

Rules for Gene Usage Inferred from a Comparison of Large-Scale Gene Expression Profiles of T and B Lymphocyte Development¹

Reinhard Hoffmann,^{2,*†} Ludovica Bruno,^{*‡} Thomas Seidl,^{*‡} Antonius Rolink,^{*§} and Fritz Melchers^{*¶}

Ribonucleic acid expression profiles of seven consecutive stages of mouse thymocyte development were generated on high density oligonucleotide arrays. Previously known expression patterns of several genes were confirmed. Ten percent (1,304 of more than 13,000) of the monitored genes were found with 99% confidence to be differentially expressed across all T cell developmental stages. When compared with 1,204 genes differentially expressed in five consecutive B lineage developmental stages of bone marrow, >40% (546 genes) appeared to be shared by both lineages. However, when four pools of functionally distinct cell stages were compared between B and T cell development, DJ-rearranged precursor cells and resting immature precursor cells before and after surface Ag receptor expression shared less than 10%, mature resting lymphocytes between 15 and 20%, and only cycling precursors responding to precursor lymphocyte receptor deposition shared >50% of these differentially expressed genes. Three general rules emerge from these results: 1) proliferation of cells at comparable stages is in majority executed by the same genes; 2) intracellular signaling and intercellular communication are effected largely by different genes; and 3) most genes are not used strictly at comparable, but rather at several, stages, possibly in different functional contexts. *The Journal of Immunology*, 2003, 170: 1339–1353.

B and T lymphocytes are closely related differentiation lineages. Both originate from a common lymphoid progenitor and develop through a strikingly similar sequence of cellular stages before becoming mature lymphocytes expressing Ag-specific receptors. These cellular differentiation programs are ordered by stepwise rearrangements of Ig H and L chain gene loci in B cells and of TCR γ -, δ -, β -, and α -chains in T cells (1, 2).

During this developmental process, lymphoid progenitors, precursors, and immature and mature cells express stage- and lineage-specific surface markers. Five stages can be distinguished in B cell development by expression of B220, *c-kit*, CD25, and IgM (3). In T cell development, the expression of CD25, CD44, CD4, and CD8 distinguishes eight cellular differentiation stages (Fig. 1A) (2). At specific stages of these programs, i.e., when precursors express precursor receptors composed of TCR β -chains or the pre-T α chains and μ H chains or surrogate L chain, respectively, they are induced to proliferate. Hence, FACS separation and purification of different stages of T and B cell development use specific marker expression as well as cell size to identify cycling vs resting cells at different stages of development.

The two developmental pathways of TCR α/β -expressing T cells and of B cells are comparable at many stages (Fig. 1). In B cell development, five stages can be distinguished. From early progenitors, in which D_H to J_H segments are being rearranged on the H chain loci, the first type of cells of this sequence, called pre-BI cells, is generated. They are D_H to J_H rearranged on both H chain alleles, and express the receptor tyrosine kinase *c-kit*, but not yet CD25. A fraction of them is in cell cycle, i.e., consists of large cells. Pre-BI cells express the VpreB and λ 5 genes encoding the surrogate L chain, ready to form the pre-B cell receptor (BCR)³ as soon as V to DJ rearrangements are initiated on the H chain loci.

In T cell development, eight stages can be distinguished. Two of them, the double-negative 1 (DN1) and DN2 stages, are probably early progenitor stages in the process of commitment to the T lineage or to other hemopoietic lineages present in the thymus. DN3 cells are D β J β rearranged, probably also on both alleles, and express pre-TCRs with *pre-T α* and TCR γ -chains. They are prepared to form either TCR γ/δ or TCR α/β receptors. These DN3 cells appear at a stage comparable to the pre-BI cell stage in B cell development.

Once V to DJ rearrangements are initiated and the in-frame rearranged loci allow the production of IgH or TCR β -chains, respectively, precursor receptors can be assembled if the chains are capable of pairing with each other (4). The deposition of the precursor receptors in the surface membrane of the pre-BII cells, the DN4 and double-positive large (DPL) cells, respectively, stimulates proliferation of these cells (5). In both lineages, an early stage of this proliferative expansion can be distinguished from a late stage. In pre-BII cells, sterile transcription of the L chain loci has been initiated, while early stages do not yet transcribe these loci. Early stages of the pre-TCR-dependent proliferating stages are the

*Basel Institute for Immunology, Basel, Switzerland; [†]Department of Bacteriology, Max-von-Pettenkofer-Institute, München, Germany; [‡]Institute of Cancer Research, Chester Beatty Laboratories, London, United Kingdom; and [§]Pharmazentrum, Department of Immunology, and [¶]Biozentrum, Department of Cell Biology, Universität Basel, Basel, Switzerland

Received for publication June 28, 2002. Accepted for publication November 12, 2002.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¹ The Basel Institute for Immunology was founded and supported by F. Hoffmann-LaRoche AG.

² Address correspondence and reprint requests to Dr. Reinhard Hoffmann, Department of Bacteriology, Max-von-Pettenkofer-Institute, Pettenkoferstrasse 9A, D-80336 Munich, Germany. E-mail address: R_hoffmann@m3401.mpk.med.uni-muenchen.de

³ Abbreviations used in this paper: BCR, B cell receptor; DN, double negative; DP, double positive; DPL, DP large; DPS, DP small; EST, expressed sequence tag; SP, single positive; SOM, self-organizing map.

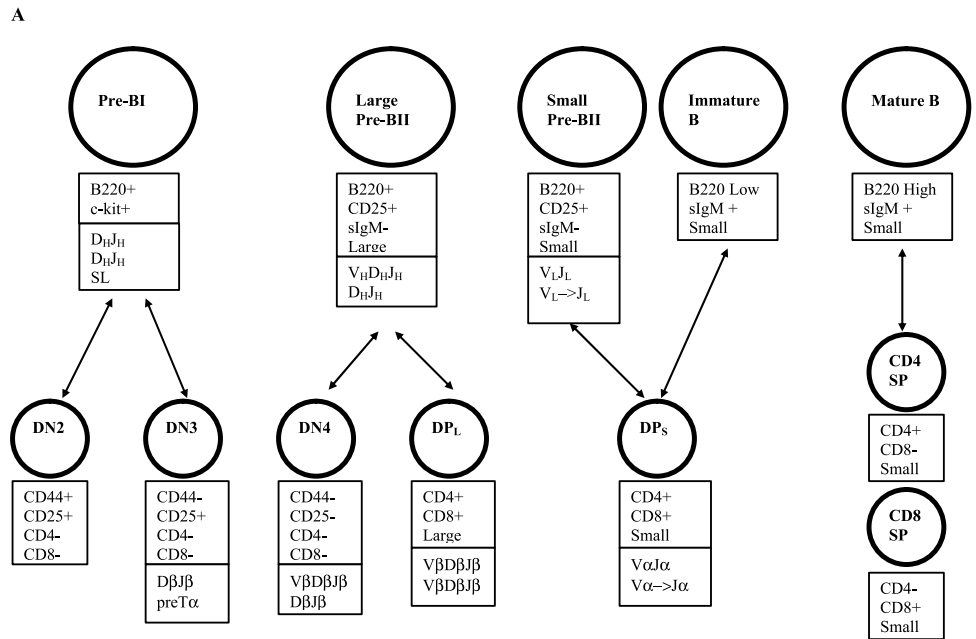
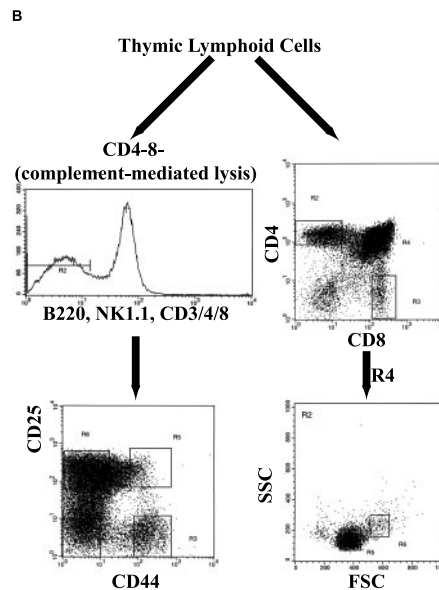


FIGURE 1. Surface markers and sorting strategy for isolation of thymic subsets. *A*, Synopsis of cellular populations and surface markers used for isolation of lymphocyte developmental stages. *Top*, B cell development; *bottom*, T cell development. Cellular stages are depicted as circles. Markers used for separation of the stages appear in the *upper boxes*; Ig and TCR gene loci rearrangement status appear in the *lower boxes*. Arrows connect corresponding stages between B and T cell development. SL, surrogate L chain. *B*, Cell-sorting strategy for separation of T cell precursors. To obtain DN thymocytes, CD4- and CD8-positive cells were removed by complement-mediated lysis from thymus single cell suspension. Remaining lymphocytes were stained with a panel of lineage markers (B220, NK.1.1, CD3, CD4, CD8), and negative cells (R2 gate in *top left panel*) were gated and analyzed for CD25 and CD44 surface expression (*bottom left*). For the isolation of DP and SP subsets, single cell suspension was stained with CD4 and CD8 mAbs (*top right*); SP cells were sorted according to gates R2 for CD4⁺ and R3 for CD8⁺. DP thymocytes (gate R4 in *top right panel*) were further resolved into large (R5) and small (R6) subsets (*bottom right*).



DN4 cells, while the later stages are the DPL cells, i.e., they are separable by the expression of CD4 and CD8. Practically all these cells are in cell cycle.

Because the expression of surrogate L chain in large pre-BII cells, and of pre-T α in DN4/DPL cells, is turned off once they have entered proliferation, the capacity of these cells to proliferate is limited by the amount of pre-BCR or pre-TCR, respectively, made in the preceding stage of development. Hence, after several divisions, these cycling cells fall into a resting state, where they begin rearrangements of the Ig κ and Ig λ L chain loci or TCR α -chain gene locus. At both the IgL and the TCR α -chain gene loci, secondary rearrangements can continue in Ag receptor-expressing immature cells, i.e., in immature B cells and immature small CD4⁺CD8⁺ (DNS) thymocytes, possibly influenced by the quality and specificity of the Ag receptors expressed on individual cells (6).

Both B and T lineage cells undergo repertoire selection at these Ag receptor-expressing immature stages (7, 8). Negative selection deletes cells with those Ag receptors that recognize autoantigens with high

avidity in the primary lymphoid organs, bone marrow or thymus, in which they are generated. Immature T and B cells also are positively selected. B cells with low avidity for autoantigens appear to be selected into the B1 compartment, prevalent in the gut-associated lymphoid organs. Double-positive small (DPS) cells recognize processed, proteolytically degraded peptides of autoantigens complexed in either MHC class I or II molecules on the surface of APCs with low avidity. This induces additional contacts of either MHC class I and CD8, or MHC class II and CD4, and leads to the down-regulation of the expression of CD4 in CD8-contacted cells, or of CD8 in CD4-contacted cells, leading to CD4 or CD8 single-positive (SP) cells. It is evident that immature lymphocytes must be a heterogeneous mixture of cells under different selective pressures, reacting by induction of cell death or survival to these pressures.

Once T and B lymphocytes have been selected to survive, they become, as mature cells, functionally more homogeneous again. CD4⁺ T cells, CD8⁺ T cells, B1 B cells, and conventional B cells are resting cells, awaiting the invasion of foreign Ags.

We have shown previously that a very large number of genes is differentially expressed during B cell development (9). Although a role in both B and T cell development has been established for some genes (such as the Ag receptors Ig and TCR, their anchoring molecules $I\alpha/I\beta$ and CD3, the precursor receptors surrogate L chain and pre-T α , the recombination machinery RAG-1/-2, and some others), it is not known how many genes are coexpressed during the development of B and T cells. It might be expected, due to the similarity of the cellular stages of B and T cell development and their functional programs, that the molecular execution of these programs could show a substantial overlap in genes that are specifically up-regulated at corresponding stages. To address these questions, we analyze in this work the gene expression programs of T cell development in the thymus of mice by generating high density oligonucleotide array-based gene expression profiles (10, 11) from highly purified, ex vivo isolated thymic T cell precursors similarly to those described earlier for bone marrow B cell precursors (9). We then compare these gene expression profiles of T cells at seven stages of their development in thymus with the previously generated profiles of five stages of B cell development in bone marrow.

Materials and Methods

Isolation of mouse T cell precursors

Cell suspensions from 4-wk-old B6 mice (10 mice per experiment) were divided in two portions. One portion was used for the isolation of DP and SP thymocytes and was directly stained with CD4 and CD8 mAbs. The second portion of the sample was used to isolate DN cells. In particular, DN thymocytes were obtained by depleting CD4⁺CD8⁺ cells with RL172.4 (anti-CD4) and 3.168.8.1 (anti-CD8) supernatants in DMEM for 10 min at 37°C. Complement (Low Tox-M rabbit complement; Cedarlane Laboratories, Hornby, Ontario, Canada) was added subsequently, and the suspension was left for 45 min at 37°C. Live cells were recovered by Ficoll density-gradient centrifugation. Following CD4 and CD8 depletion, samples were stained with a cocktail of lineage-specific mAbs (B220, NK11, CD3, CD4, CD8) APC conjugated together with CD25 (FITC conjugated) and CD44 (PE conjugated). For each population, 150,000 cells were sorted into TRIzol RNA isolation reagent. Each experiment was replicated five times. Aliquots were sorted into buffer for postsort reanalysis. All cell populations were $\geq 99\%$ pure, except for DPL cells, which were $\geq 95\%$ pure due to occasional doublets.

RNA amplification and hybridization to oligonucleotide probe arrays

In vitro transcription-based RNA amplification was performed, as described earlier (9, 12, 13). A total of 20 μ g of labeled cRNA was fragmented and hybridized to the Affymetrix (Santa Clara, CA) GeneChip mU1k array set, consisting of two separate arrays interrogating a total of 13,104 sequences, including control genes. GeneChip arrays were washed, stained, and scanned, according to the manufacturer's recommendations, and images were analyzed with Affymetrix GeneChip version 3.2 software. As a measure of relative gene expression, the average difference value (average of the differences in fluorescence intensity between matching and mismatched oligonucleotide probes within a gene-specific probe set) was recorded.

Statistical evaluation of replicate experiments

Expression level values for individual chips were normalized to the same overall average difference, as described earlier (9). Differentially expressed genes were identified, as described earlier (9), using ANOVA with Kruskal-Wallis test statistics. Genes were considered differentially expressed at a confidence level of 99%, and with a change of at least 2-fold, with a difference in normalized average difference values of at least 100 U. Before calculation of change factors and expression level differences, all values below 20 were set to 20 to eliminate insignificant signals. A previously described B cell data set (9) was reanalyzed using the same criteria. For analysis of constitutively expressed genes, the Affymetrix absolute call, describing whether a probe set is present (P), absent (A), or marginally present (M), was used as follows: first, the letter designations were enumerated so that P became 1, M became 0.5, and A became 0. Next, the average over the five replicate experiments per cellular population exam-

ined was calculated, and a gene was considered as expressed if this average reached at least 0.5. A gene was considered constitutively expressed if the average absolute call reached 0.5 in all cellular populations examined.

Clustering and functional annotation

Self-organizing map (SOM)-based cluster analysis was performed with the program of Tamayo et al. (14): row-wise normalization to a mean of 0 and a SD of 1 across experimental conditions was performed before clustering. Hierarchical clustering was performed with the programs of Eisen et al. (15) using uncentered Pearson's correlation as the similarity metric and average linkage clustering. Before clustering, expression level values were normalized to a mean of 0 and SD of 1 across experimental conditions, separately for the B and T cell data set, using Microsoft Excel. Functional annotations were generated by manual review of article abstracts in the National Library of Medicine's PubMed database about the gene in question or $\sim 100\%$ sequence homologues identified by BLAST in the nucleotide nr database.

RT-PCR

Cellular populations were sorted into PBS, as described above (200,000 cells/population). Cells were pelleted by centrifugation, and RNA was isolated with TRIzol reagent (Invitrogen, San Diego, CA), according to manufacturer's recommendations. Random hexamer-primed cDNA synthesis was performed with Superscript II reverse transcriptase (Invitrogen), according to manufacturer's specifications. RT-PCR was performed in a LightCycler instrument using SYBR green (Roche, Basel, Switzerland) (modifier-1, pre-T α , lymphotoxin- β , *edg-1*, and hypoxanthine phosphoribosyltransferase) or in a Biometra Personal Thermocycler (integrin β_7). Primers were designed using the Primer3 software suite (http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi). Primer sequences were (all 5'→3'): *edg-1*, TTTCCATCGCCATCTCTAC and TTGACGCCCA CATTAACAG; integrin β_7 , TGCCTGTGTATCAGGAGCTG and CTC ACCCTCCGTCTTCTCAG; lymphotoxin- β , TACTACTGTCTGCTGTGC and TCCTGGAAGCATTGGATCTC; modifier-1, GCCCTGACCTTATT GCTGAG and CCTCGTGGCTTTTCTGACTC; pre-T α , ACTGGGTC ATGCTTCTCCAC and ACCAGACAGGGTTGTCAAGG; and hypoxanthine phosphoribosyltransferase, GCTGGTGAAGGACCTCC and CACAGGACTAGAACACCTGC. Annealing temperature was 60°C. Specificity of the product was confirmed by agarose gel electrophoresis or by both agarose gel electrophoresis and melting curve analysis (for LightCycler PCR).

Results

Precursor lymphocyte subpopulations from bone marrow and thymus were separated and purified by FACS, as published previously (9) and as shown in Fig. 2. To study gene expression profiles in murine T cell development, total cellular RNA was extracted from 1.5×10^5 cells of eight consecutive T lymphocyte lineage subpopulations (Fig. 1, *bottom*). mRNA was amplified by two subsequent cycles of cDNA synthesis and in vitro transcription (9, 12, 13). The RNA samples were hybridized to

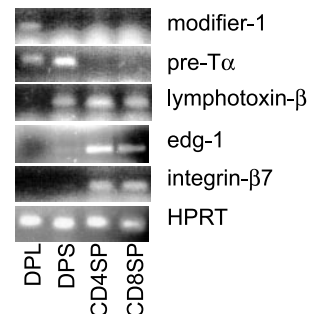


FIGURE 2. Confirmation of gene expression patterns by RT-PCR. Cellular populations were purified, and RNA was prepared, as described in *Materials and Methods*. RT-PCR was performed either on a LightCycler or a conventional thermocycler. For LightCycler PCRs, the specificity of the amplification product was additionally confirmed by DNA melting curve analysis (data not shown).

high density oligonucleotide arrays interrogating 13,026 transcripts (5,231 known genes and 7,795 expressed sequence tags (ESTs)). Five independent replicate experiments were performed, and differential expression was assessed on the basis of statistical algorithms (see *Materials and Methods* for details).

DN1 thymocytes are excluded from the analyses

An initial evaluation of the T cell gene expression data set reveals that the DN1 cell population (R3 gate in Fig. 1B) expresses multiple markers specific for the NK cell, B cell, T cell, and myeloid lineages. This might reflect multilineage gene expression in cells before lineage commitment (16). It is, however, more likely that the DN1 cell population is heterogeneous and might consist of a mixture of different lineage-committed precursors (17). Therefore, we have excluded this population from all following analyses.

Expression patterns of known T cell development-related genes

Table I shows the array-based gene expression values of a series of genes known to be involved in T cell development. CD4 and CD8 show the expected expression patterns, i.e., begin to be detectable in DPL cells and continue to be expressed into the mature, SP cells, in which CD8 is turned off in CD4-positive cells and vice versa.

Of the four TCR chain-encoding loci, TCR γ is expressed highly at the earliest, DN2 stage and is turned off in DPL and all subsequent, TCR β - and TCR α/β -expressing stages. Similarly, TCR δ is expressed highly at the earliest, DN2 stage and turned off a little later than TCR γ , i.e., at the DPS stage of TCR α/β T cell development. TCR β is also expressed from the earliest, DN2 cell stage on. It is not clear, however, why these probe sets do not detect TCR β -chains in CD8 SP T cells. As expected, TCR α begins to be

Table I. Array-based gene expression values of genes with preestablished expression patterns during mouse T cell development

Probe Set ^a	Name ^b	DN2 ^c	DN3	DN4	DPL	DPS	CD4	CD8	p Value (H) ^d	Max Ratio ^e	Max Diff ^f
x04836_s.at	Mouse mRNA for T cell surface glycoprotein CD4 (L3T4)	20	20	160.4	1441.6	2375.4	1001.2	20	0.0022	118.77	2355.4
y00157_s.at	Murine Lyt-2.2 gene for differentiation Ag	20	20	497.4	21547.6	29362.4	647.2	16023.2	0.00049	1468.12	29342.4
M12839_g.at	Mouse TCR germline γ -chain gene C7.5-region	14420.8	14577.2	7994.2	690.8	775.2	1030.4	984	0.00013	21.101911	13886.4
m12815_at	Mouse TCR active γ -chain V10.8B-J-C10.8 mRNA, from cytotoxic T cell line 5/10-13	3683.2	3944.8	1552.6	20	20	20	20	0.00017	197.24	3924.8
m54996_s.at	<i>Mus musculus</i> TCR γ -chain mRNA, complete CDs, clone MNG8	19644.6	25837.6	17084.2	908.4	1030.4	1659.2	1185.2	0.00019	28.442977	24929.2
L36135_i.at	<i>M. musculus</i> germline TCR δ -chain, C region, exon 4	5139.8	4949.2	3682.6	3362.6	966.4	657.8	396.8	0.00073	12.953125	4743
M26053_at	Mouse TCR germline β -chain gene constant region (CT), exon 1	6650	1966.4	1623.4	5400.6	13591	5936.2	276.4	0.00523	49.171491	13314.6
ET63436_at	TCR V β = TCR β -chain variable region (transgenic mice, Genomic, 804 nt)	20	103.4	321.8	803.8	898.4	419.6	20	0.00714	44.92	878.4
ET61537_g.at	<i>M. musculus</i> B10.BR TCR α -chain precursor mRNA, complete CDs	20	20	20	211.2	1693	2439.4	20	0.00205	121.97	2419.4
m16118_at	Mouse mRNA for TCR insulin (A-chain) reactive α -chain VJC, complete CDs	1793.4	1685	3320	6563	17621.4	21263.8	10627.4	0.0001	12.619466	19578.8
u16958_s.at	<i>M. musculus</i> pre-TCR α -type chain precursor mRNA, complete CDs	414.8	1625.6	997	2844.2	1855.4	565.4	216	0.00027	13.167593	2628.2
M29475_s.at	Mouse recombination activating protein (RAG-1) mRNA, complete CDs	1145.2	12454.6	5582.8	14282.2	30140.6	2636	3251.6	0.00006	26.319071	28995.4
M64796_s.at	<i>M. musculus</i> RAG-2 protein (RAG-2) mRNA, complete CDs	20	181.6	20	4629.6	5820.4	20	20	0.00044	291.02	5800.4
x02339_s.at	Mouse mRNA for T3 δ chain of T3/TCR glycoprotein	881.8	10226.4	6703.8	7544.4	13543.8	14939	12508.6	0.00108	16.941483	14057.2
x12729_s.at	Mouse mRNA for TCR δ -chain	5805.8	5266.6	5963	5612	2345.2	1805.4	4889.8	0.00442	3.3028692	4157.6
j02990_s.at	Mouse mRNA for TCR, t3- ϵ glycoprotein	2634	5638.8	4449.8	5073.6	8012.6	7967.6	5525.4	0.02577	3.04198937	5378.6
j04967-3.at	<i>M. musculus</i> TCR ζ -chain (Tcrz) mRNA, complete CDs	1970	2694	3741.4	7232.4	11238.8	9548.8	5194.6	0.00011	5.7049746	9268.8

^a Unique Affymetrix probe set identifier.

^b GenBank description.

^c DN2/DN3/DN4/DPL/DPS/CD4/CD8: average difference value in the respective stage.

^d Value of p from Kruskal-Wallis test.

^e Ratio between the highest and lowest expression level.

^f Difference between the highest and lowest expression level.

expressed in the DPL stage of development and continues to be expressed throughout all further stages of α/β T cell development. Again, it is not clear why one of the two TCR α probe sets on the arrays does not detect TCR α expression in CD8 SP T cells. The TCR ϵ -chain is expressed in all stages examined, but shows an increase in expression level between DN2 and DN3 cells. The δ - and ζ -chains of the CD3 signaling complex are, as expected, expressed from the earliest, DN2 cell stage on throughout development.

Pre-T α is expressed in DN3 cells, appears to be down-regulated in DN4, but up-regulated again in DPL and DPS cells before being turned off in CD4 and CD8 SP cells. Hence, pre-T α might be expressed in some TCR α/β -expressing cells. This would be different from surrogate L chain expression in B cell development, which is turned off in large pre-BII cells and all subsequent stages (18). However, the expression pattern of pre-T α could be confirmed by RT-PCR (Fig. 2).

The RAG genes of the rearrangement machinery show the expected pattern with high expression in DN3 cells, down-regulation in DN4 cells, and peak levels in DP cells.

These results are generally well in accordance with previous studies; Fig. 2 shows an additional confirmation of gene expression patterns by RT-PCR. However, CD44, used for sorting of the cell populations, is not detected by any of the two probe sets on the arrays, presumably due to very inefficient hybridization to the oligonucleotides chosen.

Constitutively expressed genes in B and T cell development

Because B and T cells are closely related cell types, it can be expected that they share a large proportion of the genes that are not developmentally regulated, but constitutively expressed. We therefore identified genes that are detected throughout B and T cell development. A total of 2770 genes are expressed in all T cell developmental stages, and 1830 genes are expressed in all B cell developmental stages. The two lineages share 1612 constitutively expressed genes. Thus, almost 90% of the genes that are expressed in developing B cells are also expressed in developing T cells. This also holds true when comparing individual corresponding developmental stages, e.g., DPL and large pre-BII cells.

Expression patterns of genes differentially expressed during T cell development

A total of 1314 probe sets specific for 535 genes and 779 ESTs are detected as differentially expressed using nonparametric (Kruskal-Wallis) ANOVA at a 99% confidence level. This compares with 1204 differentially expressed probe sets specific for 594 genes and 610 ESTs in murine B cell development using the same threshold criteria. All of these genes are listed in supplementary Table I. Fig. 2a shows expression patterns of coexpressed genes generated by self-organizing maps of differentially expressed genes in T cell development. Twenty different, although sometimes closely related, patterns could be identified. Clusters containing genes preferentially expressed in DN2 (clusters 1, 2), DPL (clusters 10, 11), DPS (clusters 12, 13), CD4 (clusters 15, 6), and CD8 SP (cluster 17) cells can be readily identified. However, most of the genes contained in these clusters are not specifically expressed in one cell population, but show substantial expression levels in more than one stage. No cluster with genes preferentially expressed in DN3 or DN4 cells exists. Genes up-regulated in these stages are also expressed either in DN2 cells (clusters 3, 4, 5) or in all stages up to DPL cells (cluster 6).

Correlations between gene function and expression pattern

The 535 known genes were functionally annotated by review of article abstracts available in the PubMed database. This resulted in putative or definitive annotations for 493 genes; no function could be deduced for 43 genes. Table II correlates the expression pattern from Fig. 3 with this functional annotation. Earlier precursors (DN2 to DPL cells) express genes involved in general metabolism, protein folding, DNA replication, RNA processing, chromatin structure, molecular transport, cell cycle regulation, cytoskeleton structure, and membrane behavior. Signaling molecules, in contrast, appear at similar numbers in most of the expression patterns without clear dominance in either early or late precursors. Genes involved in transcriptional regulation are also represented in most expression patterns, but show a clear peak in DPS cells (eight genes in cluster 0). Later precursors predominantly express genes involved in cytoskeleton modification, intercellular communication, protein degradation, ion transport, and cell-cell contact or adhesion. Thus, in the subsequent stages of T cell development, different functional classes of genes are active in a characteristic sequence.

Most genes involved in cell cycle regulation (78% of cell cycle-related genes differentially expressed in T cell development) and DNA replication (82%) are differentially expressed in both lineages. However, genes involved in intercellular communication (35% for receptors and secreted factors, respectively), general metabolism (39%), signaling (40%), cell-cell contact/adhesion (40%), and transcriptional regulation (52%) are much less coexpressed between the two lineages (Table II).

Key molecular programs responsible for differentiation of precursor cells through different cellular stages are shared between B and T cell development. We thus hypothesize that B and T cell precursors up-regulate a similar set of genes in similar developmental stages. We use the SOM-based clusters shown in Fig. 3 to address this question. For each of the two lineages, we define four pools of functionally related cell stages, i.e., DN2 and DN3 as T pool 1; DN4 and DPL as T pool 2; DPS as T pool 3; and CD4SP and CD8 SP as T pool 4. In B cell development, we define pre-BI cells as B pool 1, large pre-BII cells as B pool 2, small pre-BII and immature B cells as B pool 3, and mature B cells as B pool 4.

Comparison of cellular stages before V to DJ rearrangement and pre-TCR/pre-BCR expression: pool 1

The clusters 1, 2, and 3 of Fig. 3A contain 209 genes up-regulated in DN2 and DN3 cells. Pre-BI-specific genes are contained in clusters 1, 2, 3, and 5 of Fig. 3B, a total of 194 genes. Strikingly, only 18 probe sets are common between these two groups, 10 of which are ESTs (Tables III and IV). Five of the ESTs are now contained in UniGene clusters together with known genes. These genes coexpressed by very early B and T cell precursors stem from a variety of functional classes. Two genes are cell surface receptors involved in intercellular communication. One of those, the G protein-coupled receptor 97, has previously been described as being specifically expressed in early lymphocytes, but shows no phenotype when deleted (19). The other, complement component 1 q-binding protein, is present in virtually all cellular compartments and involved in various ligand-mediated cellular responses (20). Also contained in this group is the suppressor of cytokine signaling 2, inhibitory for the Janus kinase/STAT pathway in a cytokine-specific way. The HES-1 gene is a transcriptional repressor involved in the notch pathway, a key regulator of lymphoid development (21).

A total of 150 genes are up-regulated in DN2 or DN3 cells that are not differentially expressed in B cell development. Similarly,

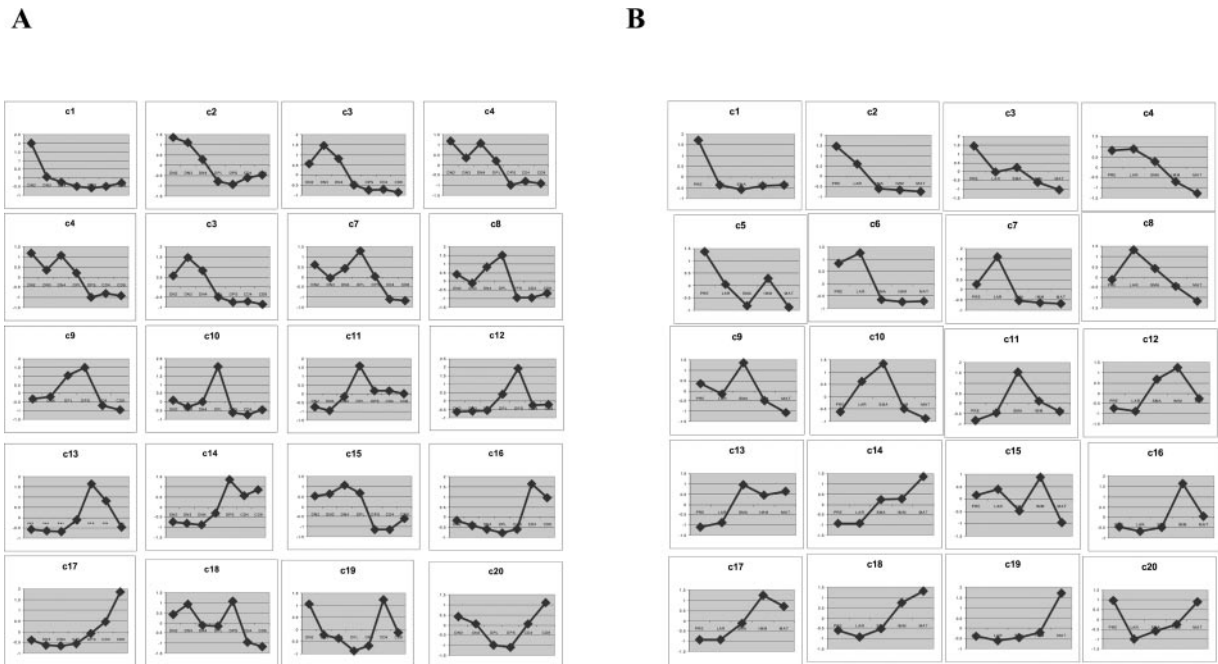


FIGURE 3. Expression patterns of differentially expressed genes in murine T cell development (A) and murine B cell development (B), as identified by self-organizing maps. The cluster number is indicated on the *top* of each diagram. Standardized expression level values are indicated on the y-axis, and developmental stages on the x-axis, from the earliest precursor stages on the *left* to the mature B or T cells on the *right*. The entire list of differentially expressed genes is available as supplementary material.

92 genes are up-regulated in pre-BI cell that are not differentially expressed during T cell development. These groups contain characteristic cell adhesion molecules (carboxypeptidase A and

EGP314 in T cell precursors; glycoprotein-70, α catenin, and P-selectin glycoprotein ligand 1 in B cell precursors), illustrating stage-specific adhesion properties to different microenvironments.

Table II. *Numbers of genes with known functions differentially expressed in the seven developmental stages of T cells with the same patterns of changes^a*

Cluster	Redox Metabolism	Cytochromes	General Metabolism	Chaperones	DNA Replication	RNA Processing	Chromatin Structure	Molecular Transport	Cell Cycle Regulation	Cytoskeleton Components	Membrane Behavior	IG and Signaling	Related
1	1		5						1	2	2	7	
2	1		3	2	1	1		2	1			4	
3	1	1	5	1	1				1			3	
4			10	1	3			1				2	
5			4	3	2					1		1	
6			6	1	9	3	2	3	2			6	
7			1		1		1		1			6	
8				1	10	4	5		3		1	1	
9			1	1	5		1	4	2	1	1		1
10					2		4	1	8	4	3	7	
11							1		1			1	
12			5		1				2	1		6	2
13	1		2				1					7	
14	1					1		1			1	3	
15			2									3	
16			2		1				1	2		6	
17	1		4						2		3	6	
18			1		1				1			3	
19	1											3	
20			5		1				1		1	2	
Total ^b	7	1	56	10	38	9	15	12	27	11	12	77	3
Common ^c	6	1	22	4	31	4	10	6	21	5	6	31	1
% Common ^d	86	100	39	40	82	44	67	50	78	45	50	40	33

Table continues

Similarly, characteristic sets of cell surface receptors (FceRI, IL-3R, Thy-1, endoglin in pre-BI cells, IL-2R, TNFR-1, 4F2 Ag H chain, somatostatin receptor, CTLA-4 in DN2 or DN3 T cell precursors) can be distinguished, indicating that different cytokines are active in the different compartments of early B and T cell development.

Precursor lymphocytes expanding by proliferation after pre-TCR/pre-BCR expression: pool 2

After rearrangement on the Ig H chain and TCRβ locus and surface deposition of a pre-TCR/pre-BCR, cells undergo a pre-BCR- and a pre-TCR-mediated proliferative burst, respectively. In Fig. 3A, the clusters 4, 5, 6, 7, 8, 10, and 11 contain 516 genes up-regulated in DN4 or DPL cells. In Fig. 3B, the clusters 6, 7, and 8 contain 301 genes up-regulated in large pre-BII cells. A total of 161 genes are common between these two groups (Tables III and IV). As expected, many of these genes are involved in DNA replication and cell cycle control as well as chromatin structural proteins and cytoskeleton components. In fact, no known cell cycle-related gene is only expressed in either large pre-BII or DN4/DPL cells. Moreover, large pre-BII and DN4/DPL cells share the expression of many signaling molecules. Although some of these are poorly characterized (fug1, pMELK), many of them are known to transmit cell cycle-related signals (*ect-2*, *sak-a*, *CDC25*, *nek-2*, *stk-1*). Thus, the molecular programs underlying prelymphocyte receptor-mediated proliferative expansion appear to use the same genes in B and T cell precursors.

One possible mechanism for closing of the H chain locus for further rearrangements, i.e., for allelic exclusion is, as mentioned earlier (9), modification of locus accessibility by chromatin reorganization. Earlier studies identified several genes with potential

chromatin-remodeling activity as specifically up-regulated in large pre-BII cells, in which allelic exclusion occurs in the B cell lineage (9). Of these, the modifier-1 RNA is up-regulated not only in large pre-BII, but also in DN4 and DPL cells (confirmed by PCR in Fig. 2), in which allelic exclusion at the TCRβ locus is likely to occur in the T cell lineage. DN4 and DPL cells express additional genes involved in chromosome reorganization (*Cenpc*, *BRCA-2*, and *ZW10*), but these genes have to date not been associated with transcriptional silencing by regulation of accessibility. These expression analyses suggest, therefore, that the modifier-1 gene might be involved in allelic exclusion.

Precursor lymphocytes at the stage of TCRα/IgL chain gene rearrangements: pool 3

After proliferative expansion, precursor B and T cells fall into a resting, noncycling state and are induced to the second wave of rearrangements at the Ig L chain/TCRα loci. A total of 200 genes in clusters 12, 13, and 14 are up-regulated in DPS cells (Fig. 3A), while 325 genes in clusters 9, 10, 11, 12, 16, and 17 are up-regulated in either small pre-BII cells or immature B cells (Fig. 2B). Twenty-three genes form the common subset (Tables III and IV). Among those, two have antiapoptotic activity (*EAT/MCL-1*, *bcl-x*). *TIS-21* is known for its cell cycle-repressing activity (22). The only signaling molecule, inositol-1, 4, 5-triphosphate receptor, is a key transmitter of Ag receptor-derived signals (23), pointing to a possible common function in one of the steps of repertoire selection and/or maturation after completion of Ig L chain/TCRα rearrangements. Interestingly, one EST is homologous to and contained in one UniGene cluster with the newly described DNA polymerase, a TdT homologue producing random

Table II. Continued

Transcriptional Regulation	TCR	Extracellular Matrix	Cytoskeleton Modification	Intercellular Communication	Protein Degradation	MHC	Ion Transport	Unknown	VDJ-Recombination	Apoptosis	Cell-Cell Contact/Adhesion	Anti-microbial	Total
3			1		3			1		2	1		29
2	3			8		1		4					33
3	1			2	1			1		1	1		23
2	1							1		2			23
4			1					2			2		20
5	1			1	2			4		1	1		47
1					2			3					16
4						1							30
4	2			1			1	1	1				27
3				1			1	3		1			38
				1	1			3			1		10
1		1		5	3	1	1		1	2	1		33
8	2			3		1		3		1	2		31
3		1	1		2		1	4			1		20
5	4		2	4	1	1					2		25
5		1	1	8	4	11	2	4		1	2		51
3				5		2	1	6		2	1		36
1				2				1	1				11
2				3							3	1	13
1				2	1	1		2			2		19
60	14	3	6	46	20	20	8	43	3	13	20	1	535
31	0	0	3	16	7	6	4	18	3	4	8	0	248
52	0	0	50	35	35	30	50	42	100	31	40	0	

^a The clusters of genes with shared expression patterns identified in the analysis shown in Fig. 3A 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 function groups appear in columns. The numbers of genes identified for each functional group in each cluster are given in bold.

^b Row- and column-wise totals, respectively.

^c Number of genes in the respective functional group that are differentially expressed in both B and T cell development.

^d Percentage of genes differentially expressed in T cell development that are also differentially expressed in B cell development (“Common” divided by “Total” multiplied by 100).

Table III. *Genes coexpressed between B and T cell precursors at similar developmental stages^a*

Probe Set ^b	Description ^c	T-Cell Cluster ^d	B-Cell Cluster ^e	UniGene ^f
aa044561_s.at	EST	10	18	
AA271402_at	EST	10	18	G protein-coupled receptor 97
aa561252_s.at	EST	10	15	
Msa.24217.0_f.at	EST	10	19	
U88327_s.at	<i>M. musculus</i> suppressor of cytokine signalling-2 (SOCS-2) mRNA, complete CDs	10	18	
w82380_s.at	EST	10	15	Cyclin-dependent kinase inhibitor 1A (P21)
M13018_f.at	Mouse cysteine-rich intestinal protein (CRIP) mRNA, complete CDs	11	18	
Msa.18493.0_f.at	EST	11	18	
Msa.554.0_f.at	Mouse CRIP mRNA, complete CDs	11	18	
u12273_s.at	<i>M. musculus</i> apurinic/aprimidinic endonuclease (Apex) gene, complete CDs	11	14	
X70100_f.at	<i>M. musculus</i> mal1 mRNA for keratinocyte lipid-binding protein	11	19	
Z22593_s.at	<i>M. musculus</i> fibrillarlin mRNA	11	18	
AA217762_s.at	EST	15	15	
aa544043_s.at	EST	15	15	Acid phosphatase 6, lysophosphatidic
D16464_s.at	Mouse helix-loop-helix factor HES-1 gene	15	14	
Msa.1561.0_f.at	<i>M. musculus</i> thioredoxin-dependent peroxidase reductase (tpx) mRNA, complete CDs	15	19	
Msa.5443.0_s.at	EST	15	18	Complement component 1, q subcomponent binding protein
u53547_s.at	MMU53547 mouse thymus expressed sequences <i>M. musculus</i> cDNA, mRNA sequence	15	15	Uridine-cytidine kinase 2
U90123_s.at	<i>M. musculus</i> HN1 (Hn1) mRNA, complete CDs	12	7	
L00606_at	<i>M. musculus</i> MHC class I Qa-1a antigen mRNA, complete CDs	12	6	
u20497_s.at	<i>M. musculus</i> p19 protein mRNA, complete CDs	12	7	
Msa.22043.0_f.at	EST	12	11	
M27844_f.at	<i>M. musculus</i> calmodulin synthesis (CaM) cDNA, complete CDs	13	7	
U39302_s.at	<i>M. musculus</i> 26S proteasome subunit 4 ATPase mRNA, complete CDs	13	11	
U42190_s.at	<i>M. musculus</i> G/T-mismatch binding protein (Gtmbp) mRNA, complete CDs	13	11	
Msa.2456.0_at	Mouse mRNA for modifier 1 protein	13	7	
U63337_s.at	<i>M. musculus</i> cyclin-dependent kinase-2 α (Cdk2- α) mRNA, complete CDs	13	11	
Msa.16618.0_s.at	EST	13	7	
aa177891_at	EST	13	11	
Msa.3318.0_at	Mouse mRNA encoding DNA (cytosine-5)-methyltransferase (EC 2.1.1.37)	14	11	
Y15522_s.at	<i>M. musculus</i> mRNA for MNUDC protein	14	11	
U27014_s.at	<i>M. musculus</i> sorbitol dehydrogenase precursor mRNA, partial CDs	14	6	
L32836_f.at	<i>M. musculus</i> (clone C7/B9) S-adenosyl homocysteine hydrolase (ahcy) mRNA, complete CDs	14	11	
Msa.1580.0_f.at	<i>M. musculus</i> (clone C7/B9) S-adenosyl homocysteine hydrolase (ahcy) mRNA, complete CDs	14	11	
Msa.3027.0_s.at	<i>M. musculus</i> lymphocyte specific helicase mRNA, complete CDs	14	11	
X62154_s.at	<i>M. musculus</i> mRNA for P1 protein (P1.m)	14	11	
AA124895_s.at	EST	14	6	
Msa.23515.0_f.at	EST	14	7	
aa407907_rc.at	EST	14	11	
aa409333_rc_s.at	EST	14	11	
L21973_s.at	EST	14	11	
Msa.41695.0_f.at	EST	14	11	
U69488_s.at	<i>M. musculus</i> viral envelope like protein (G7e) gene, complete CDs	16	6	
U13262_s.at	<i>M. musculus</i> myelin gene expression factor (MEF-2) mRNA, partial CDs	16	7	
U51037_s.at	<i>M. musculus</i> 11-zinc-finger transcription factor (CTCF) mRNA, complete CDs	16	7	
L11316_s.at	Mouse oncogene (ect2) mRNA, complete CDs	16	7	
I29479_s.at	<i>M. musculus</i> serine/threonine kinase (sak-a) mRNA, complete CDs	16	7	

(Table continues)

Table III. *Continued*

Probe Set ^b	Description ^c	T-Cell Cluster ^d	B-Cell Cluster ^e	UniGene ^f
u08110_s.at	<i>M. musculus</i> RNA1 homolog (Fug1) mRNA, complete CDs	16	7	
u15562_s.at	<i>M. musculus</i> CDC25 (Cdc25) mRNA, complete CDs	16	7	
u80932_s.at	<i>M. musculus</i> serine/threonine kinase Ayk1 (ayk1) mRNA, complete CDs	16	7	
u95610_s.at	<i>M. musculus</i> nimA-related kinase 2 (Nek2) mRNA, complete CDs	16	7	
x95351_s.at	<i>M. musculus</i> mRNA for pMELK protein	16	7	
D12646_s.at	Mouse kif4 mRNA for microtubule-based motor protein KIF4, complete CDs	16	7	
D49544_s.at	Mouse mRNA for KIFC1, complete CDs	16	7	
d55720_s.at	Mouse mRNA for nuclear pore-targeting complex, complete CDs	16	7	
Y09632_s.at	<i>M. musculus</i> mRNA for rabkinesin-6	16	7	
d12513_f.at	Mouse mRNA for DNA topoisomerase II, complete CDs	16	7	
Msa.47.0_f.at	<i>M. musculus</i> DBA/2 DNA topoisomerase α fusion protein mRNA, partial CDs	16	7	
Msa.1263.0_f.at	Mouse mRNA for β -tubulin (isotype M β 5)	16	7	
x04663_f.at	Mouse mRNA for β -tubulin (isotype M β 5)	16	7	
m35153_s.at	Mouse lamin B mRNA, complete CDs	16	11	
AF012709_s.at	<i>M. musculus</i> centromere protein A (Cenp-a) mRNA, complete CDs	16	7	
X58069_s.at	Mouse mRNA for histone H2A.X	16	7	
af002823_s.at	<i>M. musculus</i> mitotic checkpoint protein kinase (Bub1) mRNA, complete CDs	16	7	
M86377_s.at	Mouse esk kinase mRNA, complete CDs	16	7	
U19596_s.at	<i>M. musculus</i> Cdk4 and Cdk6 inhibitor p18 protein mRNA, complete CDs	16	7	
X58708_g.at	Mouse cycB mRNA for cyclin B	16	7	
X64713_s.at	<i>M. musculus</i> mRNA for cyclin B1	16	7	
X66032_s.at	<i>M. musculus</i> mRNA for cyclin B2	16	7	
u27177_s.at	<i>M. musculus</i> p107 (p107) mRNA, complete CDs	16	11	
X82786_s.at	<i>M. musculus</i> mRNA for Ki-67	16	11	
aa049623_s.at	EST	16	7	
aa139030_s.at	EST	16	7	
aa183033_s.at	EST	16	7	
aa184798.at	EST	16	7	
aa189313_s.at	EST	16	7	
aa214783_s.at	EST	16	7	
aa218046_s.at	EST	16	7	
aa222854_s.at	EST	16	7	
aa268341_s.at	EST	16	7	
aa285607_s.at	EST	16	7	
aa289122_s.at	EST	16	7	
aa407201_rc_s.at	EST	16	7	
aa408511_rc.at	EST	16	7	
aa590750.at	EST	16	7	
aa590750_g.at	EST	16	7	
aa592163_s.at	EST	16	7	
AA673431_rc_f.at	EST	16	7	
aa690055.at	EST	16	7	
AC002393_geneC8_f.at	EST	16	7	
c77497_rc.at	EST	16	7	
C78101_rc.at	EST	16	7	
C78640_rc.at	EST	16	7	
c78700_rc.at	EST	16	7	
c79407_rc.at	EST	16	7	
c80926_rc.at	EST	16	7	
c80926_rc_g.at	EST	16	7	
Msa.11185.0_f.at	EST	16	7	
Msa.15090.0_s.at	EST	16	7	
Msa.15849.0_s.at	EST	16	7	
Msa.17532.0_f.at	EST	16	7	
Msa.18569.0_s.at	EST	16	7	
Msa.33600.0_f.at	EST	16	7	
Msa.4014.0_f.at	EST	16	7	
Msa.6932.0_f.at	EST	16	7	
aa182195.at	EST	16	11	
z31235_f.at	EST	16	11	

(Table continues)

Table III. *Continued*

Probe Set ^b	Description ^c	T-Cell Cluster ^d	B-Cell Cluster ^e	UniGene ^f
U52951_s.at	<i>M. musculus</i> putative transcriptional regulator mEnx-1 mRNA, complete CDs	17	7	
Z46757_s.at	<i>M. musculus</i> mRNA for high mobility group 2 protein	17	7	
D21099_s.at	Mouse mRNA for STK-1 (serine/threonine kinase), complete CDs	17	7	
X97796_s.at	<i>M. musculus</i> mRNA homologous to <i>S. cerevisiae</i> RAD54	17	7	
af004105_s.at	<i>M. musculus</i> BM28 homolog mRNA, complete CDs	17	11	
d17384_s.at	Mouse mRNA for DNA polymerase alpha catalytic subunit p180	17	11	
D26091_s.at	House mouse; <i>Musculus domesticus</i> female mammary carcinoma mRNA for mCDC47, complete CDs	17	11	
K02927_s.at	Mouse ribonucleotide reductase subunit M1 mRNA, complete CDs	17	11	
m14223_s.at	Mouse ribonucleotide reductase M2 subunit mRNA, complete CDs	17	11	
u36475_s.at	<i>M. musculus</i> breast and ovarian cancer susceptibility protein (Brca1) mRNA, complete CDs	17	11	
U42385_s.at	<i>M. musculus</i> fibroblast growth factor inducible gene 16 (FIN16) mRNA, complete CDs	17	7	
Msa.2344.0.f.at	Mouse mRNA for HMG-17 chromosomal protein	17	11	
X12944_f.at	Mouse mRNA for HMG-17 chromosomal protein	17	11	
U08215_s.at	<i>M. musculus</i> Hsp70-related NST-1 (hsr.1) mRNA, complete CDs	17	7	
ET61411_s.at	<i>M. musculus</i> thymopoietin γ mRNA, complete CDs	17	7	
z26580_s.at	<i>M. musculus</i> mRNA for cyclin A	17	7	
X75888_s.at	<i>M. musculus</i> mRNA for cyclin E	17	11	
aa175756_s.at	EST	17	7	
aa189300_s.at	EST	17	7	
aa189902_s.at	EST	17	7	
AA261527_at	EST	17	7	
aa266783_s.at	EST	17	7	
aa267955_s.at	EST	17	7	
aa269806_s.at	EST	17	7	
aa407116_rc_s.at	EST	17	7	
aa409629_rc_at	EST	17	7	
aa537407_at	EST	17	7	
aa545124_at	EST	17	7	
aa673484_rc_s.at	EST	17	7	
C77864_rc_f.at	EST	17	7	
R75011_rc_s.at	EST	17	7	
Msa.33213.0.f.at	EST	17	7	
Z31300_s.at	EST	17	7	
aa177744_at	EST	17	11	
aa200689_at	EST	17	11	
aa200970_s.at	EST	17	11	
aa238367_at	EST	17	11	
aa267296_s.at	EST	17	11	
aa407737_s.at	EST	17	11	
AA426917_s.at	EST	17	11	
aa545124_g.at	EST	17	11	
aa561108_at	EST	17	11	
AA673176_rc_f.at	EST	17	11	
C80742_rc_s.at	EST	17	11	
C81007_rc_f.at	EST	17	11	
Msa.23918.0_s.at	EST	17	11	
U89876_s.at	<i>M. musculus</i> transcriptional coactivator ALY (ALY) mRNA, complete CDs	18	7	
AF016583_g.at	<i>M. musculus</i> checkpoint kinase Chk1 (Chk1) mRNA, complete CDs	18	11	
x56045_s.at	Mouse mRNA (clone λ -19) for hypothetical protein A	18	11	
M58558_s.at	Murine sm D small nuclear ribonucleoprotein sequence	18	11	

(Table continues)

Table III. *Continued*

Probe Set ^b	Description ^c	T-Cell Cluster ^d	B-Cell Cluster ^e	UniGene ^f
U75680_s.at	<i>M. musculus</i> histone stem-loop binding protein (SLBP) mRNA, complete CDs	18	11	
j04627_s.at	Mouse NAD-dependent methylenetetrahydrofolate dehydrogenase-methenyltetrahydrofolate cyclohydrolase mRNA, complete CDs	18	11	
D00812_s.at	<i>M. musculus</i> mRNA for 30-kDa subunit of replication protein A, complete CDs	18	11	
d13544_s.at	Mouse mRNA for primase small subunit, complete CDs	18	11	
D26089_s.at	Mouse mRNA for mcdc21 protein, complete CDs	18	11	
D26090_s.at	Mouse mRNA for mCDC46 protein, complete CDs	18	11	
X53068_s.at	Mouse mRNA for proliferating cell nuclear Ag	18	11	
AF034610_s.at	<i>M. musculus</i> nuclear autoantigenic sperm protein mRNA, complete CDs	18	11	
u58633_s.at	<i>M. musculus</i> p34 cdc2 kinase mRNA, complete CDs	18	7	
c79202_rc.at	EST	18	6	
aa050541_s.at	EST	18	7	
aa175185_s.at	EST	18	7	
aa204524_s.at	EST	18	7	
aa475191_s.at	EST	18	7	
aa259483.at	EST	18	11	
aa614954.at	EST	18	11	
aa689977_f.at	EST	18	11	
c78306_rc.at	EST	18	11	
Msa.21737.0_f.at	EST	18	11	
aa237376_s.at	EST	19	11	
c77199_rc.at	EST	19	11	
C81593_rc_f.at	EST	19	11	
U35623_s.at	<i>M. musculus</i> EAT/MCL-1 mRNA, complete CDs	0	8	
M63695_s.at	Mouse CD1.1 mRNA, complete CDs	0	0	
Msa.2938.0_s.at	<i>M. musculus</i> mRNA for selenoprotein P	0	1	
Msa.1690.0_g.at	<i>M. musculus</i> I- κ B α chain mRNA, complete CDs	0	0	
U17259_s.at	<i>M. musculus</i> p19 mRNA, complete CDs	0	1	
aa189937_s.at	EST	0	9	Semaphorin G
u78031_s.at	<i>M. musculus</i> apoptosis inhibitor bcl-x (bcl-x) gene, exon 3 and complete CDs	4	2	
M64292_s.at	Mouse TIS21 gene, complete CDs	4	0	
X03690_s.at	<i>M. musculus</i> mRNA for Ig heavy chain constant region mu(b) allele	4	5	
Msa.2935.0_s.at	<i>M. musculus</i> mRNA for inositol 1,4,5-trisphosphate receptor (type 2)	4	0	
M29475_s.at	Mouse recombination activating protein (RAG-1) mRNA, complete CDs	4	1	
aa288448_g.at	EST	4	8	
aa408085_rc_s.at	EST	4	0	
aa529960_s.at	EST	4	0	
aa673574.at	EST	4	5	
C80574_rc_s.at	EST	4	2	
R74735_rc_s.at	EST	4	0	
Msa.16595.0_s.at	EST	4	1	
Msa.26238.0_s.at	EST	4	2	
m89956_s.at	<i>M. musculus</i> calcium-binding protein (pp52) mRNA, complete CDs	5	8	
aa185666_s.at	EST	5	8	
aa290454_s.at	EST	5	0	DNA polymerase mu
C77346_rc.at	EST	5	9	
x73359_s.at	<i>M. musculus</i> mAES-1 mRNA	1	16	
u21795_g.at	<i>M. musculus</i> common cytokine receptor γ -chain gene, complete CDs	1	13	
u16985_s.at	<i>M. musculus</i> lymphotoxin- β mRNA, complete CDs	1	16	
aa184116_s.at	EST	1	13	
aa638800_s.at	EST	1	13	α -adducin

(Table continues)

Table III. *Continued*

Probe Set ^b	Description ^c	T-Cell Cluster ^d	B-Cell Cluster ^e	UniGene ^f
C78154_rc.at	EST	1	13	Ral guanine nucleotide dissociation stimulator,-like 2
aa185262_s.at	EST	1	16	
aa189785_s.at	EST	1	16	
aa289470_s.at	EST	1	16	Actin-cross-linking protein 7
aa407468_rc.s.at	EST	1	16	
m59821_s.at	Mouse growth factor-inducible protein (pip92) mRNA, complete CDs	2	13	
U20238_s.at	<i>M. musculus</i> GTPase-activating protein GAPIII mRNA, complete CDs	2	16	
u89269_s.at	<i>M. musculus</i> preprodiptidyl peptidase I mRNA, complete CDs	2	13	
d87910_s.at	<i>M. musculus</i> mRNA for PA28 β subunit, complete CDs	2	13	
U40811_s.at	<i>M. musculus</i> orphan G-protein-coupled receptor (edg-1) gene, complete CDs	2	13	
u05837_s.at	<i>M. musculus</i> mouse β -N-acetylhexosaminidase α -subunit (Hexa) gene, exon 14 and partial CDs	2	13	
m95632_s.at	Mouse integrin β_7 subunit mRNA, complete CDs	2	16	
Msa.463.0.at	Mouse bcl-2 gene encoding mbcl-2- β	2	16	
AA170444.at	EST	2	13	
aa189534_f.at	EST	2	13	
aa255186_s.at	EST	2	13	Cathepsin S
aa275273_f.at	EST	2	13	
aa175794_s.at	EST	2	16	
aa177433_s.at	EST	2	16	
AA190042_s.at	EST	2	16	
AA266394.at	EST	2	16	Oxysterol binding protein 2
X13460_s.at	Mouse mRNA for p68 protein of the lipocortin family	3	13	
u18975_f.at	<i>M. musculus</i> β -1,4 N-acetylgalactosaminyltransferase (Ggm2) mRNA, complete CDs	3	16	
aa137972_s.at	EST	3	13	
AA178227.at	EST	3	13	CD83
Msa.8157.0_s.at	EST	3	13	Cathepsin S
aa178516_s.at	EST	3	16	
aa408820_rc.at	EST	3	16	
c79715_rc.at	EST	3	16	
w81812_s.at	EST	3	16	

^a Shown are genes up-regulated of both B and T cell precursors at corresponding developmental stages, i.e., in DN2/DN3 and pre-BI cells, in DN4/DPL and large pre-BII cells, in DPS and small pre-BII/immature B cells, and in mature B cells and CD4/CD8 SP cells.

^b Unique Affymetrix identifier.

^c GenBank header.

^d SOM cluster number from Fig. 3A.

^e SOM cluster number from Fig. 3B.

^f Name of full-length gene contained in the same UniGene cluster as the respective EST.

insertions of nucleotides (24). One might speculate about a possible role of this polymerase in recombination events at the TCR and/or in Ig gene loci.

Among the 142 known genes that are specifically up-regulated in DPS cells and the 213 known genes that are specifically up-regulated in small pre-BII or immature B cells, one can, again, find

Table IV. *Total numbers of differentially expressed genes up-regulated at comparable cellular stages and number of genes common to both T- and B-lineage cells at the same state*

Cell Differentiation Stage		Genes Up-regulated in T Cells ^a	Genes Up-regulated in B Cells ^b	Common Genes ^c	% Common in T Cells ^d	% Common in B Cells ^d
T Cell development	B Cell development					
DN2, DN3	Pre BI	209	194	18	9	9
DN4, DPL	Large Pre-BII	516	301	161	31	53
DPS	Small Pre-BII, Immature B	200	325	23	11	7
CD4 ⁺ , CD8 ⁺	Mature B	245	168	35	14	21

^a Number of genes with peak expression in the T cell developmental stages indicated in the first column as determined by SOM-based cluster analysis in Fig. 3A.

^b Number of genes with peak expression in the B cell developmental stages indicated in the second column as determined by SOM-based cluster analysis in Fig. 3B.

^c Number of genes up-regulated by corresponding stages of B and T cell development (common subset between columns 3 and 4).

^d % common in B/T cells: percentage of the genes up-regulated by corresponding stages of both B and T cell development from the total number of genes up-regulated in the respective stages of B and T cell development.

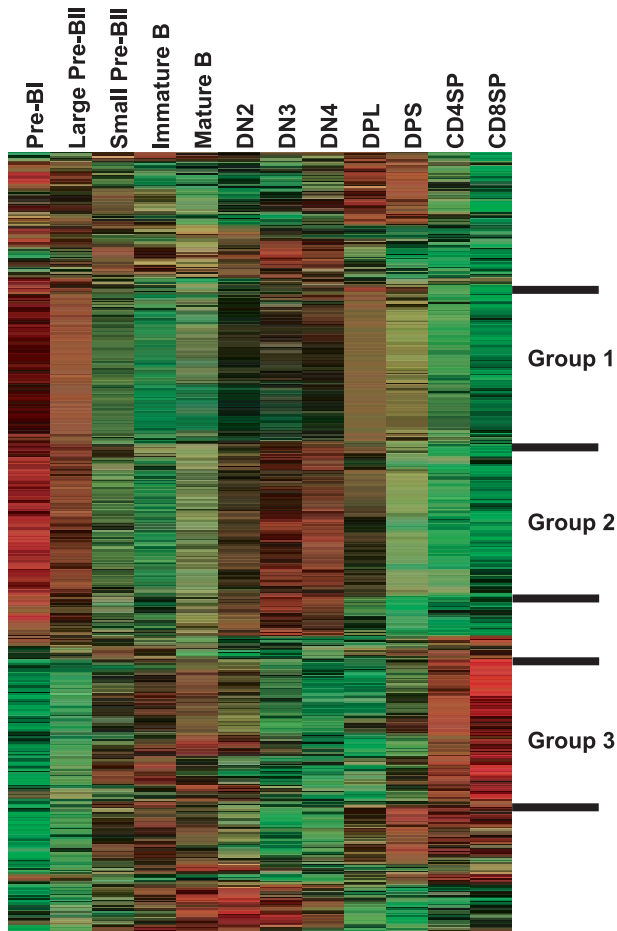


FIGURE 4. Expression patterns of 546 genes differentially expressed in both T and B cell development. For better visualization of global patterns, hierarchical clustering was performed according to the method of Eisen et al. (15). Genes were normalized to a mean of 0 and SD of 1 across experimental conditions, separately for B and T cell precursors, before clustering. Red indicates expression level above; green indicates expression level below the row-wise mean, respectively. The entire list of genes is available as supplementary material.

characteristic signatures of gene expression. Of the genes that have been described as potentially involved in cell cycle arrest in the B cell lineage, only EAT/MCL-1 is up-regulated in the respective stage of T cell development, while the other (MyD118) is not. If we assume that some of the molecular programs underlying cell cycle arrest are similar in the two lineages, this points to an important role of EAT/MCL-1.

A large number of cell surface receptors and signaling molecules are up-regulated in T pool 3 cells on the one hand, and in B pool 3 cells in contrast, while almost none are shared in expression between the B and the T cell pool. This indicates that the signaling capacities in these cells increase, but the extracellular stimuli and intracellular signal transduction pathways appear to be different and unique for B and T cell precursor stages.

Mature lymphocytes: pool 4

Regarding mature T and B cells, 245 genes are up-regulated in CD4 or CD8 SP (T pool 4) cells (clusters 15, 16, and 17 in Fig. 3A), whereas 168 genes are up-regulated in mature B (B pool 4) cells (clusters 18 and 19 in Fig. 3B). Of these, 35 known genes and ESTs (Tables III and IV) are expressed both in the B and in the T cell lineage. Among those, the known expression pattern of *bcl-2*

is confirmed (25). Cell surface molecules up-regulated by mature lymphocytes of both lineages include the common cytokine receptor γ -chain, *edg-1* (previously associated with vascular maturation) (26), lymphotoxin β , and integrin β_7 (see also PCR confirmation in Fig. 2). However, much like in small pre-BII and DPS cells, the vast majority of cell surface receptors and signaling molecules are not shared between mature B and T cells. It is well known, however, that very similar signal transducers are downstream of the Ag receptors of B and T cells, respectively. A thorough analysis of expression patterns of signal-transducing molecules downstream of the B cell Ag receptor revealed that these genes are generally not differentially expressed during development (data not shown). Thus, they are likely not to appear in the lists of genes described in this work, because we focus exclusively on developmentally regulated genes.

Expression patterns of genes shared between B and T cell development

These analyses show that corresponding stages of B and T cell development generally coexpress only very small numbers of genes, with the exception of many cell cycle-related genes jointly expressed by large pre-BII and DPL cells. In contrast, 546 of the genes investigated in this study are differentially expressed during both B and T cell development. We have used hierarchical clustering of these genes to determine their expression patterns (Fig. 4). Apparent is the large cell cycle-related cluster mentioned above as expressed in T/B pool 2 cells (group 1 in Fig. 4). Besides, a second cluster exists with genes up-regulated in pre-BI and large pre-BII cells as well as in all T cell stages up to DPL cells (group 2 in Fig. 4). A third large group consists of genes up-regulated in immature and mature B cells as well as in CD4⁺ or CD8⁺ T cells (group 3 in Fig. 4). A complete list of these genes is available as supplementary Table II.

In conclusion, comparable stages of T and B cell development at different points in their lineage differentiation share different proportions of differentially expressed genes. Only ~10% are shared between T/B pool 1 cells, as well as between T/B pool 3 cells (Table IV). Both pools express several cell communication and signaling molecules, indicating that these programs individually differ between B and T lineage cells. In contrast, T/B pool 2 cells share one-third to one-half of all differentially expressed genes (Table IV), and the majority of them have functions in dividing cells, for example DNA replication and cell cycle control. These genes should be shared by all dividing cells regardless of the differentiation lineage. Altogether, it is surprising how low the number of genes is that appear to be shared between comparable stages of B and T cell development.

Discussion

This work describes gene expression patterns in mouse thymocyte precursor populations. These expression patterns are compared with corresponding patterns in B cell development by defining four pools of functionally related cell stages: pool 1, cells before V to DJ rearrangements (DN2/DN3 and pre-BI cells); pool 2, cells undergoing pre-BCR/pre-TCR-mediated proliferative expansion (DN4/DPL and large pre-BII cells); pool 3, precursor lymphocytes at the stage of TCR α /IgL chain gene rearrangements (DPS and small pre-BII/immature B cells); and pool 4, mature lymphocytes.

One mouse has between 5×10^5 (pool 2) and 5×10^6 - 10^7 (pools 1, 3, and 4) cells in its primary lymphoid organs. Therefore, it needed RNA amplification techniques developed in our laboratory to analyze RNA expression patterns from as few as 1.5×10^5 cells. We have been able to confirm previously known expression patterns of genes that are involved in lymphoid development.

Thus, we are confident that this method allows the expected qualitative and semiquantitative detection of differentially expressed genes (Table I). Moreover, the cluster analysis used in this study groups together genes in functionally relevant ways. One particularly clear example is the detection of genes involved in DNA replication, cell cycle control, and cell division in the T and B cell pools 2, which contain, to a large majority, cells proliferating in response to the deposition of prelymphocyte receptors in their surface membrane.

It is clear that all the pools examined are inherently heterogeneous. In T pool 1, two populations separable by surface marker expression are combined, while B pool 1 is known to be composed in part of resting and in part of cycling cells (27). T pool 2 is a collection of cells distinguishable by expression levels of CD4 and CD8, while cells of B pool 2 (large pre-BII cells) are known to be a mixture of cells still expressing the surrogate L genes, but not yet transcribing the IgL chain gene loci, and others that have terminated surrogate L chain expression, but have begun sterile transcription from the L chain gene loci (27). T pool 3 is expected to be a mixture of DPS cells not yet expressing the TCR α/β molecules and others that do express this Ag-specific receptor, just as B pool 3 is a deliberate mixture of surface IgM-negative pre-BII cells and surface IgM-positive immature B cells. Further heterogeneity in these pool 3 cells can be expected in their gene expression programs because some of the Ag receptor-expressing cells are on their way to death by negative selection, others try to avoid this by receptor editing with secondary rearrangements, still others are in the process of positive selection, and yet others are ignored because their Ag-specific receptors do not find a suitable autoantigen in the primary lymphoid organs (7). Finally, T pool 4 is a deliberate combination of CD4 SP and CD8 SP cells, while B pool 4 might be expected to contain conventional as well as B1-type mature B cells. It should also not be forgotten that in T cell development, ~3% of all cells in the thymus are TCR γ/δ -expressing T cells. Therefore, it is obvious that our analysis can only be regarded as a first, relatively crude overall picture of gene expression programs operative in B and T cell development. This should improve once novel markers for an even better separation of functionally distinct stages become available.

In the statistical analysis of these oligonucleotide array data, we have used a robust statistical test for detection of differentially expressed genes. It should be noted, however, that several options for the analysis of oligonucleotide array data exist, with impact on the genes that are subsequently detected as differentially expressed (28). To make comparisons between the different sets of data described in this and our earlier studies possible, we decided to use the same algorithms in each of these studies. The complete data set is, however, available at the Gene Expression Omnibus of the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/geo/>), and can therefore be subjected to analyses with other algorithms.

An earlier study indicated that B and T cells differ in only 2% of expressed genes (29). To reinvestigate this figure, we have analyzed constitutively expressed genes in our data set. First, more genes appear to be expressed in T cells than in B cells, either when analyzing individual stages or when looking at genes constitutively expressed throughout the entire developmental pathway. This might be due to a higher sensitivity of the T cell experiment due to a change in the manufacturing process of the arrays. However, this still holds true when the data set is analyzed with sophisticated statistical algorithms that are designed for correction of such differences (data not shown) (30, 31). Moreover, the absolute call parameter is designed for maximum specificity, and not necessarily for maximum sensitivity. It might thus be that some genes

scored as absent in the array-based gene expression analysis can be detected by PCR or some other, more sensitive procedure. In view of these limitations, our finding that ~90% of the genes constitutively expressed in B cell development are also constitutively expressed in T cell development is reasonably well in agreement with this earlier study.

Approximately 10% of all the 13,000 genes present on the arrays change their expression during subsequent stages of development. Between 30 and 86% of the genes in one functional group share expression in T and B lineage cells. Three classes of functional groups are distinguishable based on the coexpression between the two lymphoid developmental lineages: those that are coexpressed to <40%, those that are coexpressed to ~50%, and those that are coexpressed to >75% (Table II). It becomes evident that both B and T lineage cells, and maybe other mouse cell lineages, too, use the same machinery to control DNA replication and cell cycle. In contrast, intracellular signaling pathways, cell surface receptors, and secreted factors such as cytokines are distinct between T and B lineage cells.

Furthermore, it is evident that even of those genes that are shared between B and T lineage cells, an even smaller proportion is shared between comparable stages of B and T cell development. This is particularly evident in pools 1, 3, and 4, in which only ~10–20% of the genes are shared (Table IV). Notably, this is despite the fact that these pools consist of markedly heterogeneous cell populations, and the number of shared genes might decrease even further when functionally homogeneous subpopulations of the pools described in this work are examined. This small number of genes shared at comparable stages of development may suggest that the same gene products are used at different stages of cellular development to convey different signaling qualities to a given set of cells during their developmental pathways.

Many genes included in this analysis peak in expression in the stage under consideration, but are also expressed at substantial levels in developmentally adjacent stages (Fig. 3). Only in rare cases is a cluster associated specifically with only one stage of development. Most of the expression patterns displayed in these clusters indicate that sets of genes can be combined in their expression behavior so that expression may be highest in one given cellular stage, but that adjacent or even nonadjacent stages (clusters 17, 19, and 20 in Fig. 3A) also express these genes. Particularly the bi- or multimodal expression patterns suggest that such genes could be used for different functional purposes in different cellular stages. In all these considerations, it should be clear that we cannot extrapolate these RNA expression patterns into translational expression of proteins and their subsequent posttranslational modifications.

Nevertheless, these gene expression analyses should be useful because they identify stages of cellular development by a more complete phenotype than the presently used markers. They set the stage to study changes in gene expression upon stimulation and other external influences on these cells. Another useful expansion of these expression analyses could be to compare these normal stages of lymphoid cells with leukemias and other neoplasms transformed to malignancy at these stages of development (32). Moreover, comparisons with other, more distantly related cellular differentiation programs might help to find those groups of genes with shared functions that have been preserved throughout evolution as common important programs in many, or even all, cellular developmental processes.

Despite the limitations of the current analyses of gene expression programs, and despite the limitations to separate functionally homogeneous subsets of cells for our analyses, for the first time we were able to compare the global gene expression program of two

closely related cellular differentiation lineages in their primary organs of generation. To date, no other cell developmental pathways of mice or humans have been studied with so many genes, estimated to represent at least one-third of the genome. Simple general rules for gene usage emerge: 1) proliferation of cells at comparable stages is executed by the same genes; 2) differentiation and maturation of cells, mediated by intracellular signaling, cell adhesion, and intercellular communication, are in majority executed by different sets of genes; and 3) most genes are not used strictly at comparable, but rather at several, stages, possibly in different functional contexts.

Acknowledgments

We are indebted to Tracy Haydn and Hubertus Kohler for operation of the cell sorter facility, to Ulli Certa for access to the Affymetrix machinery, and to Martin Neeb and Clemens Broger for programming of the analysis software.

References

- Melchers, F., and A. Rolink. 1999. B-lymphocyte development and biology. In *Fundamental Immunology*. W. E. Paul, ed. Lippincott-Raven, Philadelphia, p. 183.
- Benoist, C., and D. Mathis. 1999. T-lymphocyte differentiation and biology. In *Fundamental Immunology*. W. E. Paul, ed. Lippincott-Raven, Philadelphia, p. 183.
- Rolink, A., U. Grawunder, T. H. Winkler, H. Karasuyama, and F. Melchers. 1994. IL-2 receptor α chain (CD25, TAC) expression defines a crucial stage in pre-B cell development. *Int. Immunol.* 6:1257.
- Ten Boekel, E., F. Melchers, and A. G. Rolink. 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.
- Rolink, A. G., T. Winkler, F. Melchers, and J. Andersson. 2000. Precursor B cell receptor-dependent B cell proliferation and differentiation does not require the bone marrow or fetal liver environment. *J. Exp. Med.* 191:23.
- Yu, W., H. Nagaoka, M. Jankovic, Z. Misulovin, H. Suh, A. Rolink, F. Melchers, E. Meffre, and M. C. Nussenzweig. 1999. Continued RAG expression in late stages of B cell development and no apparent re-induction after immunization. *Nature* 400:682.
- Melchers, F., E. ten Boekel, T. Seidl, X. C. Kong, T. Yamagami, K. Onishi, T. Shimizu, A. G. Rolink, and J. Andersson. 2000. Repertoire selection by pre-B-cell receptors and B-cell receptors, and genetic control of B-cell development from immature to mature B cells. *Immunol. Rev.* 175:33.
- Von Boehmer, H., I. Aifantis, O. Azogui, J. Feinberg, C. Saint-Ruf, C. Zober, C. Garcia, and J. Buer. 1998. Crucial function of the pre-T-cell receptor (TCR) in TCR β selection, TCR β allelic exclusion and $\alpha\beta$ versus $\gamma\delta$ lineage commitment. *Immunol. Rev.* 165:111.
- Hoffmann, R., T. Seidl, M. Neeb, A. Rolink, and F. Melchers. 2002. Changes in gene expression profiles in developing B cells of murine bone marrow. *Genome Res.* 12:98.
- Wodicka, L., H. Dong, M. Mittmann, M. H. Ho, and D. J. Lockhart. 1997. Genome-wide expression monitoring in *Saccharomyces cerevisiae*. *Nat. Biotechnol.* 15:1359.
- Lockhart, D. J., H. Dong, M. C. Byrne, M. T. Follettie, M. V. Gallo, M. S. Chee, M. Mittmann, C. Wang, M. Kobayashi, H. Horton, and E. L. Brown. 1996. Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat. Biotechnol.* 14:1675.
- Luo, L., R. C. Salunga, H. Guo, A. Bittner, K. C. Joy, J. E. Galindo, H. Xiao, K. E. Rogers, J. S. Wan, M. R. Jackson, and M. G. Erlander. 1999. Gene expression profiles of laser-captured adjacent neuronal subtypes. [Published erratum appears in 1999 *Nat. Med.* 5:355.] *Nat. Med.* 5:117.
- Eberwine, J., H. Yeh, K. Miyashiro, Y. Cao, S. Nair, R. Finnell, M. Zettel, and P. Coleman. 1992. Analysis of gene expression in single live neurons. *Proc. Natl. Acad. Sci. USA* 89:3010.
- Tamayo, P., D. Slonim, J. Mesirov, Q. Zhu, S. Kitareewan, E. Dmitrovsky, E. S. Lander, and T. R. Golub. 1999. Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. *Proc. Natl. Acad. Sci. USA* 96:2907.
- Eisen, M. B., P. T. Spellman, P. O. Brown, and D. Botstein. 1998. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA* 95:14863.
- Hu, M., D. Krause, M. Greaves, S. Sharkis, M. Dexter, C. Heyworth, and T. Enver. 1997. Multilineage gene expression precedes commitment in the hemopoietic system. *Genes Dev.* 11:774.
- Godfrey, D. I., J. Kennedy, P. Mombaerts, S. Tonegawa, and A. Zlotnik. 1994. Onset of TCR- β gene rearrangement and role of TCR- β expression during CD3⁻CD4⁻CD8⁻ thymocyte differentiation. *J. Immunol.* 152:4783.
- Karasuyama, H., A. Rolink, Y. Shinkai, F. Young, F. W. Alt, and F. Melchers. 1994. The expression of Vpre-B/ λ 5 surrogate light chain in early bone marrow precursor B cells of normal and B cell-deficient mutant mice. *Cell* 77:133.
- Sleckman, B. P., W. N. Khan, W. Xu, C. H. Bassing, B. A. Malynn, N. G. Copeland, C. G. Bardoni, T. M. Breit, L. Davidson, E. M. Oltz, et al. 2000. Cloning and functional characterization of the early-lymphocyte-specific Pb99 gene. *Mol. Cell Biol.* 20:4405.
- Ghebrehiwet, B., B. L. Lim, R. Kumar, X. Feng, and E. I. Peerschke. 2001. gC1q-R/p33, a member of a new class of multifunctional and multicompartmental cellular proteins, is involved in inflammation and infection. *Immunol. Rev.* 180:65.
- Takebayashi, K., Y. Sasai, Y. Sakai, T. Watanabe, S. Nakanishi, and R. Kageyama. 1994. Structure, chromosomal locus, and promoter analysis of the gene encoding the mouse helix-loop-helix factor HES-1: negative autoregulation through the multiple N box elements. *J. Biol. Chem.* 269:5150.
- Iacopetti, P., M. Michelini, I. Stuckmann, B. Oback, E. Aaku-Saraste, and W. B. Huttner. 1999. Expression of the antiproliferative gene TIS21 at the onset of neurogenesis identifies single neuroepithelial cells that switch from proliferative to neuron-generating division. *Proc. Natl. Acad. Sci. USA* 96:4639.
- Campbell, K. S. 1999. Signal transduction from the B cell antigen-receptor. *Curr. Opin. Immunol.* 11:256.
- Dominguez, O., J. F. Ruiz, T. Lain de Lera, M. Garcia-Diaz, M. A. Gonzalez, T. Kirchhoff, A. C. Martinez, A. Bernad, and L. Blanco. 2000. DNA polymerase μ (Pol μ), homologous to TdT, could act as a DNA mutator in eukaryotic cells. *EMBO J.* 19:1731.
- Lang, J., B. Arnold, G. Hammerling, A. W. Harris, S. Korsmeyer, D. Russell, A. Strasser, and D. Nemazee. 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.
- Liu, Y., R. Wada, T. Yamashita, Y. Mi, C. X. Deng, J. P. Hobson, H. M. Rosenfeldt, V. E. Nava, S. S. Chae, M. J. Lee, et al. 2000. Edg-1, the G protein-coupled receptor for sphingosine-1-phosphate, is essential for vascular maturation. *J. Clin. Invest.* 106:951.
- Grawunder, U., T. M. Leu, D. G. Schatz, A. Werner, A. G. Rolink, F. Melchers, and T. H. Winkler. 1995. Down-regulation of RAG1 and RAG2 gene expression in preB cells after functional immunoglobulin heavy chain rearrangement. *Immunity* 3:601.
- Hoffmann, R., T. Seidl, and M. Dugas. 2002. Profound effect of normalization on detection of differentially expressed genes in oligonucleotide microarray data analysis. *Genome Biol.* 3:RESEARCH0033.
- Hedrick, S. M., D. I. Cohen, E. A. Nielsen, and M. M. Davis. 1984. Isolation of cDNA clones encoding T cell-specific membrane-associated proteins. *Nature* 308:149.
- Li, C., and W. H. Wong. 2001. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc. Natl. Acad. Sci. USA* 98:31.
- Li, C., and W. H. Wong. 2001. Model-based analysis of oligonucleotide arrays: model validation, design issues and standard error application. *Genome Biol.* 2:research0032.1.
- Klein, U., Y. Tu, G. A. Stolovitzky, M. Mattioli, G. Cattoretti, H. Husson, A. Freedman, G. Inghirami, L. Cro, L. Baldini, et al. 2001. Gene expression profiling of B cell chronic lymphocytic leukemia reveals a homogeneous phenotype related to memory B cells. *J. Exp. Med.* 194:1625.