Changes in Gene Expression Profiles in Developing B Cells of Murine Bone Marrow

Reinhard Hoffmann,[1,3,4] Thomas Seidl,[1] Martin Neeb,[2] Antonius Rolink,[1] and Fritz Melchers[1,4]

[1]*Basel Institute for Immunology, 4005 Basel, Switzerland;* [2]*PRBT-S, F. Hoffmann-La Roche Ltd, Basel, Switzerland*

Gene expression profiles of five consecutive stages of mouse B cell development were generated with high-density oligonucleotide arrays from as few as 2×10^4 ex vivo isolated and flow-cytometrically purified cells. Between 2.8% and 6.8% of all genes change on differentiation from one cellular stage to the next by at least twofold. The entire pathway involves differential expression of 10.7% of all genes. Previously known expression patterns of 15 genes (like surrogate light chain, RAG-1/2, MHC class II, mel-14 antigen) are confirmed. The gene expression patterns of the proliferating pre-BI and large pre-BII cells on the one hand, and the resting immature and mature B cells on the other hand, are most similar to each other. Small pre-BII cells display a pattern that is transitional between these two groups. Most of the genes expressed in early precursors are involved in general processes, like protein folding or cell cycle regulation, whereas more mature precursors express genes involved in more specific molecular programs (cell surface receptors, secreted factors, and adhesion molecules, among others). Between 19 and 139 genes share a given expression pattern. Combining knowledge about gene function and expression pattern allows identification of novel candidate genes potentially involved in self-maintenance of pre-BI cells, allelic exclusion and pre-B cell receptor signaling in large pre BII cells, cell-cycle arrest of small pre-BII cells, propensity toward apoptosis or anergization in immature B cells, propensity toward cell division and activation in mature B cells, and stage-specific interactions with stromal cells in the bone marrow.

[The sequence data described in this paper have been submitted to the Gene Expression Omnibus (GEO) at the National Center for Biotechnology Information (NCBI) under accession number GSE13. Online supplementary material available at www.genome.org.]

Mouse B-lymphocytes develop from progenitors and precursors in bone marrow in a sequence that can be ordered by changing status of their immunoglobulin gene rearrangements (Tonegawa 1983; ten Boekel et al. 1995). Cell cycle status and the differential surface expression of c-kit, CD25, IgM, and IgD on B220+ cells distinguish five consecutive stages of development (Rolink et al. 1994). Therefore, fluorescence-activated cell sorting (FACS) can be used to purify five cell populations that follow each other in progressive differentiation: (1) D_H-J_H rearranged, c-kit+CD25-cycling pre-BI cells; (2) $V_HD_HJ_H$-rearranged, c-kit-CD25+-cycling large pre-BII cells; (3) $V_HD_HJ_H$- and V_LJ_L-rearranged c-kit-CD25+ resting small pre-BII cells; (4) sIgM+ resting immature; and (5) sIgM+IgD+ resting mature B-cells (Melchers and Rolink 1999).

In this developmental sequence of cells, pre-BI cells also express surrogate light chain encoded by VpreB and λ5 genes (Karasuyama et al. 1994), and the rearrangement machinery encoded by the RAG-1, RAG-2 (Grawunder et al. 1995), and TdT (Melchers and Rolink 1999) genes. As soon as one allele has been rearranged productively, somatic recombination is stopped, preventing additional rearrangements on the second allele. This process is termed allelic exclusion (Melchers and Rolink 1999). The μ heavy chain derived from a productively $V_HD_HJ_H$-rearranged IgH chain locus has to pair with the surrogate light chain to form a pre-BCR on the surface of large pre-BII cell (ten Boekel et al. 1997). Expression of the surrogate light chain and of the rearrangement machinery is then turned off (Grawunder et al. 1995). The pre-BCR induces two to five divisions of large pre-BII cells (Rolink et al. 2000). As the pre-BCR is diluted by these divisions, the cells come to rest as small pre-BII cells, the expression of the rearrangement machinery is turned on again and V_L segments are rearranged to J_L segments on the κL and λL chain gene loci. As soon as an L chain has paired with the pre-existing μ heavy chain, IgM can be deposited on the surface to give the cell the status of an immature B cell. Autoantigens select the emerging repertoire of immature B cells negatively to delete high-affinity autoreactive cells and may also select positively to differentiate low-affinity autoreactive cells into the B1 cell compartment (Nemazee et al. 2000). Immature B cells keep the rearrangement machinery up-regulated to allow for secondary rearrangements at the IgL chain gene loci with which they can change, thereby edit the specificity of autoreactive cells (Yu et al. 1999). During this differentiation program in the bone marrow, B-cell precursors interact with different cell types (osteoblasts, osteoclasts, reticular stromal cells, dendritic cells, and others) in a possibly stage-specific manner (Melchers and Rolink 1999). Immature B cells, finally, leave the bone marrow for the spleen where they mature to sIgM+ sIgD+ B cells.

These cellular stages of B-cell differentiation have been described in detail. Very little, however, is known about the molecular mechanisms controlling the various functions of B-lineage cells in this development. The self-renewing capacity of pre-BI cells with the ability to long-term proliferate on

[3]Present address: Max-von-Pettenhofer-Institut, Pettenhofer Str. 9A, 80336 Munich, Germany.
[4]Corresponding authors.
E-MAIL r_hoffmann@m3401.mpk.med.uni-muenchen.de; FAX 49-89-5160-5202.
F. Melchers FAX 41-61-605-1300.
Article and publication are at http://www.genome.org/cgi/doi/10.1101/gr.201501.

12:98–111 ©2002 by Cold Spring Harbor Laboratory Press ISSN 1088-9051/01 $5.00; www.genome.org

stromal cells with IL-7, the signal transduction from the pre-BCR for proliferative expansion and possibly for allelic exclusion (both taking place in large pre-BII cells), cell-cycle exit in small pre-BII cells, and susceptibility to autoantigen signaling in immature B cells all are processes lacking a precise molecular description to date. Some of the genes involved in such mechanisms can be expected, however, to show specific changes in expression during B-cell development. Such genes appear to be good candidate genes that could control this development.

High-density oligonucleotide arrays allow the analysis of a large number of genes expressed as mRNA in parallel (Lockhart et al. 1996). Using higher-level data analysis algorithms, expression profiles of different cellular stages can be compared, and genes potentially involved in biological processes of interest can be identified by their expression patterns (Eisen et al. 1998; Tamayo et al. 1999). Previously, a large number of cells has been required to perform such analysis, essentially excluding the use of this technology for cells that are too infrequent in vivo, as in this case for the development of B cells in mouse bone marrow. Here, we describe the modification of an in-vitro-transcription-based RNA amplification procedure (Eberwine et al. 1992; Luo et al. 1999) that now allows us to perform array-based expression profiling from as few as 5×10^4 cells or less, lowering the amount of necessary input material by a factor of 1000. This has enabled us, for the first time, to analyze gene expression patterns of five consecutive stages of a mammalian developmental program using ex vivo isolated cells.

RESULTS

To study gene expression profiles in murine B-cell development, total cellular RNA was extracted from 5×10^4 to 1.5×10^5 cells of each of five consecutive B-lymphocyte lineage subpopulations (Fig. 1). mRNA was amplified by two subsequent cycles of cDNA synthesis and in vitro transcription (Eberwine et al. 1992; Luo et al. 1999). The RNA samples were hybridized to high-density oligonucleotide arrays interrogating 13,026 transcripts [5231 known genes and 7795 expressed sequence tags (ESTs)]. Five independent replicate experiments were performed, and differential expression was assessed on the basis of statistical algorithms (see Methods). The raw, non-normalized gene expression values for each of the individual experiments are presented in supplementary Table 1 (available as an on-line supplement at http://www.genome.org).

We first determined how many genes change on differentiation from one cellular stage to the next. Between pre-BI and large pre-BII, large pre-BII and small pre-BII, small pre-BII and immature B cells, and immature and mature B-cells, a total of 1028, 1604, 996, and 873 genes have a *t*-test *P* value of 0.02 or less. Of these genes, 488, 885, 462, and 362 genes change at least twofold; 133, 321, 154, and 121 genes change at least fivefold; and 39, 110, 58, and 43 genes change at least 10-fold, respectively. Therefore, between 2.8% and 6.8% of all genes present on the arrays change on differentiation from one stage to the next. A complete list of genes changing at least fivefold together with the expression level values and associated *t*-test *P* values is available as supplementary Table 2 and supplementary Table 3 contains the information for all probe sets represented on the arrays (tables available as an on-line supplements at http://www.genome.org).

Because only ~45% of the genes reaching statistical significance at the 98% confidence level also fulfill the fold change and absolute difference criteria, the type I error rate (number of genes falsely detected to be differentially expressed) can be expected to be well below the 2% that the confidence criterion alone allows.

Employing analysis of variance (ANOVA), 1734 genes reach statistical significance at the 98% level. Of these, 1406 genes were identified as differentially expressed by a factor of two or more during mouse B-cell development. The entire information about the ANOVA analysis is presented in supplementary Table 4 (available as an on-line supplement at http://www.genome.org). A hierarchical cluster analysis (Eisen et al. 1998) performed on this subset of genes is shown in Figure 2. Except for large pre-BII cells, all B-cell precursor stages up-regulate a specific set of genes. This is indicated on the left side of the plot. Genes up-regulated in large pre-BII cells are already expressed in pre-BI cells. Those genes are often cell-cycle related (cyclins, brca-1, ki-67, PCNA), involved in DNA replication (Topoisomerase II, helicase, DNA polymerase, ligase, and primase), or cytoskeleton components (tubulins)(see the bottom two clusters in Fig. 2, right). This is consistent with the fact that pre-BI and large pre-BII cells are cycling while the subsequent three cellular stages are resting. In Figure 2, functionally related genes tend to show similar expression patterns and come to lie in close proximity to each other. Examples include parts of immunoglobulin genes (expression in immature and mature B cells, top cluster in Fig. 2, right) and MHC class II sequences (expressed specifically in mature B cells, second from top in Fig. 2, right).

Comparing the gene expression patterns across the five cellular stages, cycling pre-BI and large pre-BII on the one hand, and resting immature and mature B cells on the other hand, appear to be most similar to each other (Fig. 2). Small pre-BII cells appear to be intermediate between these two groups, expressing some genes expressed in pre-BI and large pre-BII cells (bottom half in Fig. 2) and some others expressed in immature and mature B cells (top half in Fig. 2). Therefore, a major change in gene expression takes place after the large pre-BII cell stage, probably because proliferation ceases and new differentiation programs are initiated in resting cells.

Figure 3 shows a tree diagram representing correlations between gene expression profiles from individual replicate experiments of the five developmental stages examined. All of the stagewise replicate experiments form tight clusters, indicating small intra-stage variability as compared with variability between stages. The tree reflects the relationships between stage-specific gene expression profiles as mentioned above, as the pre-BI and large pre-BII cells on the one side and the immature and mature B cells on the other cluster together on distal branches of the tree, respectively.

Next, the 1406 differentially expressed transcripts were grouped according to their expression pattern by self-organizing maps (Tamayo et al. 1999). The entire data set is available as supplementary Table 5 (available as an on-line supplement at http://www.genome.org). Twenty different, though sometimes closely related, patterns could be identified (Fig. 4). We compared these patterns with published data for genes whose expression patterns during mouse B-cell development have been determined earlier (Grawunder et al. 1995; Melchers and Rolink 1999). Fifteen such known genes [VpreB, λ5, terminal deoxynucleotidyltransferase, c-kit (all c1), RAG-1/2 (c10 and c8 in Fig. 4, respectively), μ heavy chain constant region (c11), CD40 (c14), b29 (Igβ), mb-1 (Igα), κ light chain constant region (all c15), CD19 (c16), Igδ

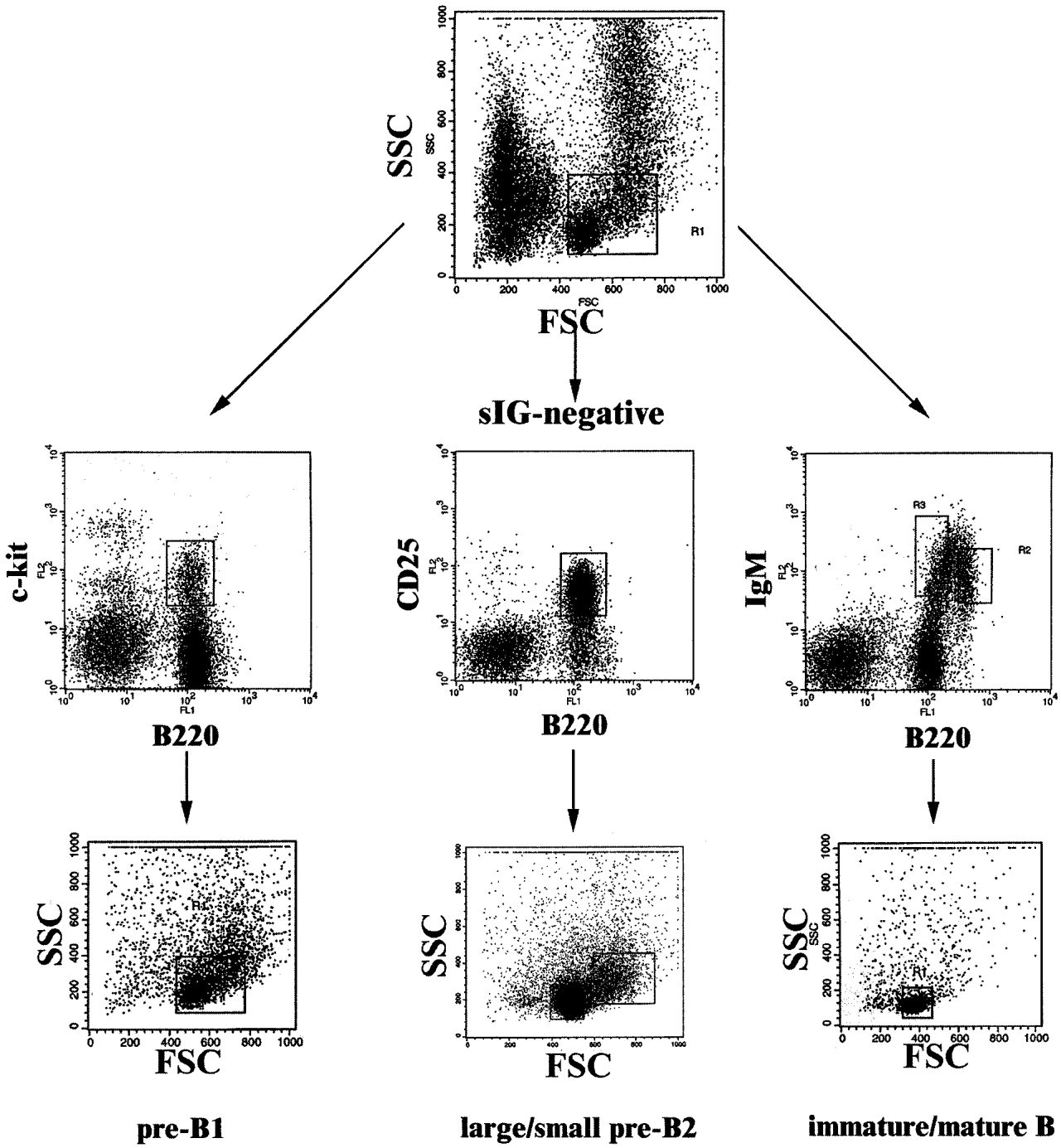

Figure 1 Flow cytometric purification of mouse B-cell precursors in three steps. Step 1 (*top* row): Enrichment of lymphoid cells. Shown is a forward–sideward scatter plot of total femoral bone marrow cells. The boxed area indicates the lymphoid gate. Second step (*middle* row): Enrichment of c-kit+B220+ pre-BI cells, CD25 +B220+ pre-BII cells, and sIgM+B220+ immature/mature B cells. Surface marker staining of three aliquots of lymphoid-gated bone marrow cells is shown. Pre-BI and immature/mature B cells were sorted as indicated by the gates displayed as boxes in the *top* and *middle* rows. Third step (*bottom* row): Separation of pre-BII cells according to cell size. Shown are forward–sideward scatter plots of cells gated as in the *middle* row. Pre-BII cells were separated according to cell size into the large pre-BII cells (*right* box) and small pre-BII cells (*left* box). Note that pre-BI cells consist of a small and a large subpopulation, which were not separated. Immature and mature B cells were sorted according to the gates shown in the *middle* row, *right* panel. As both populations consist of homogeneous small cells, only one back-gated forward–sideward scatter plot is shown in the *right* panel of the *bottom* row.

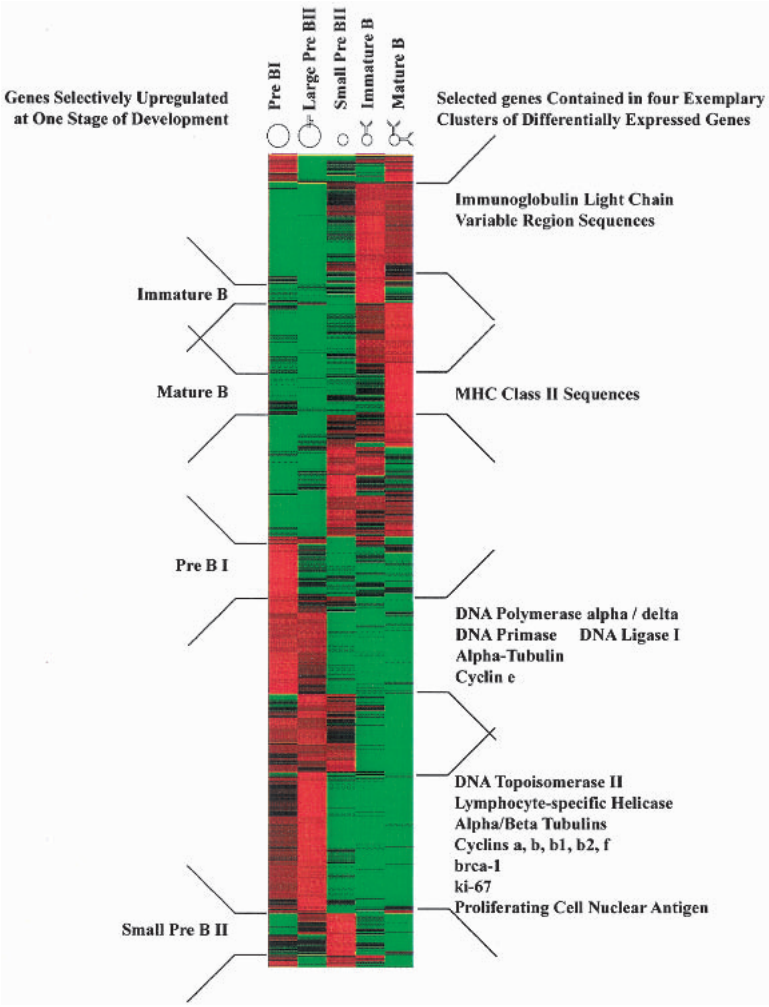

Figure 2 Hierarchical cluster analysis of 1406 differentially expressed genes as detected by ANOVA with 98% confidence (measured by Kruskal-Wallis statistics) with a change of at least twofold and a difference of at least 100 average difference units (for details see Methods). Genes are organized in rows, whereas the five columns represent the five developmental stages from the most immature, pre-BI compartment on the *left* to the most mature compartment, mature B cells on the *right*. The expression level of every gene for every developmental stage has been normalized. Green denotes a normalized expression level below, black near to, and red above the mean. Developmental stage-specific clusters are indicated on the *left* and clusters of selected functional classes of genes are indicated on the *right* side of the plot.

constant region, MHC class II sequences, mel-14 antigen (all c17)] were examined, and all the shown array-based expression patterns are in concordance with the published data (Table 1).

A total of 656 of the differentially expressed transcripts were genes; the remainders were ESTs. After review of published data, a definitive or putative function could be assigned to 625 of the 656 known genes. The 31 remaining genes were often cloned as growth factor inducible genes, and no follow-up studies have been performed.

A synopsis of functional classifications of genes correlated with their cluster membership is shown in Table 2. It shows that early precursors express genes involved in V(D)J-recombination, protein folding and degradation, general and redox metabolism, DNA replication, chromatin and cytoskeleton structure, RNA processing, cell cycle regulation, and molecular transport (Table 2; clusters 1–4 in Fig. 4). In contrast, more mature precursors express genes involved in intercellular communication (cell surface receptors and secreted factors), transcriptional regulation, immunoglobulins, complement components, cytoskeleton modification, intercellular contact and adhesion, MHC components, and apoptosis (Table 2; clusters 14–17 in Fig. 4). Therefore, the majority of genes up-regulated by more immature precursors are involved in general metabolic processes, whereas more mature precursors up-regulate genes involved in specific molecular programs.

The correlation of gene expression pattern with biochemical function of the respective protein product also allowed us to identify candidate genes that could be involved in functional properties of B-cells above.

Pre-BI cells, for example, have self-renewing capacity and are capable of long-term proliferation in the presence of bone marrow stromal cells and IL-7 or IL-3 (Winkler et al. 1995). These capacities are lost on differentiation to the pre-BII cell stage. Therefore, genes involved in these processes are probably specifically expressed in pre-BI cells, that is, in clusters 1 or 2. As can be seen in Table 2, these clusters contain a high proportion of genes involved in general metabolism and DNA replication, but four signaling molecules (p116RIP, T-cell specific clone U2, TIAM-1, and SOCS-2), five cell-surface receptors (thromboxane A2 receptor, FcεRI, endoglin, common beta chain of IL-3/IL-5/GM-CSF receptors, and c-kit) and three genes involved in cell–cell contact or adhesion (gp-70, alpha-catenin, and lectin L14) (Table 3A) are candidate genes potentially involved in the pre-BI specific processes.

In large pre-BII cells, V(D)J-recombination on the heavy chain locus is inhibited after one allele has been productively rearranged, a process termed allelic exclusion (Melchers and Rolink 1999). After surface expression of a pre-BCR consisting of the surrogate light chain components λ5 and VpreB and the newly rearranged heavy chains, these cells undergo a proliferative burst (Rolink et al. 2000). Allelic exclusion and proliferative burst are likely to require specific signals. Genes involved in these processes are likely to be found in clusters 3, 4, 5, or 9. Cluster 3 contains many genes involved in DNA replication, and cluster 4 contains many genes involved in cell-cycle regulation. In addition, however, these clusters also contain many genes encoding signaling molecules, with a total of 23 signaling components specifically up-regulated in large pre-BII cells (Table 3B and supplementary Table 5, available as an on-line supplement at http://www.genome.org). A number of these (CDC25, ect2, STK-1, sak-a, nek-2, and ran) are known to transmit signals that are important for cell-cycle progression (Coutavas et al. 1994; Fode et al. 1994; Wickramasinghe et al. 1995; Niwa et al. 1996; Tanaka et al. 1997; Tatsumoto et al. 1999). Two of the signaling components up-regulated in pre-BII cells (calmodulin and the calmodulin syn-

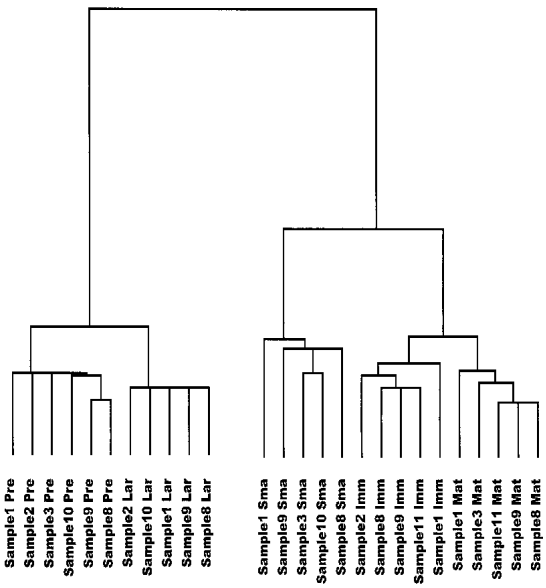

Figure 3 Tree diagram displaying correlations between gene expression patterns from individual replicate experiments of the five stages of B-cell differentiation, based on the set of 1406 differentially expressed genes. Shorter branches indicate more similar gene expression profiles. (Sample) Number of the sample (consisting of cells pooled from four mice) from which the cell populations have been purified by FACS, numbered consecutively from the beginning of the study; (Pre) pre-BI cells; (Lar) large pre BII cells; (Sma) small pre-BII cells; (Imm) immature B cells; (Mat) mature B cells.

thesis cDNA) are involved in calcium signaling (Chin and Means 2000), whereas the rest of the signaling molecules (PP2Abalpha3, citron, fug1, pLK, ayk1, pMELK, MKP-1, and STAM) remain rather poorly characterized. Therefore, signal transduction from the preBCR might involve some of these genes.

For allelic exclusion, one possible part of the total mechanism is structural reorganization of chromatin that would make the immunoglobulin heavy chain loci inaccessible for further recombination. In fact, the clusters 3, 4, and 5 contain seven, six, and three genes, respectively, involved in chromatin structure (Table 3B). Five of these genes code for histones, two more for chromosomal structural proteins; it is likely that the expression pattern of these genes simply reflects the cycling status of large pre-BII cells. Four genes, however, are involved in transcriptional regulation via structural changes: the modifier-1 protein is involved in heterochromatin formation (Singh et al. 1991), whereas SRG-3 is associated with the SWI–SNF complex to modify locus accessibility (Jeon et al. 1997). The retinoblastoma-binding proteins RbAp46 and RbAp48 are involved in histone acetylation-dependent transcriptional regulation (Zhang et al. 2000).

Small pre-BII cells are characterized by cell-cycle arrest and induction of immunoglobulin light chain rearrangements. For cell-cycle arrest, at least two mechanisms can be deduced from the chip data (Table 3C). The major cell-cycle regulator, PCNA, is highly expressed in pre-BI and large pre-BII cells, decreasing in more mature cell stages (cluster 3). MyD118, known to inhibit cell cycle progression by heterodimerizing with PCNA (Vairapandi et al. 2000), shows a peak of expression in small pre-BII cells (cluster 11). A second PCNA-interacting partner with similar function (EAT/MCL-1) (Fujise et al. 2000) is up-regulated from immature B cells onward (cluster 14). Together, these two genes might be involved in sustained cell-cycle arrest after exit from the large pre-BII cell stage. A second cell cycle regulatory pathway involves chk-1 (clusters 2, 3), phosphorylating cdc25 (clusters 4, 5) in response to DNA damage signals (Sanchez et al. 1997), possibly generated during aberrant heavy chain V(D)J-recombination. This renders cdc25 inactive and generates a binding site for 14-3-3 (cluster 7), maintaining cdc25 in an inactive state and effectively blocking cell-cycle progression (Fu et al. 2000). The clusters 10 and 11, however, containing genes with peak expression in small pre-BII cells, contain four more cell-cycle repressing genes (gadd45, gas7, c-jun, and tis-21).

The pool of immature B cells is exposed to selective pressures by autoantigens. These undergoing negative selection are likely to be on their way toward apoptotic cell death or anergization, and are prevented of entering the cell cycle on BCR crosslinking. Other immature B cells are expected to be in the process of positive selection, possibly also initiated by autoantigens. BCR crosslinking should signal cell survival and proliferation in these cells (Melchers and Rolink 1999). Genes involved here are likely to be found most prominently in cluster 13, but also in cluster 14 (Table 3D). Cluster 13 contains CD72, a negative regulator for B-cell responsiveness (Parnes and Pan 2000); c-fes, a non-receptor tyrosine kinase involved in negative regulation of cytokine-induced activation in macrophages (Hackenmiller et al. 2000); and PAC-1, a tyrosine phosphatase that specifically dephosphorylates the MAP-kinases erk-1 and erk-2 (Gerondakis et al. 1994). Cluster 14 includes genes for the ITIM-bearing inhibitory receptor PIRB1 (Maeda et al. 1998); the inhibitory BCR coreceptor FcγRIIB; calcineurin, a Ca++-dependent protein phosphatase (Rusnak and Mertz 2000); and two genes involved in inhibition of cAMP-derived signals, the cyclic nucleotide phosphodiesterase PDE7A2 and the inhibitor of cAMP-dependent protein kinase. Susceptibility to apoptotic stimuli, most prominently differing between immature and mature B cells, appears to be guided by the differential expression of survival-mediating bcl-2 homologs. Therefore, although bcl-x peaks in small Pre-BII cells, PS-2 and the inhibitor of apoptosis-1 are expressed from small pre-BII cells onward (cluster15), and EAT/MCL-1 is up-regulated in immature B cells with some sustained expression in mature B cells (cluster 14), bcl-2 beta is specifically upregulated in mature B cells (cluster 17). This reconfirms that bcl-2 might be the major regulator of the difference in response to apoptotic stimuli between immature and mature B cells (Lang et al. 1997).

In contrast to immature B cells, mature B cells proliferate in response to stimulation by foreign antigens (Melchers and Rolink 1999). Although this might in part be attributable to the down-regulation of the cell-cycle inhibitors (mentioned above), as well as to the expression of bcl-2 (preventing apoptotic cell death), genes contained in cluster 17 provide additional candidate molecules potentially involved in keeping mature B cells in an activation-prone status (Table 3D). Cluster 17 contains 10 signaling molecules, for example annexin V (a collagen-regulated calcium channel; von der Mark and Mollenhauer 1997), type 1 ryanodine receptor (an intracellular calcium release channel; Giannini et al. 1995), the mitogen activated protein kinase kinase kinase 8 (activating NFκB; Lin et al. 1999), TANK (another NFκB activator downstream of TRAF in both CD40 and TNF-receptor II pathways; Cheng and Baltimore 1996), the cytokine-induced tyrosine kinase jak-2, and the cAMP synthesizing adenylyl-cyclase type VII. Also in

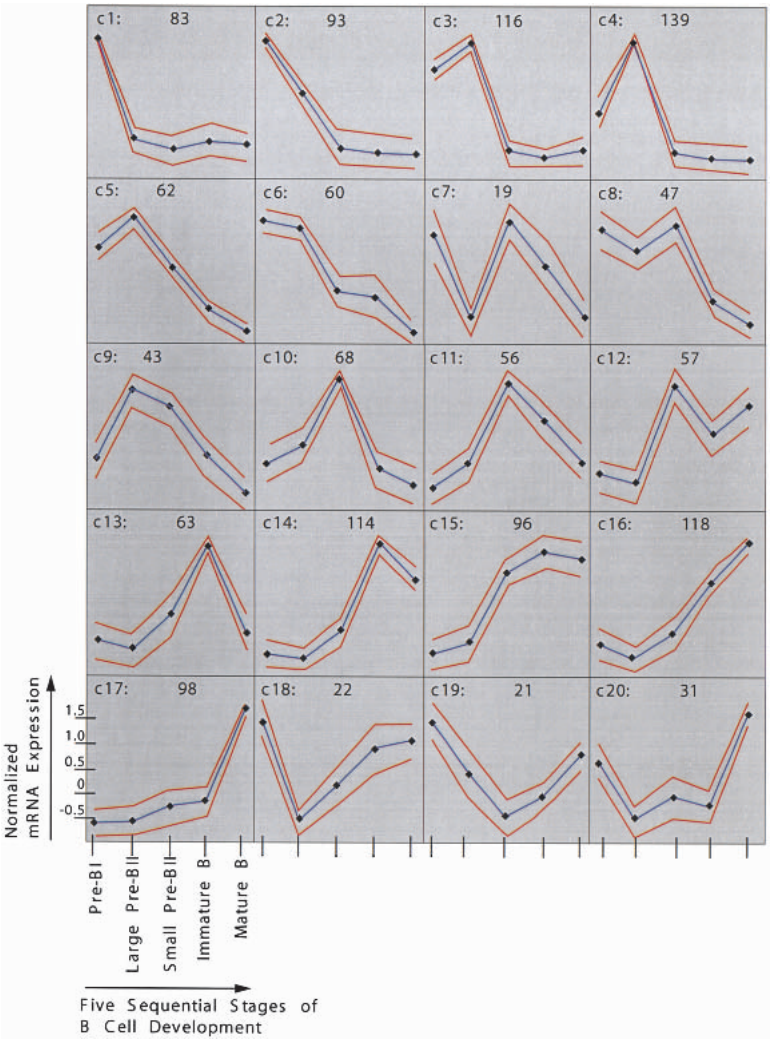

Figure 4 Gene expression patterns identified by self-organizing maps. The cluster numbers are indicated on the *top left* of each cluster diagram, and the number of genes plus ESTs belonging to every cluster is indicated on the *top right*. Expression levels are shown on y-axis and developmental stages on x-axis. Dots indicate developmental stages, with the most immature, pre-BI cells on the *left* and the most mature, that is, mature B cells, on the *right*. Expression level of each gene was normalized to have mean = 0 and SD = 1 across developmental stages. Blue and red lines indicate expression level means and standard deviations, respectively.

this cluster are the surface receptors for Interferon-gamma, the chemokine receptors CCR7 and CXCR5, and the T-cell activating protein TAP. The latter has been shown previously to be decreased in anergic B cells (Glynne et al. 2000). Mature B cells also express immunosuppressive receptors, however, like the IL-10 receptor (cluster 17), and immature B cells also express activating receptors like PIRA6 (cluster 14). Therefore, these signal-transducing molecules might be wired differently in these functionally heterogeneous B cell populations to signal either for activation or for silencing.

During differentiation in the bone marrow, B-cell precursors interact with various cell types in a stage-specific manner. Immature B cells are ready to leave the bone marrow to migrate to the spleen, subsequently maturing into mature, recirculating B cells (Melchers and Rolink 1999). These processes are likely to be guided by the differential expression of adhesion molecules. Table 3E lists genes expressed differentially throughout the five stages of B cell development that are involved in cell-to-cell contact and adhesion. Cluster 8 contains adhesion molecules that might be involved in retaining the B-cell precursors in the bone marrow (DGCR-6 and P-selectin glycoprotein ligand 1). The different intramedullary cellular interactions of B-cell precursors might be governed by molecules contained in clusters 1 and 2 (pre-BI-specific: Ig-related glycoprotein-70, alpha-catenin, and L14 lectin), cluster 5 (up-regulated in large pre-BII cells: fertilization antigen-1), and cluster 10 (up-regulated in small pre-BII cells: reelin). Cluster 12 and 15 contain adhesion molecules (galectin-3, Stra1/ ephrin B1, ICAM-2) expressed in small pre-BII, immature and mature B cells that are about to, or have already left the bone marrow. In addition, some other adhesion molecules appear to be expressed more specifically in only one of these three cellular stages. In immature B cells, these are CD37 and ICAM-1 (clusters 13 and 14). Mature B cells specifically upregulate C33/R2/IA4, CD22, LFA-1, VCAM-1, novel stromal cell protein, P-selectin/MEL-14 antigen, and integrin beta-7 subunit (clusters 16 and 17). However, there are also a number of cell adhesion molecules that are expressed both in pre-BI as well as in immature and mature B cells (clusters 18 and 20: Ly-6C.2, clusterin, Cd73, sialic acid O-acetylesterase). It remains to be determined which adhesive properties should be preserved for different microenvironments of B lymphocyte development.

DISCUSSION

This study describes gene expression patterns of five consecutive cellular stages in mouse B-cell development. A large number of differentially expressed genes is identified, and a combination of previous knowledge about B-cell development, gene expression pattern analysis, and functional annotation allows us to identify genes potentially involved in specific processes taking place in mouse B-cell precursors. These candidate genes can be subjected to more detailed functional analysis by heterologous expression, by functional inactivation employing antisense technology, or by gene inactivation via targeted deletion.

Two factors underline the validity of the gene expression data set. First, genes whose expression patterns have been identified earlier are detected in our analysis in a manner concordant with previous results (Rolink et al. 1994; Grawunder et al. 1995; Melchers and Rolink 1999). Second, analysis of co-expressed genes groups together genes with similar function in a manner consistent with what is known about these cells; genes involved in cell-cycle regulation, for example, although never formally investigated in the setting of B-cell development, are highly expressed in the cycling pre-BI and large pre-BII cells only. A number of aspects, however, require special attention.

First, before the entire murine genome sequence is avail-

Table 1. Array-Based Gene Expression Values of Genes with Pre-Established Expression Patterns During Mouse B Cell Development

Affy_id	Descr	p(H)	AD_PRE	SD_PRE	AD_LAR	SD_LAR	AD_SMA	SD_SMA	AD_IMM	SD_IMM	AD_MAT	SD_MAT	Clus
x05556_s_at	mouse pre-b lymphocyte v(preb)1 mma (lambda5 locus).	0.00047	4391	572	740	299	586	214	228	215	231	137	1
Msa.2379.0_at	murine ig-related lambda(5) gene (exon 2) transcribed selectively in pre-b-lymphocytes	0.00016	3443	2152	387	317	57	145	−146	107	−88	58	1
X04123_s_at	mouse mma for terminal deoxynucleotidyltransferase (tdt).	0.00071	794	276	107	31	66	40	70	61	3	44	1
X68670_s_at	*MUS. musculus* mma for terminal deoxynucleotidyltransferase.	0.00073	7557	548	252	151	88	112	3	34	13	56	1
y00864_s_at	mouse c-kit mma.	0.00658	262	82	75	111	75	75	23	69	2	71	1
M64796_s_at	*MUS musculus* rag-2 protein (rag-2) mma, complete cds.	0.00195	539	407	241	391	571	147	−140	183	−153	116	8
M29475_s_at	mouse recombination activating protein (rag-1) mma, complete cds.	0.00004	2156	818	1145	535	4909	1888	125	71	−83	114	10
X03690_s_at	*MUS musculus* mRNA for Ig heavy chain constant region mu(b) allε	0.00156	5817	1353	7963	1576	10997	2298	10876	2696	5495	2063	11
m94351_at	*Mus musculus* immunoglobulin lambda chain (igl) mma, complete cds.	0.00004	884	245	324	58	5148	1514	7201	821	1846	586	13
m83312_s_at	mouse cd40 mma, complete cds.	0.00059	−90	101	−140	161	−78	48	370	84	302	119	14
J03857_s_at	mouse b-cell-specific glycoprotein (b29) mma, complete cds.	0.00025	4423	1889	5847	1024	12045	1913	13044	2474	10699	884	15
x13450_s_at	murine b lymphocyte lineage restricted mb-1 mma.	0.00174	1567	403	1287	625	3716	1726	4202	1948	3159	1209	15
V00802_f_at	mouse gene fragment for kappa-immunoglobulin (constant region) (from cell line mopc21).	0.00018	987	407	4366	2050	15428	2065	19794	7964	20198	2714	15
m62553_s_at	mouse cd19 gene, exons 6–15.	0.00677	639	466	673	241	1025	636	1301	511	1583	222	16
Msa.2376.0_at	mouse gene fragment for delta-immunoglobulin encoding one of two alternate c-termini (exon 8?)	0.01846	44	54	61	44	92	69	87	90	250	80	17
Msa.803.0_s_at	part of messenger ma for mouse delta-immunoglobulin (codes for part of exon 8–one of two alternate c-termini)	0.0001	416	190	359	51	865	139	1363	500	5400	1919	17
V00790_at	part of messenger ma for mouse delta-immunoglobulin (codes for part of exon 8–one of two alternate c-termini).	0.00142	31	54	90	110	162	122	188	88	846	165	17
m25324_s_at	Mouse peripheral lymph node-specific homing receptor (MEL-14 antigen) mRNA, complete cds.	0.00017	1637	379	739	156	981	583	417	117	7757	1520	17

(Affy_ID) Unique Affymetrix probe set identifier; (Descr) gene description; p(H) *P*-value corresponding to the Kruskal-Wallis *H* value; (AD_Pre/Lar/Sma/Imm/Mat) mean Average Difference value of the five independent replicate experiments of Pre-BI, Large Pre-BII, Small Pre-BII, immature, and Mature B cells, respectively; (SD_) standard deviation among the five independent replicate experiments for the indicated population; (Clus) Cluster number from Figure 4.

Table 2. Numbers of Genes with Known Functions Differentially Expressed in the Five Developmental Stages of B Cells with the Same Patterns of Changes

Cluster	VDJ-Recombination	Unknown	Chaperones	Protein Degradation	General Metabolism	Redox Metabolism	DNA Replication	Chromatine Structure	Cytoskeleton Components	RNA Processing	Cell Cycle Regulation	Molecular Transport	Signalling	Cytochromes	Extracellular Matrix	Membrane Behaviour	Ion Transport	Intercellular Communication	-Receptors	-Secreted Factors	Transcriptional Regulation	IG and related	Complement	Cytoskeleton Modification	Cell-Cell Contact/Adhesion	MHC	Apoptosis	
1	2	5	3	4	7	2	5			1	1	1	3					7	5	2	4	3			1			49
2		4	3	1	5	3	8	1		1	1	1	2				1	1		1	2			1	2			37
3					4		16	7	4	5	4	2	5					1		1	2			1				51
4		1	1		1	1	2	6	3	1	11	4	12			1					5			2				51
5				3	2	1	3	3		4			5	3	1	1		1	1		4				1			32
6				1	5	2	2			1	3	1	1	2							1							19
7		1		3	1				1			1	4	1							1							13
8	1	1		3	2	2	1	2		1	1	2	1	1			1				2				2	1		24
9		1		1	2		1		1		2		1			2	1	1	1		3			1		1		18
10	1	2	1		2	2	2			1	3	1	6				3	4	3	1	3			1	1		1	34
11		1		1	1	1	1		1		2	1	6								3	3				1		22
12		3			3					4	1		2			1					3				3	1		21
13				1			1				1	1	7			1	1	9	6	3	2	3	1		1			29
14		3			2	2							6			1	1	8	7	1	5	39		2	1		1	71
15		2		1	2	1	2						7				2	2	1	1	6	13		1	2		2	43
16		4		3	4					1	1		6	1		1	1	9	7	2	4	3	2	3	5	2		50
17		1			5							1	10		1	1		9	7	2	4	4	1		2	16	2	57
18		1	1		2				1			1	1			1		1	1		2				1			12
19				1	2	1	2	1				1	1				1											10
20		1			2								1				1				2		1		5			13
	4	31	9	23	54	18	46	20	11	20	31	18	87	8	2	10	13	53	39	14	58	69	5	12	27	22	6	656

The clusters of gene with shared expression patterns identified in the analysis shown in Figure 4 are ordered so that patterns specific for the more immature precursors are on top, and those for the most mature cells are on the bottom. The last three rows represent U-shaped expression patterns that could not unambiguously be put in order with the other patterns. Next, genes in each cluster are subdivided into functional groups as indicated in the top row of the table. These functional groups appear in columns. The numbers of genes identified for each functional group in each cluster are given in the boxes, with darker color shading corresponding to higher gene numbers. Total numbers of genes are given on the lower and right borders. Note that these total numbers differ from those shown in Figure 4, because the table contains only known genes, whereas Figure 4 also includes ESTs.

able, this type of analysis remains incomplete. Nevertheless, we consider the selection of genes analyzed here to be representative for the entire genome, since in every group of genes with similar expression pattern, the ratio of known genes to (randomly chosen) ESTs is ~1:2, reflecting the proportion on the arrays.

Second, this analysis is based on differentially expressed genes that differ by a factor of two or more during B-cell differentiation. A number of examples demonstrate a functional significance of mono- versus bi-allelic expression (Nutt et al. 1999). Different genes, however, might require different changes in expression level to become functionally relevant. Therefore, not all of the genes analyzed here might actually contribute to the phenotypes of mouse B-cell precursors.

Third, many genes appear to exert different functions in different cellular contexts. Therefore, biochemical function of a gene product has to be seen in the context of the respective interaction partners present in a particular cell type. This context can signify different cellular effects of a given gene product in different cells. The functional associations presented here can only be a first evidence that these genes function in our given cellular setting, but not how exactly they exert their function.

Fourth, whereas the cells examined here have been ordered as consecutive stages of an ordered differentiation pathway, there might be additional, small, intermediate, transitional populations not distinguishable by current surface marker analysis. Moreover, a given cell might have distinct ways to react to external stimuli. Immature B cells, for example, are known to react to BCR crosslinking in a ligand-dependent manner. Although high-avidity antigens lead to negative selection, low- and medium-avidity antigens might lead to anergization or even positive selection (Hayakawa et al. 1999; Nemazee et al. 2000). The pool of immature B cells examined here might actually consist of cells before and at different points during the selection process.

A surprisingly high number of genes, that is, almost 11% of all, are differentially expressed in this single mammalian

Table 3. Genes Potentially Involved in Stage-Specific Properties of B Cells

Table 3A. Genes Up-Regulated in Pre-BI Cells

affy_id	Description	Function	PRE	LAR	SMA	IMM	MAT	Cluster	p(H)
j03535_s_at	Mouse Ig-related glycoprotein-70 mRNA, complete cds.	Adhesion	672	97	124	141	146	1	0.00536
d90362_s_at	Mouse mRNA for cadherin-associated protein (CAP102/alpha catenin).	Adhesion	5425	3215	2703	1932	1904	2	0.00042
x66532_f_at	*M. musculus* mRNA for L14 lectin.	Adhesion	3723	2406	1151	719	365	2	0.00008
J05020_s_at	Mouse mast cell high affinity IgE receptor (Fc-epsilon-RI) gamma subunit mRNA, complete cds.	intercell communic	3322	1008	1061	1317	1353	1	0.00176
M29855_s_at	Mouse interleukin-3 receptor mRNA, complete cds.	intercell communic	219	25	−35	16	−93	1	0.00091
y00864_s_at	Mouse c-kit mRNA.	intercell communic	262	75	75	23	2	1	0.00658
D10849_s_at	Mouse mRNA for thromboxane A2 receptor, complete cds.	intercell communic	614	13	125	124	182	1	0.00089
X77952_s_at	*M. musculus* (CD1) endoglin mRNA.	intercell communic	1103	−11	−102	−79	−532	1	0.0045
I38444_s_at	*Mus musculus* (clone U2) T-cell specific protein nRNA, complete cds.	Signaling	544	84	220	259	267	1	0.00126
U05245_s_at	*Mus musculus* BALB/c invasion inducing protein (Tiam-1) mRNA, complete cds.	Signaling	1662	414	196	129	117	1	0.00062
U88327_s_at	*Mus musculus* suppressor of cytokine signalling-2 (SOC-2) mRNA, complete cds.	Signaling	506	76	38	−39	−57	1	0.00195
u73200_s_at	*Mus musculus* p116Rip mRNA, complete cds.	Signaling	1038	841	589	353	404	2	0.00288

Table 3B. Genes Potentially Involved in Specific Properties of Large Pre BII Cells (pre-BCR Mediated Proliferative Burst and Allelic Exclusion)

affy_id	Description	Function	PRE	LAR	SMA	IMM	MAT	Cluster	p(H)
I32751_f_at	Mouse (clone M1) GTPase (Ran) mRNA, complete cds.	Signaling	6427	6620	3361	3062	2627	3	0.00014
X61940_s_at	Mouse mRNA for a growth factor-inducible immediate early gene (3CH134).(=mkp1)	Signaling	199	251	77	40	171	3	0.00105
Msa.502.0_s_at	*Mus musculus* calmodulin synthesis (CaM) cDNA, complete cds.	Signaling	4117	7007	4205	2621	2414	4	0.00213
M27844_f_at	*Mus musculus* calmodulin synthesis (CaM) cDNA, complete cds.	Signaling	357	823	260	203	212	4	0.00147
u01063_f_at	*Mus musculus* pLK serine/threonine kinase mRNA, complete cds.	Signaling	1363	2411	868	717	647	4	0.00113
u08110_s_at	*Mus musculus* RNA1 homolog (Fug1) mRNA, complete cds.	Signaling	1152	1793	435	442	465	4	0.0008
L11316_s_at	Mouse oncogene (ect2) mRNA, complete cds.	Signaling	326	565	59	37	100	4	0.0003
u80932_s_at	*Mus musculus* serine/threonine kinase Ayk1 (ayk1) mRNA, complete cds.	Signaling	1366	3389	276	203	237	4	0.00022
I29479_s_at	*Mus musculus* serine/threonine kinase (sak-a) mRNA, complete cds.	Signaling	501	972	95	57	171	4	0.00024
Msa.1795.0_at	*Mus musculus* citron, putative rho/rac effector, mRNA, complete cds.	Signaling	164	355	21	−4	−43	4	0.00159
D21099_s_at	Mouse mRNA fir STK-1 (serine/threonine kinase), complete cds.	Signaling	2624	4421	565	176	265	4	0.00027
x95351_s_at	*M. musculus* mRNA for pMELK protein.	Signaling	184	559	−126	−59	−80	4	0.00032
u95610_s_at	*Mus musculus* nimA-related kinase 2 (Nek2) mRNA, complete cds.	Signaling	1044	3922	326	132	134	4	0.00017
u15562_s_at	*Mus musculus* CDC25 (Cdc25) mRNA, complete cds.	Signaling	406	900	112	85	54	4	0.00129
U27323_s_at	*Mus musculus* Cdc25a (cdc25a) mRNA, complete cds.	Signaling	268	250	170	48	16	5	0.01373
u59418_s_at	*Mus musculus* protein phosphatase 2A Balpha3 regulatory subunit mRNA, partial cds.	Signaling	1456	1791	1224	768	645	5	0.00066
Msa.1896.0_s_at	*Mus musculus* protein phosphatase 2A Balpha3 regulatory subunit mRNA, partial cds.	SIgnaling	1068	1602	1025	451	553	5	0.00028
U43900_s_at	*Mus musculus* signal transducing adaptor molecule STAM mRNA, complete cds.	Signaling	334	411	338	226	141	5	0.01482

(Table continued on following page.)

Table 3B. (*Continued*)

affy_id	Description	Function	PRE	LAR	SMA	IMM	MAT	Cluster	p(H)
Msa.1497.0_f_at	Mouse calmodulin (Cam III) mRNA, complete cds	Signaling	523	864	480	286	241	5	0.01975
x65687_s_at	*M. musculus* mRNA for serine-threonine protein kinase. (=c-akt)	Signaling	572	1005	1157	700	466	9	0.0057
Msa.1702.0_f_at	*Mus musculus* retinoblastoma-binding protein (mRbAp46) mRNA, complete cds	chromatin structure	3932	3982	1954	1705	2252	3	0.00028
U85614_s_at	*Mus musculus* SRG3 mRNA, complete cds.	chromatin structure	2304	2629	909	762	670	3	0.00054
U35141_s_at	*Mus musculus* retinoblastoma-binding protein (mRbAp48) mRNA, complete cds.	chromatin structure	2740	3483	1146	882	918	3	0.00032
Msa.2344.0_f_at	Mouse mRNA for HMG-17 chromosomal protein	chromatin structure	4356	5780	2364	1128	1102	3	0.00017
X12944_f_at	Mouse mRNA for HMG-17 chromosomal protein	chromatin structure	4596	6597	2118	1294	1043	3	0.00008
AF034610_s_at	*Mus musculus* nuclear autoantigenic sperm protein mRNA, complete cds.	chromatin structure	352	432	53	78	182	3	0.0026
Msa.2456.0_g_at	Mouse mRNA for modifier 1 protein	chromatin structure	685	817	300	131	184	3	0.00043
Msa.2456.0_at	Mouse mRNA for modifer 1 protein	chromatin structure	225	467	72	−11	−28	4	0.00059
U70494_s_at	*Mus musculus* histone H2A.Z (H2A.Z) mRNA, complete cds.	chromatin structure	5757	9094	2830	1151	1251	4	0.00009
AF012709_s_at	*Mus musculus* centromere protein A (Cenp-a) mRNA, complete cds.	chromatin structure	2119	3969	1291	505	653	4	0.00006
X58069_s_at	Mouse mRNA for Histone H2A.X.	chromatin structure	3178	7120	1040	893	1101	4	0.00048
U42385_s_at	*Mus musculus* fibroblast growth factor inducible gene 16 (FIN16) mRNA, complete cds.	chromatin structure	1063	2228	126	187	252	4	0.00022
m37736_f_at	Mouse replication-dependent histone H2A,1 gene.	chromatin structure	10962	20136	2378	811	961	4	0.00027
z11997_s_at	*M. musculus* mRNA for non-histone chromosomal high-mobility gruop 1 protein.	chromatin structure	11234	12273	8043	4015	3323	5	0.00019
z30940_f_at	M.domesticus (CD-1) mRNA for histone H2A (partial).	chromatin structure	3683	4721	1972	827	433	5	0.0016
ET63372_f_at	M.domesticus (CD-1) mRNA for histone H3 (partial).	chromatin structure	460	969	487	−110	−113	5	0.00871

Table 3C. Genes Potentially Responsible for Cell Cycle Arrest after the Large Pre BII Cell Stage

affy_id	Description	Function	PRE	LAR	SMA	IMM	MAT	Cluster	p(H)
X53068_s_at	Mouse mRNA for proliferating cell nuclear antigen.	DNA Replication	4266	4920	2296	1557	1423	3	0.00033
X54149_s_at	Mouse mRNA for MyD118, a myeloid differentiation primary response gene.	Cell Cycle Reg	196	218	523	578	315	11	0.00232
U35623_s_at	*Mus musculus* EAT/MCL-1 mRNA, complete cds.	Apoptosis	1272	1132	2801	6862	4182	14	0.00008
AF016583_at	*Mus musculus* checkpoint kinase Chk1 (Chk1) mRNA, complete cds.	Signaling	285	236	118	56	68	2	0.00638
AF016583_g_at	*Mus musculus* checkpoint kinase Chk1 (Chk1) mRNA, complete cds.	Signaling	293	332	70	−7	78	3	0.0003
u15562_s_at	*Mus musculus* CDC25 (Cdc25) mRNA, complete cds.	Signaling	406	900	112	85	54	4	0.00129
U27323_s_at	*Mus musculus* Cdc25a (cdc25a) mRNA, complete cds.	Signaling	268	250	170	48	16	5	0.01373
d87663_s_at	House mouse; Musculus domesticus mRNA for 14-3-3 epsilon, complete cds.	Signaling	1613	1134	1771	1782	887	7	0.01699
M64292_s_at	Mouse TIS21 gene, complete cds.	Cell Cycle Reg	1539	2299	6332	3889	3293	11	0.00011
x12761_s_at	Mouse mRNA for protein homologous to human c-JUN.	Cell Cycle Reg	208	130	509	113	88	10	0.01608
I28177_s_at	Mouse Gadd45 mRNA, complete cds.	Cell Cycle Reg	103	281	449	72	199	19	0.01109
U19860_s_at	*Mus musculus* growth arrest specific mRNA, clone 3544, complete cds.	Cell Cycle Reg	468	445	823	381	88	10	0.00163

(Table continued on following page.)

Table 3. (*Continued*)

Table 3D. Genes Distinguishing Immature from Mature Cells

affy_id	Description	Function	PRE	LAR	SMA	IMM	MAT	Cluster	p(H)
J04170_s_at	Mouse B-cell differentiation antigen Lyb-2.1 protein, complete cds.(=CD72)	intercell communic	1015	1929	2856	5057	1906	13	0.00076
X12616_s_at	Mouse c-fes proto-oncogene mRNA for c-fes protein.	Signaling	252	77	26	399	167	13	0.00555
I11330_s_at	*Mus musculus* protein tyrosine phosphatase (PAC-1) mRNA, complete cds.	Signaling	674	484	1090	2079	790	13	0.00213
ET62844_f_at	*Mus musculus* immunoglobulin-like receptor PIRA6 (12M1) mRNA, complete cds.	intercell communic	147	111	154	263	257	14	0.00472
u96689_s_at	*Mus musculus* immunoglobulin-like receptor PIRB1 (7M5) mRNA, complete cds.	intercell communic	112	269	245	542	557	14	0.00949
x04648_s_at	Mouse mRNA for IgG1/IgG2b Fc receptor (FcR).	intercell communic	202	504	868	1877	1865	14	0.00012
j05479_s_at	Mouse calcineurin catalytic subunit mRNA, complete cds.	Signaling	457	522	682	1039	1046	14	0.00048
U68171_s_at	*Mus musculus* cyclic nucleotide phosphodiesterase PDE7A2 (MMPDE7A) mRNA, complete cds.	Signaling	144	286	304	491	426	14	0.00328
L02241_s_at	Mouse protein kinase inhibitor (testicular isoform) mRNA, complete cds.(=inhibitor of cAMP dependent protein kinase)	Signaling	154	18	92	376	248	14	0.00264
u57325_s_at	*Mus musculus* PS-2short mRNA, complete cds.	Apoptosis	357	211	950	1000	785	15	0.00087
U35623_s_at	*Mus musculus* EAT/MCL-1 mRNA, complete cds.	Apoptosis	1272	1132	2801	6862	4182	14	0.00008
U88908_s_at	*Mus musculus* inhibitor of apoptosis protein 1 mRNA, complete cds.	Apoptosis	229	301	1611	1578	2217	15	0.00016
u78031_s_at	*Mus musculus* apoptosis inhibitor bcl-x) gene, exon 3 and complete cds.	Apoptosis	530	999	1793	444	152	10	0.00027
Msa.463.0_at	Mouse bcl-2 gene encoding mbcl-2-beta	Apoptosis	122	−38	−8	154	511	17	0.00028
L16956_at	Mouse protein-tyrosine kinase (Jak2) mRNA, complete cds.	Signaling	198	225	228	157	502	17	0.01466
L16956_at	Mouse protein-tryosine kinase (Jak2) mRNA, complete cds.	Signaling	198	225	228	157	502	17	0.01466
L16956_g_at	Mouse protein-tryosine kinase (Jak2) mRNA, complete cds.	Signaling	47	73	105	21	505	17	0.00615
u51907_s_at	*Mus musculus* TRAF family member associated NF-kappa B activator (TANK) mRNA, complete cds.	Signaling	93	103	273	247	444	17	0.00013
U12919_s_at	*Mus musculus* adenylyl cyclase type VII mRNA, complete cds.	Signaling	187	113	316	315	587	17	0.01204
D13759_s_at	Mouse mRNA for proto-oncogene protein.(=MAPKKK8)	Signaling	−37	57	−44	26	143	17	0.0056
x83932_s_at	*M. musculus* mRNA for ryanodine receptor type 1.	Signaling	−11	11	−1	96	342	17	0.00128
j05265_s_at	Mouse interferon gamma receptor mRNA, complete cds.	intercell communic	661	521	644	542	1178	17	0.01011
U53696_s_at	*Mus musculus* class II cytokine receptor 4 (CRF2–4) mRNA, complete cds.(=IL-10 receptor)	intercell communic	182	163	218	293	693	17	0.00115
L31580_s_at	Mouse G protein-coupled receptor (EBI 1) mRNA, complete cds.(+CCR7)	intercell communic	178	280	606	529	1330	17	0.00024
Msa.549.0_s_at	Murine complement receptor type 2 (CR2) mRNA, complete cds	intercell communic	−10	−10	29	98	581	17	0.00145
X71788_s_at	*M. musculus* bir-1 gene for Burkitt lymphoma receptor 1 homologue. (=CXCR5)	intercell communic	−29	−56	−40	276	1076	17	0.00037
m18184_s_at	Mouse lymphocyte differentiation antigen (Ly-6.2) mRNA, complete cds. (=TAP)	intercell communic	689	63	119	1857	10119	17	0.00006
m34163_s_at	Mouse low affinity IgE receptor (Fc-epsilon-RII) mRNA, complete cds.	intercell communic	−13	3	50	1471	6536	17	0.00037

Genes distinguishing immature from mature cells with potential implications for positive/negative selection and response to B-cell receptor crosslinking (apoptosis vs. proliferation).

(Table continued on following page.)

Table 3. (*Continued*)

Table 3E. Expression of Cell Adhesion Molecules During Mouse B Cell Development

affy_id	Description	Function	PRE	LAR	SMA	IMM	MAT	Cluster	p(H)
j03535_s_at	Mouse Ig-related glycoprotein-70 mRNA, complete cds.	Adhesion	672	97	124	141	146	1	0.00536
d90362_s_at	Mouse mRNA for cadherin-associated protein (CAP102/alpha catenin).	Adhesion	5425	3215	2703	1932	1904	2	0.00042
x66532_f_at	*M. musculus* mRNA for L14 lectin.	Adhesion	3723	2406	1151	719	365	2	0.00008
ET62820_s_at	*Mus musculus* fertilization antigen-1 mRNA, complete cds.	Adhesion	5327	6402	4779	2913	2626	5	0.00043
X91144_s_at	*M. musculus* mRNA for P-selectin glycoprotein ligand 1.	Adhesion	3315	2513	2897	1638	1776	8	0.00306
AF021031_at	*Mus musculus* Dgcr6 protein (Dgcr6) mRNA, partial cds.	Adhesion	285	248	341	174	0	8	0.00964
u24703_s_at	*Mus musculus* reelin mRNA, complete cds.	Adhesion	14	106	2028	196	21	10	0.00088
z48781_s_at	*M. musculus* Stra1 mRNA for Stra1/Eplg2 protein.	Adhesion	69	119	225	141	298	12	0.01548
j03723_rc_g_at	Mouse carbohydrate binding protein 35 mRNA, 3 end.(galectin 3)	Adhesion	39	−37	183	129	205	12	0.00751
Msa.3206.0_s_at	*Mus musculus* CD37 gene	Adhesion	940	572	1664	2548	1582	13	0.01823
x52264_s_at	<pise ICAM-1 mRNA for intercellular adhesion molecule-1.	Adhesion	173	56	68	370	223	14	0.01131
x65490_at	*M. musculus* ICAM-2 gene for intercellular adhesion molecule-2, exon 1.	Adhesion	582	1135	1269	1431	1858	15	0.00012
x14951_s_at	*M. musculus* mRNA for CD18 antigen beta subunit, leukocyte adhesion protein (LFA-1).	Adhesion	625	371	889	1217	1288	16	0.00121
X96618_s_at	*M. musculus* mRNA for novel stromal cell protein.	Adhesion	473	442	973	1475	1588	16	0.00139
d14883_s_at	Mouse mRNA for C33/R2/IA4, complete cds.	Adhesion	549	1165	920	2140	2563	16	0.00427
x67783_s_at	*M. musculus* VCAM-1 mRNA.	Adhesion	40	−62	51	197	222	16	0.00348
I16928_s_at	*Mus musculus* differentiation antigen (CD22) mRNA, complete cds.	Adhesion	73	125	359	3098	3232	16	0.00038
m95632_s_at	Mouse integrin beta-7 subunit mRNA, complete cds.	Adhesion	140	214	71	277	1226	17	0.00257
m25324_s_at	Mouse peripheral lymph node-specific homing receptor (MEL-14 antigen) mRNA, complete cds.	Adhesion	1637	739	981	417	7757	17	0.00017
m18466_f_at	mouse lymphocyte differentiation antigen Ly-6C.2 mRNA, complete cds.	Adhesion	1226	474	453	743	1173	18	0.00026
ET61114_f_at	House mouse; Musculus domesticus mRNA for Ly-6C variant, complete cds.	Adhesion	896	434	327	454	1250	20	0.00155
Msa.918.0_f_at	Mouse lymphocyte differentiation antigen Ly-6C.2 mRNA, complete cds	Adhesion	935	334	303	767	1319	20	0.00262
Msa.164.0_at	*Mus musculus* ecto-5-mucleotidase (Cd73) mRNA, complete cds	Adhesion	235	116	276	141	310	20	0.01277
I08235_s_at	*Mus musculus* clusterin mRNA, complete cds.	Adhesion	486	303	373	328	843	20	0.01645
u40408_s_at	*Mus musculus* lysosomal sialic acid O-acetylesterase mRNA, complete cds.	Adhesion	140	11	48	−22	128	20	0.00526

(Affy_id) unique Affymetrix probe set identifier; (Description) gene description; (Function) functional classification; (PRE/LAR/SMA/IMM/MAT) average expression level (mean of average differences of five independent replicate experiments) in Pre-BI, large Pre-BII, small Pre-BII, immature, and mature B cells, respectively; (Cluster) cluster designation from Fig. 4. p(H) indicates *P*-value for differential expression (Kruskal-Wallis test).

developmental pathway. This indicates that there must be substantial redundancy in usage of genes between different cellular differentiation programs. Additional gene expression studies of other lymphoid and nonlymphoid cellular differentiation programs shall identify genes that are used in many developmental pathways and distinguish them from those that are truly pathway-specific. These unique genes appear to be the best candidates for genetic modification in germline or somatic cells, leading to interpretable functional changes.

METHODS

Purification of B-Cell Precursors

Total femoral bone marrow cells of five- to six-week-old C57/BL6 mice ($n = 4$ per experiment) were aliquotted into three parts, stained, and sorted as shown in Figure 1. A total of 50,000 (pre-BI, large pre-BII) or 150,000 (small pre-BII, immature and mature B cells) cells were sorted directly into TRIzol RNA isolation reagent (Life Technologies) at 50,000 cells/500 µL TRIzol. A cell purity of ≥98% was routinely achieved.

RNA Amplification

Total cellular RNA was precipitated with 5 µg *Escherichia coli* tRNA as coprecipitant, and pellets were resuspended in 9 µL DEPC-H_2O. In vitro transcription-based RNA amplification was performed essentially as described earlier (Eberwine et al. 1992; Luo et al. 1999). Briefly, double-stranded cDNA was synthesized using the T7-tagged oligo-dT primer (5′-GGCCAGTGAATTGTAATACGACTCACTATAGGG AGGCGG$(T)_{24}$-3′), and, after cleanup with Microcon YM-50 columns, a first in vitro transcription reaction was done using the Promega RiboMax T7 kit. RNA was cleaned up again, and a random-hexamer primed first strand cDNA synthesis was

performed. The resulting RNA–DNA hybrid was treated with RNAseH, heat denatured, and the T7-oligo-d(T)-Primer was annealed. Second-strand cDNA synthesis was the performed with a mix of Klenow and T4 DNA polymerases. After cleanup, a second in vitro transcription was performed, incorporating biotinylated CTP and UTP, respectively. RNA was purified with RNeasy columns (QIAGEN).

Hybridization of Affymetrix GeneChip Probe arrays

Affymetrix Mu11k GeneChips, interrogating 13,104 genes and ESTs on two different arrays, were processed according to the manufacturer's recommendations. Approximately 20 pairs of oligonucleotides (probe pairs) represent every gene or EST in a probe set, with one sequence being complementary to the target sequence and the other having a 1-bp mismatch in a central position. This serves as an internal control for hybridization specificity. The relative abundance is reported as the averaged difference of the fluorescence intensity values between the perfectly matched and the mismatched oligonucleotides, resulting in the so-called "average difference" value (Lockhart et al. 1996; Wodicka et al. 1997).

Data Analysis

To enable statistically valid data analysis, five independent replicate experiments were performed. Analysis of this data set consisted of three steps. First, normalization of the individual arrays; second, identification of differentially expressed genes between adjacent developmental stages using pair-wise comparisons; and third, identification of sets of differentially expressed genes sharing similar changes across the sequence of five B-cell developmental stages.

First, individual GeneChip expression results were normalized as follows. All average difference values of every chip were summed up, and the mean of these sums across all chips of the same layout was calculated. The ratio of the actual average difference sum for any given chip and the mean of all average difference sums across all chips with the same layout served as a correction factor for this chip, with which all the average difference values were multiplied. This process was performed separately for the two different array types used (Mu11k sub A and sub B arrays, respectively). It results in normalized average difference values, so that all chips of the same array layout have the same overall average difference.

Second, differentially expressed genes between developmentally adjacent stages were identified as follows. In this analysis, average difference values beyond the values of most other individual hybridization values seen in five replicate experiments were removed using the Nalimov outlier test at 95% confidence level (Kaiser and Gottschalk 1972). Means and standard deviations of the remaining values were calculated for every gene separately for every developmental stage. Probability of differential expression was calculated with unpaired *t*-test statistics. Genes were considered differentially expressed if the *t*-test results in a confidence level of at least 98%, and if they changed at least twofold with a difference in mean average difference value of at least 100.

Third, sets of genes were identified that shared similar changes in their expression patterns during the five B-cell developmental stages. For this multicomparison analysis, ANOVA with H-(Kruskal-Wallis) test statistics was performed to identify genes with a statistically significant differential expression in at least one of all the possible pairwise comparisons. Threshold criteria were set as above. Hierarchical Cluster Analysis was performed according to the method of Eisen et al. (1998) after row normalization to a mean of 0 and a sum of value squares of 1. Genes with similar expression patterns were identified with self-organizing maps (Tamayo et al. 1999).

As an additional quality criterion, the homogeneity of the hybridization signal across the 20 gene-specific probe pairs was assessed for all genes mentioned in this paper as follows. First, individual probe pairs are scored as "positive" (PM intensity > MM intensity) or "negative" (PM intensity < MM intensity) according to standard Affymetrix procedures. Next, the "purity" value was calculated as follows.

Purity = [(number of pairs in the probe set) − least (number of positive pairs, number of negative pairs) × 2)]/ number of pairs in the probe set.

This value indicates how well probe pairs agree in their signal. In our analyses, the purity value had to be ≥0.7 in at least four replicate experiments of the population with the highest expression level to be considered significant.

ACKNOWLEDGMENTS

We are indebted to Mark Dressing, Annette Pickert, Hubertus Kohler, and Tracy Hayden for cell sorter operation. The Basel Institute for Immunology was founded and supported by F. Hoffmann-LaRoche Ltd.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

REFERENCES

Cheng, G. and Baltimore, D. 1996. TANK, a co-inducer with TRAF2 of TNF- and CD 40L-mediated NF-kappaB activation. *Genes & Dev.* **10:** 963–973.

Chin, D. and Means, A.R. 2000. Calmodulin: A prototypical calcium sensor. *Trends Cell Biol.* **10:** 322–328.

Coutavas, E.E., Hsieh, C.M., Ren, M., Drivas, G.T., Rush, M.G., and D'Eustachio, P.D. 1994. Tissue-specific expression of Ran isoforms in the mouse. *Mamm. Genome* **5:** 623–628.

Eberwine, J., Yeh, H., Miyashiro, K., Cao, Y., Nair, S., Finnell, R., Zettel, M., and Coleman, P. 1992. Analysis of gene expression in single live neurons. *Proc. Natl. Acad. Sci.* **89:** 3010–3014.

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.* **95:** 14863–14868.

Fode, C., Motro, B., Yousefi, S., Heffernan, M., and Dennis, J.W. 1994. Sak, a murine protein-serine/threonine kinase that is related to the *Drosophila* polo kinase and involved in cell proliferation. *Proc. Natl. Acad. Sci.* **91:** 6388–6392.

Fu, H., Subramanian, R.R., and Masters, S.C. 2000. 14–3–3 proteins: Structure, function, and regulation. *Annu. Rev. Pharmacol. Toxicol.* **40:** 617–647.

Fujise, K., Zhang, D., Liu, J.L., and Yeh, E.T. 2000. Regulation of apoptosis and cell cycle progression by MCL1: Differential role of PCNA. *J. Biol. Chem.* **275:** 39458–39465.

Gerondakis, S., Economou, C., and Grumont, R.J. 1994. Structure of the gene encoding the murine dual specificity tyrosine–threonine phosphatase PAC1. *Genomics* **24:** 182–184.

Giannini, G., Conti, A., Mammarella, S., Scrobogna, M., and Sorrentino, V. 1995. The ryanodine receptor/calcium channel genes are widely and differentially expressed in murine brain and peripheral tissues. *J. Cell Biol.* **128:** 893–904.

Glynne, R., Ghandour, G., Rayner, J., Mack, D.H., and Goodnow, C.C. 2000. B-lymphocyte quiescence, tolerance and activation as viewed by global gene expression profiling on microarrays. *Immunol. Rev.* **176:** 216–246.

Grawunder, U., Leu, T.M., Schatz, D.G., Werner, A., Rolink, A.G., Melchers, F., and Winkler, T.H. 1995. Down-regulation of RAG1 and RAG2 gene expression in preB cells after functional immunoglobulin heavy chain rearrangement. *Immunity* **3:** 601–608.

Hackenmiller, R., Kim, J., Feldman, R.A., and Simon, M.C. 2000. Abnormal Stat activation, hematopoietic homeostasis, and innate immunity in c-fes−/− mice. *Immunity* **13:** 397–407.

Hayakawa, K., Asano, M., Shinton, S.A., Gui, M., Allman, D., Stewart, C.L., Silver, J., and Hardy, R.R. 1999. Positive selection of natural autoreactive B cells. *Science* **285:** 113–116.

Jeon, S.H., Kang, M.G., Kim, Y.H., Jin, Y.H., Lee, C., Chung, H.Y., Kwon, H., Park, S.D., and Seong, R.H. 1997. A new mouse gene, SRG3, related to the SWI3 of *Saccharomyces cerevisiae*, is required for apoptosis induced by glucocorticoids in a thymoma cell line. *J. Exp. Med.* **185:** 1827–1836.

Kaiser, R. and Gottschalk, G. 1972. Ausreissertest nach Nalimov. In *Elementare Tests zur Beurteilung von Messadaten*, pp. 18–21. Bibliographisches Institut, Mannheim, Wien, Zürich.

Karasuyama, H., Rolink, A., Shinkai, Y., Young, F., Alt, F.W., and Melchers, F. 1994. The expression of Vpre-B/lambda 5 surrogate light chain in early bone marrow precursor B cells of normal and B cell-deficient mutant mice. *Cell* **77:** 133–143.

Lang, J., Arnold, B., Hammerling, G., Harris, A.W., Korsmeyer, S., Russell, D., Strasser, A., and Nemazee, D. 1997. Enforced Bcl-2 expression inhibits antigen-mediated clonal elimination of peripheral B cells in an antigen dose-dependent manner and promotes receptor editing in autoreactive, immature B cells. *J. Exp. Med.* **186:** 1513–1522.

Lin, X., Cunningham Jr., E.T., Mu, Y., Geleziunas, R., and Greene, W.C. 1999. The proto-oncogene Cot kinase participates in CD3/CD28 induction of NF- kappaB acting through the NF-kappaB-inducing kinase and IkappaB kinases. *Immunity* **10:** 271–280.

Lockhart, D.J., Dong, H., Byrne, M.C., Follettie, M.T., Gallo, M.V., Chee, M.S., Mittmann, M., Wang, C., Kobayashi, M., Horton, H., and Brown, E.L. 1996. Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat. Biotechnol.* **14:** 1675–1680.

Luo, L., Salunga, R.C., Guo, H., Bittner, A., Joy, K.C., Galindo, J.E., Xiao, H., Rogers, K.E., Wan, J.S., Jackson, M.R., and Erlander, M.G. 1999. Gene expression profiles of laser-captured adjacent neuronal subtypes [published Erratum appears in Nat. Med. 1999, Mar., **5(3):** 355]. *Nat. Med.* **5:** 117–122.

Maeda, A., Kurosaki, M., Ono, M., Takai, T., and Kurosaki, T. 1998. Requirement of SH2-containing protein tyrosine phosphatases SHP-1 and SHP-2 for paired immunoglobulin-like receptor B (PIR-B)-mediated inhibitory signal. *J. Exp. Med.* **187:** 1355–1360.

Melchers, F. and Rolink, A. 1999. B-Lymphocyte development and biology. In *Fundamental immunology* (ed. W.E. Paul), pp. 183–224. Lippincott-Raven, Philadelphia, New York.

Nemazee, D., Kouskoff, V., Hertz, M., Lang, J., Melamed, D., Pape, K., and Retter, M. 2000. B-cell-receptor-dependent positive and negative selection in immature B cells. *Curr. Top. Microbiol. Immunol.* **245:** 57–71.

Niwa, H., Abe, K., Kunisada, T., and Yamamura, K. 1996. Cell-cycle-dependent expression of the STK-1 gene encoding a novel murine putative protein kinase. *Gene* **169:** 197–201.

Nutt, S.L., Vambrie, S., Steinlein, P., Kozmik, Z., Rolink, A., Weith, A., and Busslinger, M. 1999. Independent regulation of the two Pax5 alleles during B-cell development. *Nat. Genet.* **21:** 390–395.

Parnes, J.R. and Pan, C. 2000. CD72, a negative regulator of B-cell responsiveness. *Immunol Rev* **176:** 75–85.

Rolink, A., Grawunder, U., Winkler, T.H., Karasuyama, H., and Melchers, F. 1994. IL-2 receptor alpha chain (CD25, TAC) expression defines a crucial stage in pre-B cell development. *Int. Immunol.* **6:** 1257–1264.

Rolink, A.G., Winkler, T., Melchers, F., and Andersson, J. 2000. Precursor B cell receptor-dependent B cell proliferation and differentiation does not require the bone marrow or fetal liver environment [see Comments]. *J. Exp. Med.* **191:** 23–32.

Rusnak, F. and Mertz, P. 2000. Calcineurin: Form and function. *Physiol. Rev.* **80:** 1483–1521.

Sanchez, Y., Wong, C., Thoma, R.S., Richman, R., Wu, Z., Piwnica-Worms, H., and Elledge, S.J. 1997. Conservation of the Chk1 checkpoint pathway in mammals: Linkage of DNA damage to Cdk regulation through Cdc25 [see Comments]. *Science* **277:** 1497–1501.

Singh, P.B., Miller, J.R., Pearce, J., Kothary, R., Burton, R.D., Paro, R., James, T.C., and Gaunt, S.J. 1991. A sequence motif found in a *Drosophila* heterochromatin protein is conserved in animals and plants. *Nucleic Acids Res.* **19:** 789–794.

Tamayo, P., Slonim, D., Mesirov, J., Zhu, Q., Kitareewan, S., Dmitrovsky, E., Lander, E.S., and Golub, T.R. 1999. Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. *Proc. Natl. Acad. Sci.* **96:** 2907–2912.

Tanaka, K., Parvinen, M., and Nigg, E.A. 1997. The in vivo expression pattern of mouse Nek2, a NIMA-related kinase, indicates a role in both mitosis and meiosis. *Exp. Cell Res.* **237:** 264–274.

Tatsumoto, T., Xie, X., Blumenthal, R., Okamoto, I., and Miki, T. 1999. Human ECT2 is an exchange factor for Rho GTPases, phosphorylated in G2/M phases, and involved in cytokinesis. *J. Cell. Biol.* **147:** 921–928.

ten Boekel, E., Melchers, F., and Rolink, A. 1995. The status of Ig loci rearrangements in single cells from different stages of B cell development. *Int. Immunol.* **7:** 1013–1019.

———. 1997. Changes in the V(H) gene repertoire of developing precursor B lymphocytes in mouse bone marrow mediated by the pre-B cell receptor. *Immunity* **7:** 357–368.

Tonegawa, S. 1983. Somatic generation of antibody diversity. *Nature* **302:** 575–581.

Vairapandi, M., Azam, N., Balliet, A.G., Hoffman, B., and Liebermann, D.A. 2000. Characterization of MyD118, Gadd45, and proliferating cell nuclear antigen (PCNA) interacting domains. PCNA impedes MyD118 AND Gadd45-mediated negative growth control. *J. Biol. Chem.* **275:** 16810–16819.

von der Mark, K. and Mollenhauer, J. 1997. Annexin V interactions with collagen. *Cell Mol. Life Sci.* **53:** 539–545.

Wickramasinghe, D., Becker, S., Ernst, M.K., Resnick, J.L., Centanni, J.M., Tessarollo, L., Grabel, L.B., and Donovan, P.J. 1995. Two CDC25 homologues are differentially expressed during mouse development. *Development* **121:** 2047–2056.

Winkler, T.H., Melchers, F., and Rolink, A.G. 1995. Interleukin-3 and interleukin-7 are alternative growth factors for the same B-cell precursors in the mouse. *Blood* **85:** 2045–2051.

Wodicka, L., Dong, H., Mittmann, M., Ho, M.H., and Lockhart, D.J. 1997. Genome-wide expression monitoring in *Saccharomyces cerevisiae*. *Nat. Biotechnol.* **15:** 1359–1367.

Yu, W., Nagaoka, H., Jankovic, M., Misulovin, Z., Suh, H., Rolink, A., Melchers, F., Meffre, E., and Nussenzweig, M.C. 1999. Continued RAG expression in late stages of B cell development and no apparent re-induction after immunization. *Nature* **400:** 682–687.

Zhang, Q., Vo, N., and. Goodman, R.H. 2000. Histone binding protein RbAp48 interacts with a complex of CREB binding protein and phosphorylated CREB. *Mol. Cell. Biol.* **20:** 4970–4978.

Received June 19, 2001; accepted in revised form October 16, 2001.