinated at the level of gene expression. By far the largest set of coherently regulated genes in Figure 1 are those found in the proliferation signature (Alizadeh et al., 2000; Perou et al., 1999). This signature includes genes involved in cell cycle regulation (cyclins, cyclin-dependent kinases, cell cycle checkpoint proteins), DNA synthesis (nucleotide metabolism enzymes, polymerases), and protein translation (ribosomal proteins, translation elongation factors). Within this large

proliferation signature, a fascinating substructure is observed in which genes are grouped according to their function in the cell cycle, as discussed in detail below (Figure 4B).

Another functional signature comprises a set of genes involved in actin regulation. At the core of this signature are β and γ actin, which are involved in cell motility, cytokinesis, and endocytosis. Two other genes in this signature are also involved in actin polymerization based on their interaction with the Wiskott-Aldrich Syndrome protein (WASP): G25K, the human homolog of cdc24 and a ras superfamily GTPase, and WIP (WASP-interacting protein). The remaining actin signature proteins bind actin directly. ARP 2/3 induces actin polymerization, whereas GAS2 associates with actin fibers mediating cell cycle arrest. Thymosin β4 controls actin dynamics by sequestering actin.

Coordinate expression of actin regulatory genes may orchestrate the reorganization of the actin cytoskeleton, an essential step in lymphocyte activation (reviewed in Bauch et al., 2000; Dustin and Cooper, 2000; Penninger and Crabtree, 1999). Although it was observed nearly 30 years ago that crosslinking-induced “capping” of antigen receptors on B cells was dependent on actin polymerization, it was not until T cell activation was studied that the role of actin was truly appreciated. Actin plays two roles in T cell activation: formation of the synapse between the T cell and the antigen-presenting cell (APC) and organization of intracellular signaling components. Resting T cells have actin structures that sequester antigen, chemokine, and adhesion receptors in inaccessible regions of the plasma membrane. Upon activation, reorganization of actin microfilaments leads to the formation of SMACs (supramolecular activation clusters) between the T cell and the APC, thereby promoting sustained interactions between the T cell receptor (TCR) and MHC molecules and initiating intracellular signals. These TCR signals in turn activate Cdc42Hs, which causes WASP to stimulate the Arp2/3 complex, leading to remodeling of the actin cytoskeleton. Two lines of evidence point to the absolute requirement for actin cytoskeleton reorganization in promoting and sustaining T cell activation. First, drugs that disrupt actin polymerization, like cytochalasin, immediately curtail T cell activation. Second, T cells from patients with Wiskott-Aldrich syndrome, an X-linked immunodeficiency characterized by a deficiency in WASP, fail to properly proliferate or secrete IL-2 upon TCR engagement (Snapper and Rosen, 1999). The fact that actin and actin regulatory genes are coordinately regulated was not appreciated prior to microarray analysis and underscores the importance of the actin microfilament system in lymphocyte activation.

Gene Expression Analysis of Signaling Pathways

As immune cells progress through development and respond to antigenic challenge, they trigger signaling pathways that alter the activity of transcription factors, changing their gene expression profiles. Since these same signaling modules are engaged in diverse responses and cell types, the downstream target genes of these signaling events are grouped together in gene expression maps. Three signaling pathways are represented by prominent gene expression signatures in Figure 1: the interferon pathway, the immediate early pathway, and the NF-κB pathway.

A two-part signature involves interferon responsive genes. The “split” nature of this gene expression signature relates to the different biological functions of these interferon-inducible genes, which have recently been compiled in a searchable database (de Veer et al., 2001). Interferons signal primarily through the transcription factors STAT-1, STAT-2, and the IRFs (reviewed in Chatterjee-Kishore et al., 2000; Stark, 1997; Stark et al., 1998). However, in STAT-1-deficient cells, STAT-1-independent interferon signaling has been observed, indicating that interferon signals are also processed by alternate pathways (Meraz et al., 1996). STAT-1 itself can also be activated by many noninterferon stimuli (Dimberg et al., 2000; Seebach et al., 2001; Wong et al., 2001), suggesting that although the interferon-inducible genes cluster together, they may do so independent of the action of interferons themselves. The first interferon signature (IFN-1, Figure 1) contains genes involved in protein catabolism. Three components of the proteosome (macropain, LMP2, and MECL-1) are grouped with a component of the ubiquitin-conjugating system, E2L6. Components of the proteosome pathway have been observed to be coordinately transcribed in yeast (Eisen et al., 1998), and ubiquitin-conjugating enzymes have been shown to be upregulated at the protein level in T cells treated with interferon-α (Nyman et al., 2000). In addition, MHC class I and β2 microglobulin genes appear in this signature, indicating that a program dedicated to tagging, degrading, and presenting intracellular antigens is coordinately controlled. This signature also contains the interferon-responsive transcription factors, STAT-1 and IRF-1, that crossregulate one another. The second interferon signature (IFN-2) contains genes that promote resistance to viral infection. Present in this signature are the Mx genes, GTPases that mediate resistance to influenza. The IFN-2 signature also includes 2′5′-oligoadenylate synthase whose products activate RNase L, an enzyme that can degrade viral RNA. Interestingly, HEM45, a functional relative of RNase L, is also present in this signature. Another antiviral gene in this signature is interferon-inducible double stranded RNA-dependent kinase, an enzyme that terminates translation in infected cells by phosphorylating elongation factors. The IFN-2 signature reflects a coordinated effort by cells to reclaim control of their intracellular machinery upon viral infection.

When the cellular environment acutely changes, either beneficially as in the presence of growth factors, or stressfully due to heat shock or radiation, cells express a characteristic set of genes known as immediate-early genes (IEGs). The importance of these rapidly activated genes is emphasized by the fact that multiple growth factors lead to the upregulation of the same stereotypical set of IEGs and that distinct signaling pathways initiated from the same receptor independently cause the upregulation of these genes (Fambrough et al., 1999). The genes involved in the immediate-early response form a distinct cluster in Figure 1. Among these rapidly inducible genes are the prototypical IEGs c-*fos* and c-*jun*, two components of the transcription factor AP-1 (reviewed in Wagner, 2001) that is involved in activating gene expression from a host of immunologically

important genes involved in growth and differentiation (Foletta et al., 1998; Shaulian and Karin, 2001). Several other genes of diverse function make up this signature: HSP70, named for its role in heat shock responses and involved in the folding of proteins; MKP-1, a phosphatase that dephosphorylates MAP kinases; CPBP, a transcription factor that enhances initiation from TATA-less promoters; and TTP, which binds certain mRNAs and promotes their degradation. The functional logic that knits together these diverse IEGs will take more experimentation to fully understand.

Stimulation of lymphocytes through their antigen receptors initiates cell cycle entry and changes gene expression, a response generically referred to as "activation." Likewise, cells of the innate immune system responding to invading pathogens become activated and alter their gene expression profiles. Several distinct activation signatures are present in Figure 1, the most prominent of which features targets of the transcription factor NF-κB (Figure 2). NF-κB is shorthand for a family of related proteins which dimerize to form transactivating complexes (Baldwin, 1996; Ghosh et al., 1998; Karin and Ben-Neriah, 2000; Pahl, 1999; Siebenlist et al., 1994). NF-κB dimers are sequestered in the cytoplasm by interaction with inhibitory proteins (the IκBs). Various stimuli activate kinase signaling cascades that result in the phosphorylation and degradation of IκB, thereby releasing NF-κB to translocate to the nucleus where it activates transcription of target genes. Several studies have emphasized the essential role of this transcription factor family in regulating genes at critical points in lymphocyte development and activation (Gugasyan et al., 2000). Many NF-κB targets are antiapoptotic (A1, BCL-XL, BCL-2, c-IAP2), which may explain the importance of the NF-κB pathway in oncogenesis and resistance of tumors to chemotherapy (Baldwin, 2001; Rayet and Gelinas, 1999). Other genes in this signature are proinflammatory cytokines and chemokines (MIPs, IL-8, etc.) that are acutely upregulated during an immune response. The remaining genes included components of the NF-κB pathway itself: RelB, NFκB2, IκBα, and IκBϵ.

Because expression of any given gene depends upon integrating the input from multiple pathways, it is not surprising that some NF-κB target genes appear in other parts of the cluster analysis (Figures 1 and 2). An activation signature representing genes induced during B cell, T cell, and monocyte activation includes the NF-κB target Bcl-XL, and an activation signature representing genes induced in lymphocytes includes the NF-κB targets BCL-2, IRF-4, and cyclin D2. A final activation signature is comprised of a few genes expressed specifically in activated T cells, notably the cytokines IL-2 and IL-4, whose transcription relies on the activation of both NF-κB and NFAT. Thus, gene expression signature analysis alone may not provide a complete picture of the genes that are targeted by a particular signaling pathway, and the next sections detail other approaches that are additionally helpful in delineating the gene expression consequences of intracellular signaling events.

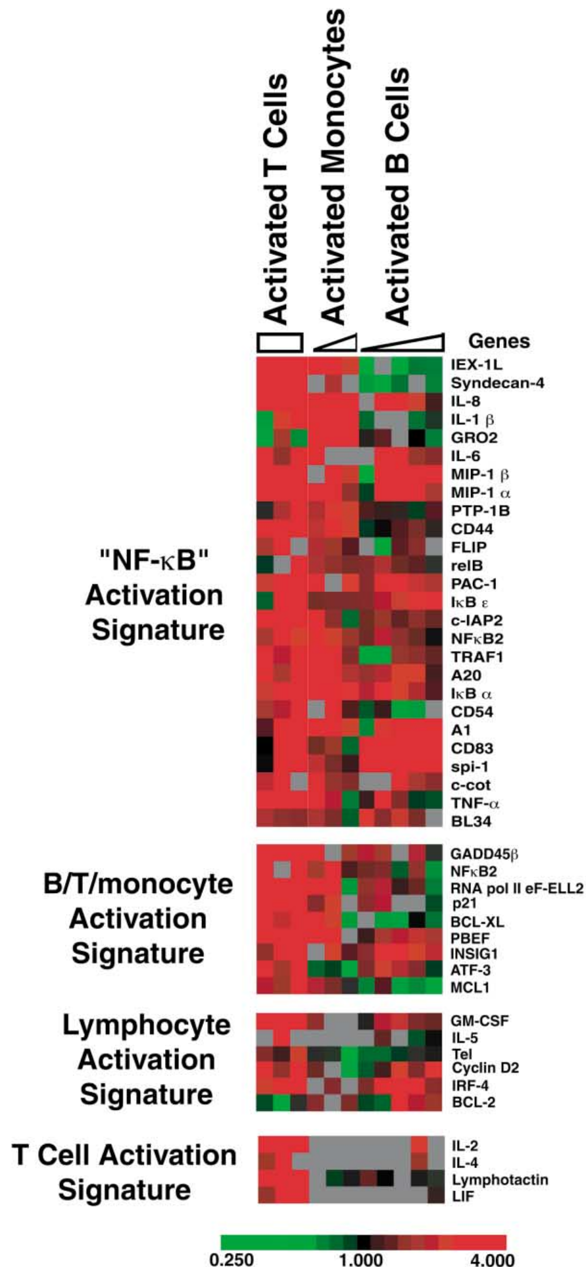

Figure 2. Gene Expression Signatures of Immune Cell Activation

Each signature represents a group of genes that cluster together based on similarity in gene expression patterns across the entire set of microarray experiments shown in Figure 1. The behavior of these genes is shown only for primary T cells activated by PMA plus ionomycin or Con A plus PHA for 3 hr, primary monocytes activated with LPS for 2 and 8 hr, and purified peripheral blood B cells activated with anti-IgM, CD40L, and IL-4 over a time course at 0, 1, 3, 6, and 24 hr. Red indicates high expression; green represents low expression; and gray indicates that the gene was not expressed at a level meeting confidence criteria.

Gene Expression Physiology of Immune Responsiveness

To reveal the downstream effects of a signaling pathway by gene expression analysis, we would ideally like to perturb the pathway discretely and measure how the system responds to the perturbation. We have many tools at our disposal: genetics (via the use of transgenic animals or the study of patients with disease), pharmacology, and molecular biology (transfection or transduction of activators or inhibitors of a pathway).

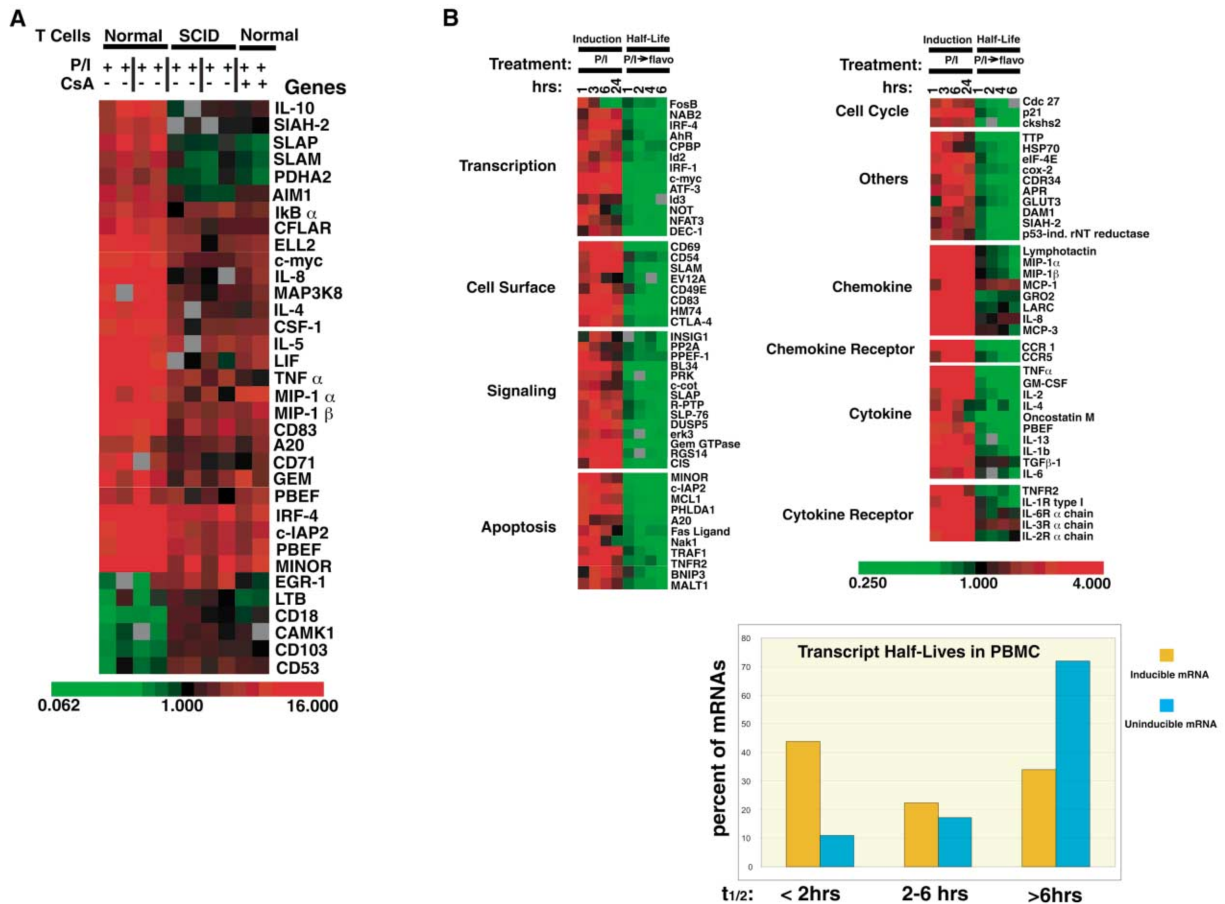

Figure 3. Gene Expression Physiology of Immune Responses

(A) Calcium-dependent T cell signaling. Microarray analysis of normal versus calcium entry-deficient (SCID) T cells is shown. T cells of normal controls and SCID patients were left unstimulated or stimulated with PMA and ionomycin (P/I) for 3 hr. In addition, some control T cells were preincubated with cyclosporin A (CsA) for 30 min prior to stimulation. Genes upregulated relative to the unstimulated samples appear red, and downregulated genes appear green. See text and Feske et al. (2001) for details.

(B) Genomic-scale analysis of mRNA stability: differential half-lives of functional classes of mRNAs. PBMC were stimulated with PMA and ionomycin (P/I) for the indicated times to identify those transcripts with rapid induction kinetics. Genes induced by P/I appear red, and genes repressed appear green. To determine the half-lives of these transcripts, cells were stimulated with P/I for 3 hr and then treated with the global transcriptional inhibitor, flavopiridol (flavo), and harvested at the indicated times. Shades of green represent the extent of decrease in mRNA abundance relative to the control sample not treated with flavopiridol. See Lam et al. (2001) for details. The bar graph shows the percentage of inducible and noninducible genes within each half-life category.

The Role of Calcium Signaling in Lymphocyte Activation

In a study of T cell receptor signaling, genetics was used to elucidate the role of calcium signals in regulating gene expression (Feske et al., 2001). A rare family of severe combined immunodeficiency (SCID) patients was found to have an impaired NFAT activation. NFAT transcription factors turn on many immunologically important genes, especially during T cell activation (Crabtree, 1999; Rao et al., 1997). In a first set of experiments, normal and SCID T cells were compared, either unstimulated or treated with PMA and ionomycin. Many genes induced in normal T cells failed to be induced in SCID patient T cells. Unexpectedly, microarray analysis revealed a set of genes whose mRNA levels were strongly downregulated by calcium signaling in normal T cells but were not as affected in SCID patient T cells (Figure 3A). Treatment of T cells with cyclosporin A (CsA) was used to assess whether calcium-regulated genes depended upon the action of the phosphatase calcineurin. Most calcium-regulated genes were dependent on calcineurin signaling, though a smaller number were less affected by CsA than by the SCID defect (e.g., MIP-1α and PBEF, Figure 3A). A single gene, TNF-α, was unique in its behavior, as it was more strongly induced in SCID patient T cells than in normal T cells treated with CsA. Interestingly, the downregulation of genes by calcium signaling was also dependent on calcineurin. Calcineurin dephosphorylates and activates NFAT, but it could conceivably affect other targets resulting in repression of genes. This example illustrates how the combination of genetics and pharmacological inhibitors can be useful in dissecting clinically relevant signaling pathways.

Similar approaches have been taken in the study of B cell signaling (Glynne et al., 2000). Anergy is a form of tolerance in which signals through the B cell receptor

are altered, thereby preventing activation upon antigen stimulation. Using a well-defined transgenic system, the HEL-Ig:HEL-transgenic mouse, the gene expression signature of activated and anergic B cells was compared. Anergic B cells showed a distinct gene expression signature marked by their failure to upregulate genes promoting cell cycle entry and preventing apoptosis, while continuing to express genes that block activation. This study also examined whether the immunosuppresive drug FK506, which interrupts signaling via the calcium/calcineurin pathway, has a gene expression effect akin to anergy. In fact, FK506 failed to mimic functional anergy since it blocked only a fraction of activation-induced genes and interfered with the induction of anergy-specific genes. New immunosuppressant agents that mimic more closely the gene expression profile of anergy will be of greater clinical utility (Glynne et al., 2000).

mRNA Stability and the Control of Immune Responses

The regulation of mRNA stability is a less well-studied aspect of gene regulation that is as important as transcriptional initiation in determining steady-state mRNA levels. Microarrays were used to assess mRNA stability on a genomic scale by studying cells treated with global inhibitors of transcription initiation (Lam et al., 2001). An unexpected relationship between the mRNA half-life of a given gene and the function of its encoded protein was uncovered. Prior to this analysis, there was no reason to assume a priori that genes related by function would also be related in terms of mRNA stability. For example, genes involved in M phase of the cell cycle (e.g., pLK, CHK1) had transcripts with short half-lives. Since M phase is typically the shortest phase of the cell cycle, it is not surprising that mRNAs encoding proteins involved in this cell cycle step would be rapidly turned over. Similarly, the class of genes encoding antiapoptotic factors (A20, MCL-1, c-IAP) was relatively enriched in labile mRNAs, suggesting that the prevention of programmed cell death is a dynamic process that must be capable of rapid modulation.

The most striking relationship uncovered was that transcriptionally inducible genes are much more likely to produce labile mRNAs than are noninducible genes (Lam et al., 2001). This relationship was found by studying the mRNA half-lives of genes that were induced during activation of peripheral blood mononuclear cells (PBMC) with PMA and ionomycin (Figure 3B). A majority of inducible transcripts have very short half-lives (<2 hr), whereas transcripts whose levels were unaffected by activation tend to be long lived (>6 hr). A clear regulatory logic emerges from this observation. Presumably, the inducible genes require both rapid up- and downmodulation of their mRNA levels in order to, on one hand, respond acutely to changes in cellular physiology but, on the other hand, terminate this response rapidly. This response profile can only be achieved by transcriptional regulation coupled with a short mRNA half-life.

The stability of mRNAs encoding cytokines, chemokines, and their respective receptors showed an interesting dichotomy (Figure 3B). Cytokine transcripts were found to have substantially shorter half-lives than those for chemokines. Conversely, mRNAs for cytokine receptors were quite stable compared to those for chemokine receptors. If one considers the roles of each of these proteins during an immune response, the logic of this regulation becomes clear. Cytokines are powerful inducers of immune responses and regulators of effector functions. To curtail potential bystander effects of cytokines, cells may limit the duration of cytokine production by rapidly degrading cytokine mRNAs. Many cells, however, are poised to respond to cytokines, hence the greater stability of mRNAs for cytokine receptors. Chemokines, on the other hand, guide cells to sites of antigenic challenge, a process that may last for several days. Cells may therefore need to synthesize chemokines for extended periods of time, and this is facilitated by the relative stability of chemokine mRNAs. Once cells have reached the appropriate site, rapid downregulation of certain chemokine receptors, promoted by rapid mRNA degradation, may allow the cells either to remain at the site or migrate in response to other chemoattractants.

The Germinal Center B Cell from Many Gene Expression Perspectives

The multiple experimental and analytical approaches to gene expression profiling presented thus far can be fruitfully combined to understand the biology of a complex immune reaction. To gain insights into the biology of the germinal center (GC) reaction, the gene expression profiles of GC B cells were analyzed using gene expression signatures that define B cell developmental stages, cell cycle dynamics, signaling pathways, and transcription factor targets (Figure 4).

GC B cells represent a unique stage of differentiation that arises when B cells encounter antigen in secondary lymphoid tissues (Figure 5). With help from T cells and dendritic cells, B cells become activated and can differentiate into short-lived plasmablasts and plasma cells. Alternatively, they may participate in the creation of the complex microenvironment called the germinal center. In the GC, rapid proliferation is coupled to a process of antigen receptor mutation and selection called affinity maturation, resulting in the production of B cells with higher affinity antigen receptors than the cells that founded the GC. After exiting the GC, GC B cells differentiate into either memory B cells or plasma cells (McHeyzer-Williams et al., 2001).

To gain insight into the biology of GC B cells, we identified genes characteristic of this developmental stage in two separate ways (Figure 4A). The upper part of this panel shows those genes that comprise the GC B cell signature as defined by the gene expression map in Figure 1. We observed, however, that certain genes believed to be GC specific, like AID, were missing from this signature. Therefore, these genes were extracted mathematically from the array data based on higher average expression in GC B cells as compared to quiescent B cells, and these genes are shown in the lower panel. These genes most likely integrate signals from multiple pathways, which causes them to migrate to distinct clusters in large gene expression maps.

The gene expression profile of GC B cells includes many previously identified differentiation markers such as CD10, CD38, BCL-6, A-myb, and the human homolog of the mouse GC-specific gene M17. In addition to those

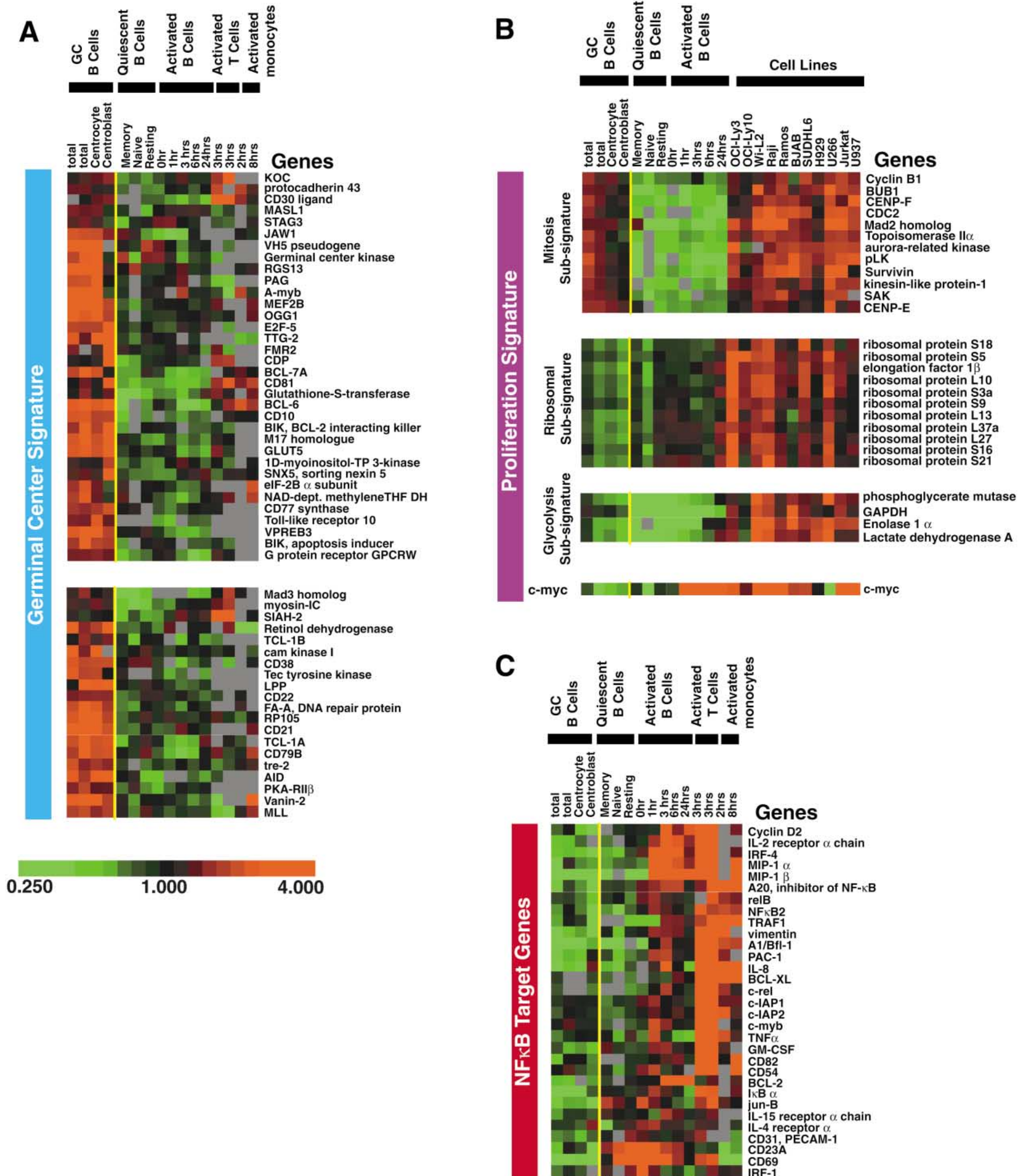

Figure 4. Multiple Microarray Perspectives on GC B Cell Biology

(A) GC B cell developmental signature. Gene expression data from primary B cells (Ma and Staudt, 2001), T cells, and monocytes, with or without activation as described in Figure 2, are presented. The GC B cell signature derived from Figure 1 is shown in the upper portion of the panel. Genes more highly expressed by GC B cells but not part of this signature were extracted mathematically and are displayed in the lower panel (see text).

(B) Expression of proliferation signature genes in GC B cells. The same samples as shown in (A) are displayed with the addition of a series of B cell lines (OCI-Ly3 through U266), as well as a T cell line (Jurkat) and a monocyte cell line (U937). Subclusters of genes involved in various functions (mitosis, protein synthesis, and glycolysis) are shown.

(C) Expression of NF-κB target genes in GC B cells. Samples are the same as in (A).

named genes illustrated in Figure 4A, the Lymphochip microarrays contain a host of novel genes that are germinal center-restricted in their expression. It is important to emphasize that the expression of these GC B cell signature genes is maintained in lymphoma cell lines derived from GC B cells (Alizadeh et al., 2000); thus, the GC B cell program can be maintained in the absence of interacting T cells or follicular dendritic cells. This observation demonstrates that the GC B cell represents a discrete stage of B cell differentiation with a stable gene expression program and not simply a particular mode of B cell activation characterized by a transient change in gene expression.

GC B cells divide every 6 hr, placing them among the most rapidly proliferating cells in the body. To understand the highly proliferative nature of GC B cells, we examined the expression of the proliferation signature genes in GC B cells (Figure 4B). Within the overall structure of the large proliferation signature, a remarkable substructure was discerned that groups genes having distinct proliferation functions together. One subcluster of genes encodes proteins that participate in the G2/M phase of the cell cycle. These genes include the cyclin/cdk pair cyclin B1/CDC2, which drives cells through G2/M phase, and pLK, which regulates cyclin B1 subcellular localization. Also found are genes encoding the centrosome proteins CENP-E and CENP-F and aurora-like kinase that help organize the mitotic spindle and genes encoding the mitotic checkpoint proteins BUB1, Mad2, and survivin. Compared to other primary B cells, GC B cells highly expressed these G2/M phase genes, reflecting the higher proportion of GC B cells at this stage of the cell cycle.

Surprisingly, two other clusters of genes in the proliferation signature, encoding proteins involved in translation and energy metabolism, were downregulated in GC B cells. The first cluster is composed almost entirely of ribosomal subunit genes, and the second is involved in energy production via the glycolytic pathway. Both sets of genes contain known targets of the transcription factor c-*myc* (Guo et al., 2000; Schuhmacher et al., 2001), and GC cells have a markedly lower level of c-*myc* mRNA than mitogenically activated blood B cells (Figure 4B). Furthermore, one of the GC B cell signature genes is a homolog of the mouse Mad3 gene that dimerizes with c-*myc* and opposes its function as a transcriptional activator. Taken together, this analysis of the proliferation signature genes revealed that GC B cells favor cell proliferation at the expense of cell growth, a previously unappreciated aspect of GC B cell physiology.

Another perspective on GC B cell physiology came from an analysis of NF-κB target genes in these cells. In addition to the NF-κB signature genes that we described above, we also assembled a list of NF-κB target genes from the published literature that were present on the Lymphochip microarray (Hinz et al., 2001; Lee et al., 1999; Pahl, 1999) (Figure 4C). It is immediately apparent that GC B cells fail to express most NF-κB target genes including many components of the NF-κB signaling pathway itself (relB, NFκB2, and IκBα). As previously discussed, many NF-κB targets are antiapoptotic (Bcl-2, Bcl-XL, A1, c-IAP-2), and these are poorly expressed in GC B cells. Thus, pathway analysis of NF-κB target genes in GC B cells suggests that these cells are poised to undergo programmed cell death. In keeping with this notion, one of the GC B cell signature genes is BIK, a member of the BCL-2 family that promotes apoptosis.

Finally, we have begun to study GC B cell physiology by manipulating individual regulatory factors in these cells and monitoring the resultant changes in gene expression. In an initial application of this approach, we identified the target genes of the transcriptional repressor BCL-6 (Shaffer et al., 2000). BCL-6 is normally expressed in GC B cells, and its misexpression is associated with many B cell non-Hodgkin's lymphomas, DLBCL in particular. To elucidate the role of BCL-6 in normal B cells, BCL-6 knockout mice were generated (Dent et al., 1997; Fukuda et al., 1997; Ye et al., 1997). In keeping with its expression pattern, ablation of BCL-6 resulted in the complete inability to form germinal centers, and this defect was autonomous to the B cell lineage.

To identify BCL-6 target genes, BCL-6 was introduced into B cell lines that lack expression of this factor or, alternatively, a dominant-negative form of BCL-6 was introduced into GC B cell-derived cell lines expressing the endogenous BCL-6 gene. Microarray analysis of these manipulated cells revealed that BCL-6 represses a discrete set of genes involved in B cell activation, inflammation, and terminal differentiation (Figures 5A and 5B). BCL-6 targets include a set of genes that are induced by signaling through the B cell antigen receptor (CD44, CD69, EBI2, cyclin D2, MIP-1α). These target genes are normally induced in activated B cells prior to plasmacytic differentiation, and thus BCL-6 may promote GC differentiation by blocking the alternative cell fate, namely local plasmacytic differentiation in the periarteriolar lymphoid sheath (Figure 5B). Perhaps the most interesting target of BCL-6 repression is another transcriptional repressor, blimp-1 (Shaffer et al., 2000), that promotes plasma cell differentiation. By repressing blimp-1, BCL-6 blocks terminal differentiation of GC B cells, a notion that was directly validated by the demonstration that B cells expressing dominant-negative BCL-6 displayed features of plasmacytic differentiation. In support of this finding, BCL-6 was shown to block plasmacytic differentiation of mouse B cells (Reljic et al., 2000). BCL-6 may also play a role in the rapid cell cycle progression of GC B cells by repressing the cdk inhibitor p27kip1 (Shaffer et al., 2000).

These findings suggest a credible mechanism by which BCL-6 causes lymphomas (Shaffer et al., 2000). BCL-6 translocations remove the normal regulatory elements in the BCL-6 promoter, thus preventing the transcriptional silencing of this gene that normally occurs during plasmacytic differentiation. A GC B cell with a BCL-6 translocation would therefore be trapped at this stage of differentiation due to repression of blimp-1 and might continue to proliferate due to repression of p27kip1. Such a B cell might accumulate further oncogenic lesions given the fact that the somatic hypermutational machinery active in GC B cells introduces double-stranded DNA breaks (Bross et al., 2000; Papavasiliou and Schatz, 2000).

Gene Expression Profiles and Lymphoid Malignancy

While the immune system is fascinating to observe as it functions normally, it is often of greater interest when

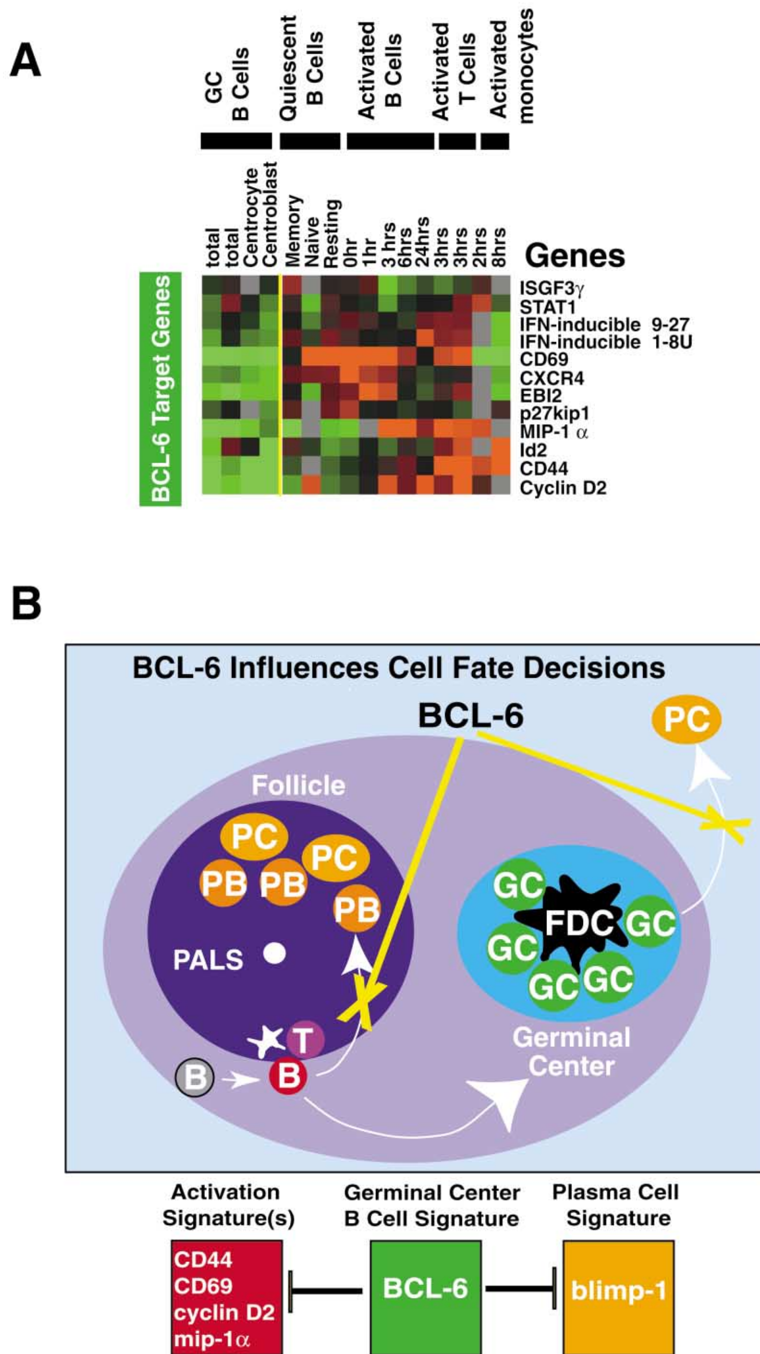

Figure 5. GC B Cells and the Role of BCL-6

(A) Targets of BCL-6 repression (Shaffer et al., 2000) are shown. Samples are the same as in Figure 4A.

(B) BCL-6 target genes and B cell fate decisions. The upper portion of the cartoon represents the possible fate choices for B cells upon encountering antigen in the periphery. B, B cell; T, T cell; PB, plasmablast; PC, plasma cell; GC, germinal center B cell; and FDC, follicular dendritic cell. The lower panel shows BCL-6 repressing genes from different signatures, in particular those involved in B cell activation and terminal differentiation.

things go awry. The most common non-Hodgkin's lymphoma, diffuse large B cell lymphoma (DLBCL), is an aggressive malignancy of mature B cells that is clinically heterogeneous, with ~40% of patients cured by current chemotherapy regimens, whereas the remainder unfortunately fail this treatment. Microarray analysis has been used to explore the molecular basis for this clinical heterogeneity (Alizadeh et al., 2000). Roughly half of the DLBCL tumor samples studied showed a prominent GC B cell gene expression signature, whereas the other

DLBCL samples instead expressed genes characteristic of mitogenically activated blood B cells. The GC B cell signature defined a group of patients with good prognosis, whereas the activated B cell signature defined a group of patients that were resistant to chemotherapy. Recent experiments have revealed that the latter group of patients had tumors that expressed genes in the NF-κB signature due to constitutive activity of the NF-κB signaling pathway in these cells (R.E. Davis et al., submitted). Interference with the NF-κB signaling pathway in cell line models of activated B-like DLBCL was lethal, suggesting that this pathway is a new molecular target for treatment of these often incurable lymphomas (R.E. Davis et al., submitted).

Concluding Remarks

The concept of a gene expression signature is somewhat fluid and operationally defined. Nonetheless, the success of these signatures in uncovering new biological principles is clear. The direct comparison of gene expression data generated on different experimental platforms and in different laboratories may be a challenging task. A more easily accomplished goal would be a gene expression signatures database that would summarize microarray experiments investigating how cells differentiate, respond to stimuli, employ transcription factors, and react to drugs. A centralized, publicly available database of gene expression signatures would make these signatures useful for the entire immunology community and might be a first step toward of genomic description of the immune system.

Acknowledgments

We apologize to those authors whose work we could not cite due to space limitations. For a comprehensive reference describing the function of individual genes in this review, we recommend OMIM at http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi. We would like to thank the members of the Staudt laboratory, past and present, and our collaborators, whose enormous efforts have made this review possible.

References

Alizadeh, A.A., and Staudt, L.M. (2000). Genomic-scale gene expression profiling of normal and malignant immune cells. Curr. Opin. Immunol. *12*, 219–225.

Alizadeh, A.A., Eisen, M., Davis, R.E., Ma, C., Sabet, H., Tran, T., Powell, J., Yang, L., Marti, G., Moore, T., et al. (1999). The lymphochip: a specialized cDNA microarray for the genomic-scale analysis of gene expression in normal and malignant lymphocytes. Cold Spring Harb. Symp. Quant. Biol. *64*, 71–78.

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). Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature *403*, 503–511.

Baldwin, A.S., Jr. (1996). The NF-kappa B and I kappa B proteins: new discoveries and insights. Annu. Rev. Immunol. *14*, 649–683.

Baldwin, A.S. (2001). Control of oncogenesis and cancer therapy resistance by the transcription factor NF-kappaB. J. Clin. Invest. *107*, 241–246.

Bassett, D.E., Jr., Eisen, M.B., and Boguski, M.S. (1999). Gene expression informatics—it's all in your mine. Nat. Genet. *21*, 51–55.

Bauch, A., Alt, F.W., Crabtree, G.R., and Snapper, S.B. (2000). The cytoskeleton in lymphocyte signaling. Adv. Immunol. *75*, 89–114.

Bross, L., Fukita, Y., McBlane, F., Demolliere, C., Rajewsky, K., and Jacobs, H. (2000). DNA double-strand breaks in immunoglobulin genes undergoing somatic hypermutation. Immunity *13*, 589–597.

Chatterjee-Kishore, M., van den Akker, F., and Stark, G.R. (2000). Association of STATs with relatives and friends. Trends Cell Biol. *10*, 106–111.

Crabtree, G.R. (1999). Generic signals and specific outcomes: signaling through Ca2+, calcineurin, and NF-AT. Cell *96*, 611–614.

de Veer, M.J., Holko, M., Frevel, M., Walker, E., Der, S., Paranjape, J.M., Silverman, R.H., and Williams, B.R. (2001). Functional classification of interferon-stimulated genes identified using microarrays. J. Leukoc. Biol. *69*, 912–920.

Dent, A.L., Shaffer, A.L., Yu, X., Allman, D., and Staudt, L.M. (1997). Control of inflammation, cytokine expression, and germinal center formation by BCL-6. Science *276*, 589–592.

Dimberg, A., Nilsson, K., and Oberg, F. (2000). Phosphorylation-deficient Stat1 inhibits retinoic acid-induced differentiation and cell cycle arrest in U-937 monoblasts. Blood *96*, 2870–2878.

Dustin, M.L., and Cooper, J.A. (2000). The immunological synapse and the actin cytoskeleton: molecular hardware for T cell signaling. Nat. Immunol. *1*, 23–29.

Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. Proc. Natl. Acad. Sci. USA *95*, 14863–14868.

Fambrough, D., McClure, K., Kazlauskas, A., and Lander, E.S. (1999). Diverse signaling pathways activated by growth factor receptors induce broadly overlapping, rather than independent, sets of genes. Cell *97*, 727–741.

Feske, S., Giltnane, J., Dolmetsch, R., Staudt, L.M., and Rao, A. (2001). Gene regulation mediated by calcium signals in T lymphocytes. Nat. Immunol. *2*, 316–324.

Foletta, V.C., Segal, D.H., and Cohen, D.R. (1998). Transcriptional regulation in the immune system: all roads lead to AP-1. J. Leukoc. Biol. *63*, 139–152.

Fukuda, T., Yoshida, T., Okada, S., Hatano, M., Miki, T., Ishibashi, K., Okabe, S., Koseki, H., Hirosawa, S., Taniguchi, M., et al. (1997). Disruption of the Bcl6 gene results in an impaired germinal center formation. J. Exp. Med. *186*, 439–448.

Ghosh, S., May, M.J., and Kopp, E.B. (1998). NF-kappa B and Rel proteins: evolutionarily conserved mediators of immune responses. Annu. Rev. Immunol. *16*, 225–260.

Glynne, R., Akkaraju, S., Healy, J.I., Rayner, J., Goodnow, C.C., and Mack, D.H. (2000). How self-tolerance and the immunosuppressive drug FK506 prevent B-cell mitogenesis. Nature *403*, 672–676.

Gugasyan, R., Grumont, R., Grossmann, M., Nakamura, Y., Pohl, T., Nesic, D., and Gerondakis, S. (2000). Rel/NF-kappaB transcription factors: key mediators of B-cell activation. Immunol. Rev. *176*, 134–140.

Guo, Q.M., Malek, R.L., Kim, S., Chiao, C., He, M., Ruffy, M., Sanka, K., Lee, N.H., Dang, C.V., and Liu, E.T. (2000). Identification of c-myc responsive genes using rat cDNA microarray. Cancer Res. *60*, 5922–5928.

Hinz, M., Loser, P., Mathas, S., Krappmann, D., Dorken, B., and Scheidereit, C. (2001). Constitutive NF-kappaB maintains high expression of a characteristic gene network, including CD40, CD86, and a set of antiapoptotic genes in Hodgkin/Reed-Sternberg cells. Blood *97*, 2798–2807.

Karin, M., and Ben-Neriah, Y. (2000). Phosphorylation meets ubiquitination: the control of NF-[kappa]B activity. Annu. Rev. Immunol. *18*, 621–663.

Lam, L.T., Pickeral, O.K., Peng, A.C., Rosenwald, A., Hurt, E.M., Giltnane, J.M., Averett, L., Zhao, H., Davis, R.E., Sathyamoorthy, M., et al. (2001). Genomic-scale measurement of mRNA turnover using the anti-cancer drug flavopiridol. Genome Biol., in press.

Lee, H.H., Dadgostar, H., Cheng, Q., Shu, J., and Cheng, G. (1999). NF-kappaB-mediated up-regulation of Bcl-x and Bfl-1/A1 is required for CD40 survival signaling in B lymphocytes. Proc. Natl. Acad. Sci. USA *96*, 9136–9141.

Ma, C., and Staudt, L.M. (2001). Molecular definition of the germinal

centre stage of B-cell differentiation. Philos. Trans. R. Soc. Lond. B. Biol. Sci. *356*, 83–89.

McHeyzer-Williams, L.J., Driver, D.J., and McHeyzer-Williams, M.G. (2001). Germinal center reaction. Curr. Opin. Hematol. *8*, 52–59.

Meraz, M.A., White, J.M., Sheehan, K.C., Bach, E.A., Rodig, S.J., Dighe, A.S., Kaplan, D.H., Riley, J.K., Greenlund, A.C., Campbell, D., et al. (1996). Targeted disruption of the Stat1 gene in mice reveals unexpected physiologic specificity in the JAK-STAT signaling pathway. Cell *84*, 431–442.

Messika, E.J., Lu, P.S., Sung, Y.J., Yao, T., Chi, J.T., Chien, Y.H., and Davis, M.M. (1998). Differential effect of B lymphocyte-induced maturation protein (Blimp-1) expression on cell fate during B cell development. J. Exp. Med. *188*, 515–525.

Nyman, T.A., Matikainen, S., Sareneva, T., Julkunen, I., and Kalkkinen, N. (2000). Proteome analysis reveals ubiquitin-conjugating enzymes to be a new family of interferon-alpha-regulated genes. Eur. J. Biochem. *267*, 4011–4019.

Pahl, H.L. (1999). Activators and target genes of Rel/NF-kappaB transcription factors. Oncogene *18*, 6853–6866.

Papavasiliou, F.N., and Schatz, D.G. (2000). Cell-cycle-regulated DNA double-stranded breaks in somatic hypermutation of immunoglobulin genes. Nature *408*, 216–221.

Penninger, J.M., and Crabtree, G.R. (1999). The actin cytoskeleton and lymphocyte activation. Cell *96*, 9–12.

Perou, C.M., Jeffrey, S.S., van de Rijn, M., Rees, C.A., Eisen, M.B., Ross, D.T., Pergamenschikov, A., Williams, C.F., Zhu, S.X., Lee, J.C., et al. (1999). Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. Proc. Natl. Acad. Sci. USA *96*, 9212–9217.

Rao, A., Luo, C., and Hogan, P.G. (1997). Transcription factors of the NFAT family: regulation and function. Annu. Rev. Immunol. *15*, 707–747.

Rayet, B., and Gelinas, C. (1999). Aberrant rel/nfkb genes and activity in human cancer. Oncogene *18*, 6938–6947.

Reimold, A.M., Iwakoshi, N.N., Manis, J., Prashanth, V., Szomolanyi-Tsuda, E., Gravallese, E.M., Friends, D., Grusby, M.J., Alt, F., and Glimcher, L.H. (2001). Plasma cell differentiation requires transcription factor XBP-1. Nature, *412*, 300–307.

Reljic, R., Wagner, S.D., Peakman, L.J., and Fearon, D.T. (2000). Suppression of signal transducer and activator of transcription 3-dependent B lymphocyte terminal differentiation by BCL-6. J. Exp. Med. *192*, 1841–1848.

Schena, M., Shalon, D., Davis, R.W., and Brown, P.O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science *270*, 467–470.

Schena, M., Shalon, D., Heller, R., Chai, A., Brown, P.O., and Davis, R.W. (1996). Parallel human genome analysis: microarray-based expression monitoring of 1000 genes. Proc. Natl. Acad. Sci. USA *93*, 10614–10619.

Schuhmacher, M., Kohlhuber, F., Holzel, M., Kaiser, C., Burtscher, H., Jarsch, M., Bornkamm, G.W., Laux, G., Polack, A., Weidle, U.H., and Eick, D. (2001). The transcriptional program of a human B cell line in response to Myc. Nucleic Acids Res. *29*, 397–406.

Seebach, F.A., Welte, T., Fu, X.Y., Block, L.H., and Kashgarian, M. (2001). Differential activation of the STAT pathway by angiotensin II via angiotensin type 1 and type 2 receptors in cultured human fetal mesangial cells. Exp. Mol. Pathol. *70*, 265–273.

Shaffer, A.L., Yu, X., He, Y., Boldrick, J., Chan, E., and Staudt, L. (2000). BCL-6 represses genes that function in lymphocyte differentiation, inflammation and cell cycle control. Immunity *13*, 191–212.

Shaulian, E., and Karin, M. (2001). AP-1 in cell proliferation and survival. Oncogene *20*, 2390–2400.

Sherlock, G. (2000). Analysis of large-scale gene expression data. Curr. Opin. Immunol. *12*, 201–205.

Siebenlist, U., Franzoso, G., and Brown, K. (1994). Structure, regulation and function of NF-kappa B. Annu. Rev. Cell Biol. *10*, 405–455.

Snapper, S.B., and Rosen, F.S. (1999). The Wiskott-Aldrich syndrome protein (WASP): roles in signaling and cytoskeletal organization. Annu. Rev. Immunol. *17*, 905–929.

Stark, G.R. (1997). Genetic analysis of interferon and other mammalian signaling pathways. Harvey Lect. *93*, 1–16.

Stark, G.R., Kerr, I.M., Williams, B.R., Silverman, R.H., and Schreiber, R.D. (1998). How cells respond to interferons. Annu. Rev. Biochem. *67*, 227–264.

Staudt, L.M. (2001). Gene expression physiology and pathophysiology of the immune system. Trends Immunol. *22*, 35–40.

Staudt, L.M., and Brown, P.O. (2000). Genomic views of the immune system. Annu. Rev. Immunol. *18*, 829–859.

Taniguchi, M., Miura, K., Iwao, H., and Yamanaka, S. (2001). Quantitative assessment of DNA microarrays—comparison with Northern blot analyses. Genomics *71*, 34–39.

Turner, C.A., Mack, D.H., and Davis, M.M. (1994). Blimp-1, a novel zinc finger-containing protein that can drive the maturation of B lymphocytes into immunoglobin-secreting cells. Cell *77*, 297–306.

Wagner, E.F. (2001). AP-1—introductory remarks. Oncogene *20*, 2334–2335.

Wong, M., Uddin, S., Majchrzak, B., Huynh, T., Proudfoot, A.E., Platanias, L.C., and Fish, E.N. (2001). Rantes activates Jak2 and Jak3 to regulate engagement of multiple signaling pathways in T cells. J. Biol. Chem. *276*, 11427–11431.

Ye, B.H., Cattoretti, G., Shen, Q., Zhang, J., Hawe, N., de Waard, R., Leung, C., Nouri-Shirazi, M., Orazi, A., Chaganti, R.S., et al. (1997). The BCL-6 proto-oncogene controls germinal-centre formation and Th2- type inflammation. Nat. Genet. *16*, 161–170.