Knockout Mouse Models of Differentiation by Transcriptional Analysis

A General Survey of Thymocyte

This information is current as of January 13, 2018.

Denis Puthier, Florence Joly, Magali Irla, Murielle Saade, Geneviève Victorero, Béatrice Loriod and Catherine Nguyen

*average

References

This article cites 64 articles, 35 of which you can access for free at:
http://www.jimmunol.org/content/173/10/6109.full#ref-list-1

Supplementary Material

http://www.jimmunol.org/content/suppl/2004/11/02/173.10.6109.DC1

Subscription

Information about subscribing to The Journal of Immunology is online at:
http://jimmunol.org/subscription

Permissions

Submit copyright permission requests at:
http://www.aai.org/About/Publications/JI/copyright.html

Email Alerts

Receive free email-alerts when new articles cite this article. Sign up at:
http://jimmunol.org/alerts
A General Survey of Thymocyte Differentiation by Transcriptional Analysis of Knockout Mouse Models

Denis Puthier, Florence Joly, Magali Irla, Murielle Saade, Geneviève Victorero, Béatrice Loriod, and Catherine Nguyen

The thymus is the primary site of T cell lymphopoiesis. To undergo proper differentiation, developing T cells follow a well-ordered genetic program that strictly depends on the heterogeneous and highly specialized thymic microenvironment. In this study, we used microarray technology to extensively describe transcriptional events regulating αβ T cell fate. To get an integrated view of these processes, both whole thymi from genetically engineered mice together with purified thymocytes were analyzed. Using mice exhibiting various transcriptional perturbations and developmental blockades, we performed a transcriptional microdissection of the organ. Multiple signatures covering both cortical and medullary stroma as well as various thymocyte maturation intermediates were clearly defined. Beyond the definition of histological and functional signatures (proliferation, rearrangement), we provide the first evidence that such an approach may also highlight the complex cross-talk events that occur between maturing T cells and stroma. Our data constitute a useful integrated resource describing the main gene networks set up during thymocyte development and a first step toward a more systematic transcriptional analysis of genetically modified mice. The Journal of Immunology, 2004, 173: 6109–6118.

With the systematic sequencing of cDNA libraries and the advent of microarray analysis, it has become possible to describe, at the tissue level, the expression pattern of thousands of genes already known or yet to be characterized. As an example, multitissue experiments using microarrays provided a broad source of information and constituted a first step in the description of gene expression profiles (1, 2). However, at the level of each organ, higher resolution is necessary to specify the histological expression pattern of individual genes. This task is currently in progress, but production of informative samples often requires microdissection or purification procedures (3). The extensive development of genetically modified mice in the past few years has led to the generation of numerous models of transcriptional perturbations, most of them exhibiting very precise developmental blockades and thus quantitative cellular disequilibrium (4). Hence, the use of perturbation models for transcriptional analyses may provide us with samples highly enriched in as yet poorly described maturation intermediates.

The thymus is the major site for T lymphocyte maturation. Anatomi- cally, it is divided into a subcapsular region, a cortex, where most of the thymocyte differentiation takes place and a medulla, where newly generated T cells undergo final processes of maturation. Each of these compartments forms a specialized stromal mi-

croenvironment that is crucial to control T cell fate. This stroma is essentially composed of epithelial cells, dendritic cells, macrophages, and fibroblasts. It provides developing T cells with essential extracellular matrix components, cell surface ligands, and soluble factors. Lymphoid progenitors derived from the bone marrow enter the thymus at the corticomedullary junction, and then migrate to the subcapsular region where they undergo multiple cycles of proliferation and progress to the CD4+CD8− double-negative stage (DN)4 through discrete maturation steps: CD44hiCD25− (DN1), CD44hiCD25+ (DN2), CD44loCD25+ (DN3), and CD44loCD25− (DN4). During DN to CD4+CD8+ double-positive (DP) transition, thymocytes relocate to the cortical region. Most DP cells (∼97%) die by neglect, because they do not recognize any of the available MHC molecules expressed by thymic stromal cells. Ultimately, DP thymocytes are subjected to negative selection events occurring both in the cortex and in the medulla, leading to the deletion of autoreactive clones and giving rise to the generation of MHC-restricted CD4-CD8 single-positive T cells (CD4+SP and CD8+SP) (5, 6). Although the molecular pathways involved in thymocyte ontogeny have been the subject of extensive studies, some of them based on microarray technology (7–9), our understanding of these mechanisms is still fragmented, and results often remain difficult to situate in an integrated and dynamic view of T cell maturation and T cell-stroma cross talk.

By taking advantage of the relative enrichment in specific thymocyte and stromal populations displayed by some genetically engineered mice, microarray analysis should reveal their characteristic transcriptional signatures. Identification of such signatures may highlight the functional specialization of each cell population and should allow us to propose a putative role for still poorly

Received for publication April 26, 2004. Accepted for publication August 17, 2004.

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 D.P. is supported by grants from the Agence Nationale de Recherche sur le Sida. This work was supported by Institut National de la Santé et de la Recherche Médicale and Ligue Nationale Contre le Cancer. Microarrays were developed and produced in the microarray platform of Marseille genopole.

2 D.P. and F.J. contributed equally to this work.

3 Address correspondence and reprint requests to Dr. Catherine Nguyen, Technologies Avancées pour le Génome et la Clinique/ERM 206, Parc Scientifique de Luminy, case 928, 13288 Marseille cedex 09, France. E-mail address: nguyen@tage.univ-mrs.fr

4 Abbreviations used in this paper: DN, double negative; DP, CD4+CD8+ double-positive thymocyte; CD4+SP, CD4 single-positive thymocyte; CD8+SP, CD8 single-positive thymocyte; KO, knockout; WT, wild type; DAPI, 4′,6′-diamidino-2-phenylindole; Ac.CD4P, activated CD4+ peripheral T cell; Ac.CD8P, activated CD8+ peripheral T cell.

Copyright © 2004 by The American Association of Immunologists, Inc.
TRANSCRIPTIONAL MICRODISSECTION OF THE MOUSE THYMUS

characterized genes. In addition, because T cell and stromal development are interdependent events, the combined analysis of genetically engineered mouse models affecting one or both compartments should underline such cross-talk events. In the present study, using a 8700 cDNA mouse microarray, we draw up the first transcriptional picture of the thymus as a whole by analyzing both whole thymi from well-defined knockout (KO) mice together with purified thymocytes.

Materials and Methods

Mice

Mice were housed in a specific pathogen-free animal facility. C57BL/6 (wild type (WT)), RAG1°, and TCR° mice were obtained from the Centre de Distribution, the Mouse et Réseau (Orléans, France). CD3°, CD8°, LAT°, and RelB° mice were kindly provided by Drs. M. Malissen and P. Naquet (Centre d’Immunologie de Marseille Luminge, Marseille, France). Thymi were isolated from adult mice that were sacrificed between 4 and 6 wk of age.

Organs and cells

Cells were cultured in RPMI 1640 supplemented with 10% heat-inactivated FCS, 0.1% penicillin, 0.1% streptomycin, 0.1% glutamine, and 2 × 10−5 M 2-ME. The LMTK fibroblastic cell line and the P388D1 macrophage cell line were both obtained from the American Type Culture Collection (Manassas, VA). MTE-1D is a thymic epithelial cell line (10). DC2.4 are immortalized dendritic cloned cells derived from bone marrow of C57BL/6 mice (11). 427.1, 130.81, and 6.1.1 thymic stromal cell lines were kindly provided by Dr. B. Knowles (The Jackson Laboratory, Bar Harbor, ME) (12). Thymocytes were separated from stroma according to standard procedures. CD25° thymocytes were isolated from RAG1° mice. After labeling with an anti-CD25 Ab, they were purified using magnetic beads (Dynal Biotech, Oslo, Norway) according to manufacturer’s guidelines. These samples contain CD44highCD25° (DN2) and CD44°CD25° (DN3) and are referred to as CD25° DN. For DP purification, thymocytes were sorted using a FACS Vantage cell sorter (BD Biosciences, Mountain View, CA) on the basis of costained profiles with an anti-CD4 (RM4-4; BD Pharmingen, San Diego, CA) and anti-CD8 (H-59-101-2) or an anti-CD4 (H129.196), for 30 min at 4°C followed by incubation with sheep anti-rat IgG magnetic beads (Dynal Biotech). The populations recovered were then stained with anti-CD4 (RM4-4) and anti-CD8 (53-6.7; BD Pharmingen) and FACS sorted.

Microarray design

For microarray preparation, the following cDNA libraries (some of which were incomplete) were used: the NIA Mouse 15K cDNA clone set, 2NbMT, NbMLN (lymph node), and NbMS (spleen). Detailed descriptions of these cDNA libraries are available at the UniGene database website (www.ncbi.nlm.nih.gov/UniGene/browse2.cgi?TAXID=10090, 2NbMT: Lib.544, 2NbMT: Lib.553, NbMLN: Lib.567, NIA 15K: Lib.8622). All of the libraries were cloned into pT3T7D-Pac vector, except for the NIA 15K Mouse cDNA clone set, which was cloned into pSPORT1 vector. The NIA Mouse 15K cDNA clone set is a rearrayed and resequenced set of 15,000 Mouse cDNA clones, providing internal controls to assess the reproducibility of gene expression measurements. Microarray preparation

PCR amplifications were performed in 96-well microtiter plates using the following primers: 5'-CCAGTACGACGTTGTAACGGC-3' and 5'-GTGGTGAGATTGCAAGCAGTAAC-3', which are specific of the polylinker sequence of both vectors used. The reactions were performed as described (16) by transferring free Escherichia coli from a growth culture with a plastic 96-pin gadget (Genetix, New Milton, U.K.) to 1 Trit-PCR mix, containing 10 nM Tris-Cl, 60 mM MgCl2, 50 mM KCl, 0.1% Triton X-100, 1.5 M betaine, 250 μM dATP, dTTP, dGTP, and dCTP, and 5 U of Taq polymerase (Promega, Madison, WI). The plates were incubated for 6 min at 94°C, before 38 cycles of 94°C for 30 s, 65°C for 45 s and 72°C for 210 s, followed by a final elongation phase at 72°C for 10 min. Amplification products were not quantitated, but their quality was systematically checked on 1% agarose gels. Eleven percent of the bacterial clones were estimated to be nonamplified, and 4% showed multiple bands after PCR amplification. Unpurified PCR products were then transferred to 384-well microplates, before being evaporated, taken up in 40 μl of distilled water, and spotted onto nylon membranes (Hybond-N+; Amersham Biosciences, Saclay, France). This step was conducted using a Micro-Grid-II arrayer (Apogen Discoveries, Cambridge, U.K.) equipped with a 64×32 biobrobotics pipethead. A vector probe hybridization (5′-TCA CACAGGAAAAGCAGTATGC-3′) was performed as previously described and showed uniform signal intensities across individual membranes (17).

RNA extraction and cDNA labeling

Total RNA was isolated using TRIzol reagent (Invitrogen Life Technologies, Carlsbad, CA), and RNA integrity was checked on denaturing agarose gels. Reverse transcription was performed as previously described, using 2 μg of total RNA in the presence of [α-32P]dCTP (18). A large excess of oligo(dT) primers was added during cDNA synthesis, and the labeled probe was annealed with poly(dA)80 to ensure the complete saturation of poly(A) tails. Hybridizations were conducted for 48 h at 68°C in 500 μl of hybridization buffer (18). After washing, arrays were exposed to phosphor imaging plates, which were scanned using a BAS 5000 (Fuji, Tokyo, Japan) at a 25-μm resolution.

Data processing and analysis

After image acquisition, hybridization signals were quantified using the locally developed Bzscan software (64). All images were carefully inspected, and spots with uncharacterized intensities due to neighborhood effects were manually excluded. For each of the 66 arrays, background intensity was calculated on the basis of the average of the 400 lowest values and subtracted. We defined distinct classes of samples, each of them containing all of the replicates of the same sample type (e.g., three replicates of RAG1° thymi define one class). Of note, all of the cell lines, including immortalized cells were grouped in the same class. For each sample, the 6000 highest values were flagged. A cDNA clone was kept for analysis if its expression values were flagged in all of the samples of at least one class. Data were filtered (80%), log transformed, and centered relative to the median for each gene and each array using the Cluster software (19). To analyze genes with the most fluctuating expression levels across the samples, only clones exhibiting a SD of at least 0.35 were kept. Hierarchical clustering (average linkage; clustering metric) was applied to the dataset, and results were visualized with the TREEview software (19). All data are MIAME compliant and have been submitted to ArrayExpress database (www.ebi.ac.uk/miameexpress; accession no. E-MEXP-96). This article contains supplemental data (Tables SI, SII, SIII, SIV, and SIV). Immunoﬂuorescence

Primary Abs used were as follows: FITC-anti-CD54/ICAM-1 (BD Pharmingen, San Diego, CA), anti-RANTES (AF478; R&D Systems, Minneapolis, MN), anti-K5 (AF138; Covance Research, Berkeley, CA), and anti-CD74 (Mab21-1). Secondary Abs used were as follows: Alexa 568 goat anti-rabbit IgG, Alexa-477 goat-anti-rat IgG, and Alexa-477 goat-anti-mouse IgG. Cryosections (8 μm) of 8-wk-old WT thymi were generated from OCT-embedded organs and mounted on glass slides. Sections were fixed with phosphate buffer containing 4% paraformaldehyde for 1 h before staining in a humidified chamber. Abs (in 1% BSA, 0.01% Triton, and 0.1 mM 4′,6-diamidino-2-phenylindole, 10 μg/ml) were applied for 60 min.
M Tris buffer (pH 7.4) were applied to the sections and incubated for 1 h and 30 min at room temperature. Slides were then incubated in blocking buffer (3% BSA in 0.01% Triton, 0.1 M Tris buffer, and 5% goat serum) for 10 min before being incubated with Alexa-conjugated-secondary Ab (1:1000 in 1% BSA, 0.01% Triton, and 0.1 M Tris buffer). Between each step, slides were washed for 5 min with Tris buffer. Tissues were counterstained with 1 μl/ml 4',6'-diamidino-2-phenylindole (DAPI) and mounted with Mowiol fluorescent mounting medium (Calbiochem, Darmstadt, Germany). Fluorescent images were acquired by Zeiss (Oberkochen, Germany) LSM 510 confocal microscopy. All images were processed after the same exposure time and under the same magnification.

Results

Transcriptional microdissection of the murine thymus

To reveal the transcriptional signatures of the cellular compartments that constitute the adult mouse thymus, we used a combination of whole thymus and of purified cell samples. We assumed that comparing the transcriptional status of both WT and KO mice would help us to reveal specific signatures of over- or underrepresented cellular compartments. The RAG1°, LAT°, and CD3-ε° (Cd3ε°) KO mice were chosen, because they display an early block in T cell development and consequently a relative enrichment in DN T cells and stromal cells (4). Thymus from TCRα° KO mice that lack mature T cells and differentiated medullary stromal cells were used to highlight the specific transcriptional signatures of these compartments. Mice invalidated for RelB, a member of the NF-κB family, were also used, because they are characterized by a disorganized medulla and a lack of medullary dendritic cells (20). In addition, RNA from the following sorted thymocyte populations were prepared: CD25+ T cells from RAG1° mice (corresponding to DN2 and DN3 stages, CD25+DN), thymocytes from WT animals (DN, DP, CD4+SP, CD8+SP), and activated peripheral T cells (Ac.CD8+P and Ac.CD4+P). Finally, cell lines and dendritic cells were also used to delineate genes involved in cellular metabolism. All samples were hybridized to a mouse microarray containing 8750 cDNA selected from embryo, thymus, lymph node, and spleen libraries. After normalization and data filtering, 4686 genes were kept for analysis. Hierarchical clustering was used to group genes on the basis of their variation of expression over samples. The same method was applied to group the experimental samples according to their transcriptional similarity. Results are presented in Fig. 1A. As depicted in Fig. 1B, the hierarchical clustering procedure gave rise to a highly relevant classification of the samples, divided in four major groups. Cell lines and activated peripheral T cells were grouped together due to their metabolic activity. Indeed, their main transcriptional characteristic relies on the high expression of genes encoding proteins involved in general metabolism (Fig. 1A, cluster G, and data not shown). The thymi from RAG1°, LAT°, and CD3ε° clustered together, whereas a third group contained all of the purified thymocytes and also naive T cells (not shown). The fourth class of samples was composed of whole thymus from WT, TCRα°, and RelB° mice. As shown in Fig. 1A, numerous transcriptional signatures clearly stood out (signatures A–G). These signatures were extracted and samples were reordered according to their types (the signature G corresponding to general metabolism will not be discussed herein) (Fig. 1C). A gene expression signature was named by either the cell type in which its component genes were expressed (for example, “DN T cells” signature) or the biological process in which its component genes are known to function (for example, the “proliferation” signature). Signatures of T cell maturation stages (DN, DP, and SP) as well as functional signatures (proliferation, TCR rearrangement)
were defined (Fig. 1C, signatures A–E). For the SP T cell signature (E), three slightly different signatures from Fig. 1A (E1, E2, E3) originally disseminated along the clustering figure were grouped together in Fig. 1C. This signature encompasses genes characteristic of both late DP and SP stages. Indeed, these genes are not, or only weakly, expressed in the TCR\(^{\gamma\delta}\) samples. In addition, it also contains genes more specifically expressed at the SP stage, whose expression is maintained in peripheral activated T cells. Of note, although we could detect some signatures specific for activated peripheral T cells, they will not be considered in the present paper.

Finally, a clear histological signature (F), differentiating the thymic stroma from developing thymocytes, was also defined.

**Proliferation status of developing thymocytes**

Close inspection of the genes composing the signature A (Fig. 1B) suggests that it corresponds to cell cycle. Bibliographic scanning of cluster A indicates that it is composed of genes that control transitions between G1, G2, and S phases, or that monitor entry and exit of mitosis: Ccna2, Cdc2, Cdcas3/Tome1, and Ube2c (Table I).

**Table I. Detailed gene contents of cell cycle-related signature (cluster A)**

<table>
<thead>
<tr>
<th>Functional Classification(^a)</th>
<th>Ug.Symbol(^b)</th>
<th>Ug.Mm(^c)</th>
<th>Ug.Hs(^d)</th>
<th>LL.ID (Hs)(^e)</th>
<th>Associated-MEG(^f)</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell cycle regulator</td>
<td>Cdc451</td>
<td>Mm.1248</td>
<td>Hs.114311</td>
<td>8318</td>
<td>MEG1485</td>
<td>Initiation of DNA replication.</td>
</tr>
<tr>
<td></td>
<td>Cdc2a</td>
<td>Mm.4761</td>
<td>Hs.334562</td>
<td>983</td>
<td></td>
<td>Involved in G1/S and G2/M phase transitions.</td>
</tr>
<tr>
<td></td>
<td>Ccna2</td>
<td>Mm.4189</td>
<td>Hs.85167</td>
<td>890</td>
<td></td>
<td>Involved in G1/S and G2/M phase transitions.</td>
</tr>
<tr>
<td></td>
<td>Ccna2</td>
<td>Mm.151315</td>
<td>Hs.234545</td>
<td>83540</td>
<td></td>
<td>Regulatory role in chromosome segregation.</td>
</tr>
<tr>
<td></td>
<td>Ccna2</td>
<td>Mm.209931</td>
<td>Hs.82906</td>
<td>991</td>
<td>MEG3165</td>
<td>APC activator during G1 and mitosis.</td>
</tr>
<tr>
<td></td>
<td>Ccna2-Grccl8</td>
<td>Mm.22228</td>
<td>Hs.30114</td>
<td>83461</td>
<td></td>
<td>Regulator of entry and exit of mitosis.</td>
</tr>
<tr>
<td>Chromatin structure</td>
<td>Hist2</td>
<td>Mm.258496</td>
<td>Hs.348668</td>
<td>126961</td>
<td></td>
<td>Histone Nucleosome structure.</td>
</tr>
<tr>
<td></td>
<td>H2afx</td>
<td>Mm.245931</td>
<td>Hs.147097</td>
<td>3014</td>
<td></td>
<td>Histone DNA repair.</td>
</tr>
<tr>
<td></td>
<td>H2afz</td>
<td>Mm.916</td>
<td>Hs.119192</td>
<td>3015</td>
<td>MEG2200</td>
<td>Histone Transcriptional control.</td>
</tr>
<tr>
<td></td>
<td>Nasp</td>
<td>Mm.7516</td>
<td>Hs.446206</td>
<td>4678</td>
<td></td>
<td>H1 histone binding protein. Histone transport.</td>
</tr>
<tr>
<td></td>
<td>Slbp</td>
<td>Mm.4172</td>
<td>Hs.251574</td>
<td>7884</td>
<td>MEG3280</td>
<td>mRNA histone binding.</td>
</tr>
<tr>
<td>Mitotic checkpoint component</td>
<td>Birc5</td>
<td>Mm.8552</td>
<td>Hs.1578</td>
<td>332</td>
<td>MEG3232</td>
<td>Apoptosis inhibitor expressed during G2 phase.</td>
</tr>
<tr>
<td></td>
<td>Bub3</td>
<td>Mm.927</td>
<td>Hs.418533</td>
<td>9184</td>
<td>MEG4512</td>
<td>Mitotic checkpoint component.</td>
</tr>
<tr>
<td></td>
<td>Ube2c</td>
<td>Mm.89830</td>
<td>Hs.93002</td>
<td>11065</td>
<td></td>
<td>Destruction of mitotic cyclins. Cell cycle progression.</td>
</tr>
<tr>
<td>Replication</td>
<td>Prim1</td>
<td>Mm.2903</td>
<td>Hs.82741</td>
<td>5557</td>
<td>MEG1249</td>
<td>Synthesis of oligoribonucleotide primers on DNA.</td>
</tr>
<tr>
<td></td>
<td>Top2a</td>
<td>Mm.4237</td>
<td>Hs.156346</td>
<td>7153</td>
<td>MEG518</td>
<td>Topoisomerase.</td>
</tr>
<tr>
<td></td>
<td>Pcna</td>
<td>Mm.7141</td>
<td>Hs.78996</td>
<td>5111</td>
<td>MEG2077</td>
<td>Cofactor of DNA polymerase δ.</td>
</tr>
<tr>
<td></td>
<td>Smc211/Fin16</td>
<td>Mm.2999</td>
<td>Hs.43910</td>
<td>10592</td>
<td>MEG2321</td>
<td>Mitotic chromosome condensation. DNA repair.</td>
</tr>
<tr>
<td>Transcriptional regulation</td>
<td>Srebfl</td>
<td>Mm.38016</td>
<td>Hs.108689</td>
<td>6721</td>
<td></td>
<td>Transcription regulator.</td>
</tr>
<tr>
<td></td>
<td>Zbh1</td>
<td>Mm.37216</td>
<td>Hs.12940</td>
<td>11244</td>
<td></td>
<td>Transcription factor.</td>
</tr>
<tr>
<td></td>
<td>Hmgb2</td>
<td>Mm.1693</td>
<td>Hs.434953</td>
<td>3148</td>
<td></td>
<td>DNA bending factor.</td>
</tr>
<tr>
<td>RNA maturation</td>
<td>Sfrs3</td>
<td>Mm.6787</td>
<td>Hs.405144</td>
<td>6428</td>
<td>MEG6285</td>
<td>Splicing factor.</td>
</tr>
<tr>
<td></td>
<td>Rrm2</td>
<td>Mm.99</td>
<td>Hs.226390</td>
<td>6241</td>
<td>MEG1938</td>
<td>Ribonucleotide reductase M2.</td>
</tr>
<tr>
<td></td>
<td>Hnrpab</td>
<td>Mm.24</td>
<td>Hs.81361</td>
<td>3182</td>
<td></td>
<td>PremRNA processing, mRNA metabolism, and transport.</td>
</tr>
<tr>
<td></td>
<td>No15a</td>
<td>Mm.29363</td>
<td>Hs.376064</td>
<td>10528</td>
<td>MEG418</td>
<td>Pre-RNA processing.</td>
</tr>
<tr>
<td></td>
<td>Dupt</td>
<td>Mm.276215</td>
<td>Hs.367676</td>
<td>1854</td>
<td>MEG674</td>
<td>Deoxouridine triphosphatase. DNTP metabolism.</td>
</tr>
<tr>
<td></td>
<td>Ddx39</td>
<td>Mm.271162</td>
<td>Hs.311609</td>
<td>10212</td>
<td>MEG1184</td>
<td>Nuclear RNA helicase.</td>
</tr>
<tr>
<td></td>
<td>Silg41</td>
<td>Mm.123859</td>
<td>Hs.30035</td>
<td>6434</td>
<td>MEG6925</td>
<td>Homolog to sfrs10. Splicing factor.</td>
</tr>
<tr>
<td>Miscellaneous</td>
<td>Pptprv/Esp</td>
<td>Mm.4450</td>
<td>Hs.77348</td>
<td>375048</td>
<td></td>
<td>Embryonic stem cell phosphatase.</td>
</tr>
<tr>
<td></td>
<td>Hpgd</td>
<td>Mm.18832</td>
<td>Hs.37748</td>
<td>3248</td>
<td></td>
<td>Hydroxyprostaglandin dehydrogenase 15-(NAD).</td>
</tr>
<tr>
<td></td>
<td>Topk-pending</td>
<td>Mm.24337</td>
<td>Hs.104741</td>
<td>55872</td>
<td>MEG2907</td>
<td>Serine/threonine kinase.</td>
</tr>
<tr>
<td></td>
<td>Lrpa1</td>
<td>Mm.4560</td>
<td>Hs.75140</td>
<td>4043</td>
<td></td>
<td>Lipoprotein receptor-associated protein.</td>
</tr>
<tr>
<td></td>
<td>Dyncl1</td>
<td>Mm.56455</td>
<td>Hs.5120</td>
<td>8655</td>
<td></td>
<td>Dynein L chain. Enzymatic activity.</td>
</tr>
<tr>
<td>Unknown</td>
<td>2310061F22Rik</td>
<td>Mm.37324</td>
<td>Hs.79077</td>
<td>9780</td>
<td>MEG4299</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2410017T11Rik</td>
<td>Mm.34567</td>
<td>Hs.46680</td>
<td>51001</td>
<td>MEG5810</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2281041H13Rik</td>
<td>Mm.269025</td>
<td>Hs.79077</td>
<td>9780</td>
<td>MEG4299</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) Functional classification is based upon bibliographic searches.

\(^b\) Mouse UniGene symbols.

\(^c\) Mouse UniGene cluster IDs (release no. 129).


\(^e\) Corresponding Human LocusLink IDs based on the HomoloGene database.

\(^f\) Corresponding associated metagenes (Associated-MEG) were retrieved from the Kimlab database (http://cmgm.stanford.edu/~kimlab/multiplespecies/Supplement/stab1.xls).
Many genes (Cdc20, Bab3, and Cdc41) are involved in microtubule-dependent processes that are hallmarks of dividing cells. Another feature of this cluster is the presence of numerous genes encoding components of chromatin structure. It includes three histones genes, Hist2 and two histones variants, H2afx and H2aef, as well as Nasp and Slbp. Interestingly several genes present in cluster A actively participate in DNA replication (Cdc41, Prim1, Top2a, Pcmu, and Smc21/Fin16).

To visualize these genes in terms of function, we took advantage of the classification proposed by Stuart et al. (21). They defined metagenes as sets of orthologous genes conserved between Homo sapiens, Drosophila melanogaster, Caenorhabditis elegans, and Saccharomyces cerevisiae. By compiling microarray experiments from these four species, they organized these metagenes into a gene-coexpression network revealing functional clusters conserved across evolution. Because mouse microarray data have not been included in their study, we searched for each gene of cluster A, the human orthologue, and then the corresponding metagene. Among the 36 genes of cluster A, 19 were found to have an associated metagene (Table I). In the conserved gene coexpression network, most of them colocalized in an area defined by Stuart et al. (21) as corresponding to cell cycle function (Fig. 2). Dutp and Rtm2, two genes known to be involved in RNA metabolism, were included in this metaclass, suggesting that they have functional links with the cell cycle. Likewise, MEG5810 corresponds to a still-uncategorized mouse gene (24100171H18Rik) that most likely participates in cell cycle control. Remaining metagenes were mainly found in two regions, highly enriched in genes devoted to general transcription and translation. Two metagenes (MEG4299/2310061F22Rik, MEG2907/Lrap1) were found scattered in the coexpression network. Although we cannot formally exclude their participation in cell cycle-related processes, they may reflect measurement errors or clone contaminations. Altogether, this result underscores the capacity of our procedure to group coexpressed genes in functional networks without any underlying assumption.

As expected, the expression levels of all the genes in cluster A fluctuated drastically among the samples. These variations clearly reflected the proliferative status of the biological samples. The highest proliferation rates were observed in activated peripheral T cells and also in WT, RelB+, and TCRα− thymi. This last observation can be correlated with the presence of both highly dividing immature single-positive thymocytes and DP cells in these thymi. Among the most immature thymocyte populations, total DN T cells, highly enriched in DN4 stage, were found to have the highest level of proliferation, compared with CD8+ DN thymi. An intermediate level of proliferation was observed in DP T cells, whereas the lowest rate of cycling characterized the SP population, except for CD8+ SP. Although this could reflect proliferative differences between CD4+ SP and CD8+ SP, this likely reveals the presence of a few contaminating and highly dividing CD8+ immature single-positive thymocytes in the cell preparation.

Notch1, Smarca4/Brg1, and their targets are highly expressed in CD25−DN T cells

Genes from cluster B were highly expressed in the thymi from mice characterized by an early blockade in T cell differentiation, in the CD25−DN T cells from RAG1− mice, and in DN T cells obtained from WT animals (Fig. 1B). Expression profiles of these genes in purified T cells are presented in Fig. 3 (only representative genes are depicted). The complete data set including all genes (n = 41) is available in Supplemental Table SI. Cluster B was highly enriched in genes previously shown to be strongly expressed during the first thymocyte developmental stages. It includes cell surface markers such as CD25, CD104/Itgb4, and the Gfra1 (22).
Importantly, several transcriptional regulators, including Notch1, Smarca4/Brg1, Dtx1/Dbetx1, and Hes1/Hairy were specifically expressed in DN T cells. Coregulated expressions of these genes pointed out some complex transcriptional regulatory networks. Dtx1 and Hes1 are two well-established targets of Notch1. Smarca4/Brg1 encoding a component of the SWI/SNF chromatin-remodeling complex showed a coordinated expression with Capn5. Interestingly, Smarca4/Brg1 is known to be recruited in the Cd25 core promoter at the Cd25+ DN stages (23). The notion of functional links between genes of cluster B was reinforced by the simultaneous expression of several genes involved in calcium-dependent pathways, as illustrated with the following: Calpain5/Capn5, Fk506-bp5, the calcium/calmodulin-dependent eukaryotic elongation factor-2 kinase/Eef2k, and Camk2a/CaMKII, which directly regulates the calcium channel, Cacna1h, also found in this cluster (24). These data underline the high enrichment in this signature of key regulators of Cd25+ DN stages of maturation. They suggest that additional unknown genes may be linked to the Notch1 or Smarca4/Brg1 pathway. These results also point out the functional link that may exist between Notch1, Smarca4/Brg1, and the calcium-signaling pathway.

Cluster C corresponds to genes up-regulated during the TCR rearrangement process

Genes from cluster C show a relatively high expression in all the thymi tested, in Cd25+ DN, and in DP T cells (Fig. 1B). In contrast, they were strongly down-regulated in Cd4 and Cd8 T cells. Five of 55 genes were excluded from cluster C, because they were not significantly different between immature thymocytes (Cd25+ DN and DP T cells) and mature thymic and peripheral T cells (Cd4+ SP, Cd8+ SP, Ac.Cd4+ P, and Ac.Cd8+ P) (p < 0.05, Student’s t test). Close inspection of cluster C showed that it includes several genes whose expression is known to diminish upon T cell maturation (Fig. 3 and supplemental data, Table SII). Several cell surface markers were found in cluster C (Cd24/Hsa, Cd3e, Cd3γ, Ly6e, Ly6d, Cd26, Cd127/I117r, and Sort1/Ntr3). Moreover, numerous nuclear factors, known to regulate TCR expression and rearrangements, were highly represented and clustered in the same expression network. Cut1 has been shown to bind matrix attachment region sequences, and it has been proposed to be involved in the control of DNA accessibility during V(D)J recombination (25). Sp2 is a zinc finger-containing protein that is able to bind the GT box element in TCR V α promoters (26). Tcf12/Heb is abundantly detected in thymocytes and is thought to regulate E-box sites present in many T cell-specific gene enhancers, including TCR-α and TCR-β (27). Runx1/AML1 has been shown to modulate the local accessibility of DNA to the V(D)J recombinase (28). In the same way, G22p1/Ku70 and Dnmt/Tdt are components of the DNA repair/V(D)J recombination machinery and are crucial for rearrangement and generation of diversity at the V(D)J junction, respectively. Importantly, Terf2/Traf2/TLP acts as a positive regulator of the Dnmt/Tdt promoter (29). The presence of Ppp1r1c/Tarp3 is in agreement with its recently suggested role in TCR rearrangement (30). Of note, 16 other still-uncharacterized genes share the same transcription profile, and some of them may be functionally related to rearrangement events.

Cluster D is mainly composed of genes highly expressed in WT, RelB+, and TCRα+ mice, and corresponds to DP T cell signature

As shown in Fig. 3, genes from cluster D are exclusively induced at high levels at the DP stage. DP thymocytes were characterized by the expression of genes encoding cell surface proteins such as Cd4, Cd8α, CCR9, Cd97, and Blrl1/Cxcr5 (Supplemental Table SIII). Among the transcriptional regulatory genes included in this cluster, E2s2 is already known to be up-regulated at the DP stage of development (31). The nuclear orphan receptor Rorc/RORγ is an important transcription factor that is tightly regulated during T cell development: it displays a high expression during the DP stage and is turned off as thymocyte maturation progresses (32). As suggested, RORγ may reduce cell division (as does cyclin G2/Ccng2) in developing thymocytes, until they are selected. Strikingly, a large number of genes involved in TCR-mediated signal transduction were induced at the DP stage, as cells become able to receive signals through their TCR. Three of them are adapter molecules that mediate intramolecular interactions. Adaptors are now known to be as crucial for lymphocyte activation as are receptors and effectors. Sh2d1a/SAP plays a dual functional role both as an adaptor for FynT and as an inhibitor of SH2 domain-mediated interactions (33). Lcp2/Slp76 is an adaptor protein whose deficiency leads to a severe block at the DN3 stage (34). Interestingly Sla/SLAP, also present in this cluster, associates with Lcp2/Slp76 and has been shown to be highly expressed in DP thymocytes. It appears to play an important role in positive selection by down-regulating the TCR at the cell surface (35). Two genes related to apoptosis phenomena followed the same expression program characteristic of the DP stage: the product encoded by Arf6/spArmer is known to protect cells from apoptosis, whereas A3cTms-1 encodes a pro-apoptotic protein containing a pyrin domain (PD) and a caspase-recruitment domain (CARD). Altogether, this cluster clearly highlights the setting up of some key components of the TCR signaling machinery at the DP T cell stage.

Cluster E defines the transcriptional status of single-positive thymocytes

Genes from cluster E revealed the terminal maturation status of SP thymocytes subjected to selection events. This signature presents a bias in genes encoding products already known to be involved in the positive selection of thymocytes (Supplemental Table SIV). This includes notably cell surface markers, components of the signal transduction machinery, transcriptional regulators, and also several members of the cytoskeletal network. Among highly expressed cell surface molecules are Cd5, Cd6, Cd52, Cd53, Ccr7, Igβ2, and Itm2a. Cd5 and Cd53 have been shown to be developmentally regulated and induced by TCR engagement from the Cd4+ Cd8+ TCRlow population stage during selection (36). Cdl6 is closely related to Cdl5, and both are physically associated at the immunological synapse. Itm2a is a cell surface-expressed glycoprotein that appears to be transcriptionally up-regulated during positive selection (37). Igβ2 is the β-chain of LFA-1, the ligand for ICAM1. As for the DP signature, numerous specific signal transducers are found and may regulate the threshold of TCR-mediated signals. Itk is involved in positive selection, presumably by fine-tuning the TCR signals and affecting the efficiency with which thymocytes complete their maturation (38). Sh2d2a/Arhip, a T cell adaptor molecule, binds to ITK and regulates TCR-mediated signals (39). The tyrosine phosphatase HCHP/ Shlp-1 regulates the strength of TCR-mediated signals in vivo, and, in turn, helps to set the threshold for thymocyte selection (40). Among transcriptional regulatory factors expressed by developing thymocytes, Tox is a member of DNA binding HMG box protein that is up-regulated during positive selection of DP thymocytes (41). Egr1 is another transcriptional regulator of interest, because it promotes positive selection of both Cd4 and Cd8 SP cells (42). The Kruppel-like zinc finger transcription factor (Klf2/Lklf) is induced during the maturation of SP thymocytes and rapidly down-regulated after lymphocyte activation. It has been proposed to maintain the selected thymocytes in a quiescent state enhancing...
their survival (43). The large representation of genes involved in the cytoskeleton architecture is another striking feature of this cluster. Actin cytoskeletal rearrangements play an essential role in shaping the cell response to extracellular signals providing structural frameworks for the spatial reorganization of signaling effectors. Components of the cytoskeleton network are found (tubulin α4 (tuba4), myosin (Myb9)) in addition to actin assembly regulators (thymosin β10 (Tmsb10), profilin 1 (Pfn1), Vasp, and calponin 2). Interestingly, ITK also regulates TCR/CD3-induced actin-dependent cytoskeletal events (44). Of the 69 genes defining this SP signature, 15 encode still-uncharacterized proteins that may be crucial regulators of the final steps of thymocyte maturation.

**Stromal genes specifically devoted to thymocyte education**

Genes from cluster F belong to the stroma signature: they are highly expressed in thymi from RAG1°, Cd3ε°, and LAT° mice, whereas relatively weakly expressed in all of the purified T cell populations tested (Fig. 1B). In addition, these genes are not highly expressed in all of the thymic epithelial or fibroblastic cell lines tested, suggesting that they are likely to correspond to more specific genes that may be devoted to the specialized task of thymocyte education. This signature contains 96 genes, and several of them have already been described in the thymic stroma (Supplemental Table SV). This signature includes notably 1) the cell surface markers CD83, Ly75/CD205, ALCAM/CD166, VCAM1, and ICAM1/CD54, 2) chemokines, TECK/CCL25, SDF1/CXCL12, Cxc116, CCL5/RANTES, and CXCL10/IP-10, 3) nuclear factors, Whn/Foxn1 and Pax1, 4) proteins known to be involved in antigenic presentation Ii/CLIP/CD74, H2-Eb1, H2-Aa, H2-Oa, Cathepsin L/Ctsl, and C2ta/CIITA, 5) intermediate filaments, Krt2–8/K8, Krt2–18/K18, Krt2–5/K5, Krt1–17/K17, and finally 6) gene products whose function remains to be elucidated, SPARC, TSCOT, thymus LIM protein (Tlmp-pending), and PRSS16. Beyond this quite exhaustive listing of T cell maturation regulators, uncharacterized or poorly described genes may be of particular interest for further studies. Strikingly, a hierarchical clustering analysis restricted to whole thymi samples indicated that this set of genes was able to distinguish between WT and TCRα° and RelB° thymi (Fig. 4A). As minor variations in the high expression values observed in RAG1°, Cd3ε°, and LAT° mice strongly influenced gene clustering, these samples were excluded from analysis. This allows us to define a new signature containing genes that display a low expression level in RelB° and TCRα° mice as compared with WT mice (Fig. 4B, cluster F.1). This signature revealed a high enrichment of genes already known to be highly or specifically expressed in the medullary stromal cells (Fig. 4C). Among the most obvious ones are the two cytotheratin genes, Krt2–5 and Krt1–17 (shown in H. sapiens), regulators or components of the Ag presentation machinery, C2ta/CIITA and some of its target genes (Invariant chain/Ii, H2-Aa, H2-Eb1, H2-Oa) (45), the cell surface marker ICAM1, CXCL10 (shown in H. sapiens), and the CXCL16 chemokine. To clearly establish the capability of our procedure to group genes specific to the medullary stroma compartment, immunofluorescent stainings for some of them were performed. As previously reported, Krt2–5/K5 and the invariant chain Ii/CLIP/CD74 are almost exclusively expressed in the medulla with some scattered positive cells in the cortex (Fig. 5, A and B) (46, 47). Concerning ICAM1/CD54, its expression is particularly strong in the medulla, although faint staining is also observed throughout the cortex as Lepique et al. (48) recently reported (Fig. 5C). Ccl5/RANTES is already known to be expressed in the thymus, although its precise expression pattern is not yet documented. Of particular interest and as underlined by its transcriptional profile, high levels of Ccl5/ RANTES chemokine immunostaining were also found in the medullary areas (Fig. 5D). Finally, the weak expression of the genes from cluster F.1 in RelB° and TCRα° thymi demonstrates that cross-talk events with SP thymocytes are required for their induction.

**FIGURE 4.** Genes defining the thymic stroma signature. A, Cluster F is composed of 96 genes that discriminate RAG1°, Cd3ε°, and LAT° from RelB°, TCRα°, and WT thymi. Corresponding gene names are given in Supplemental Table SV. B, Exclusion of RAG1°, Cd3ε°, and LAT° thymi leads to the identification of F1 signature. C, The F1 signature is composed of 19 genes expressed in medullary stromal cells.
Discussion

As systematic characterization of the expression profiles of genes across tissues are currently in progress, data are needed to improve our knowledge concerning their precise expression pattern and their putative function (3). In the thymus, notably, the role of numerous genes has yet to be elucidated. The use of whole thymus along with purified samples constitutes the major originality of our approach. Indeed, analyzing whole organs from genetically engineered mice allows us to visualize the transcriptional perturbations in targeted populations and their effects on cross-talk events with neighboring cells. However, purified cell samples are required to resolve the complexity of some of the signatures identified with total organs. As an illustration, purified CD25+ DN were necessary to separate the thymocyte and stromal components of the RAG1°-, Cd3ε−, and LAT+− derived signatures. It also proved to be useful to delineate specific signatures of poorly represented cell populations. For example, the differences that exist, at the T cell level, between TCRα− and WT mice were more clearly underlined by the sorted SP samples.

Interestingly, this procedure enables us to visualize the coordinate expression of components of regulatory networks. As an example, the DN specific signature contains several transcriptional regulators, as well as some of their identified targets. Indeed, once activated, Notch1 is able to convert the transcription factor CBF1/Su(H) from a repressor to an activator state. Hes1 and Dtx1/Deltex1 are two well-known targets of CBF1. Notch1 has also been proposed as a Cd25 regulator, although the promoter region of Cd25 does not display canonical CBF1/Su(H) binding sites (49). Because Smarca4/Brig1 is recruited in the promoter region of Cd25 (23), this supports the notion that multiple pathways overlap in this cluster. The transcriptional analysis of Notch1 and Smarca4/Brig1 conditional targeting would be of great interest to resolve such a complexity and to help us to temporally organize these transcriptional processes. Importantly, the recent findings of Raya et al. (50), establishing a calcium-dependant activation of the Notch1 pathway, is clearly in agreement with the simultaneous expression of numerous calcium-related genes together with Notch1. Furthermore, the high expression level of Hes1 and Dtx1/Deltex1 in RAG1°, Cd3ε+, and LAT+ T cells implies that, in these thymi, the microenvironment is able to deliver a Notch1 signal. Although an active Notch1 pathway has been clearly established in early DN thymocytes (51), this underlines the capacity of our approach to detect receptor-ligand activation signals and cellular cross talks. Consequently, in a more systematic analysis of genetically modified mice, this procedure could point out stromal defects in triggering developmental signals.

As underlined by the proliferation signature, some of the clusters identified include numerous genes involved in specific functions. We revealed a cluster containing important genes involved in TCR expression and V(D)J recombination processes. The co-regulation of genes whose role is still unclear such as the transcription factor Runx1/AML1 strongly suggests that they are part of this functional pathway. Runx1 is highly expressed in cortical thymocytes (52). It has been suggested to be crucial for cell fate determination during DN to DP transition and to be involved in proliferation events (53). Nevertheless, it has been previously reported that TCR-α and TCR-β genes harbor AML1 binding sites in their enhancer regions (54). Our data support the notion that it could be involved in TCR rearrangement processes or in the regulation of TCR gene expression. More generally, such functional clusters may give substantial clues to infer putative functions for poorly described genes.

One other interesting aspect of this work resides in the visualization of the initiation of the thymocyte signaling machinery. For example, we could define genes, related to a calcium signaling pathway and specifically expressed at the DN stage (Ccap5, FK506-bp5, Eef2k, Camk2a/CaMKII, and Cacna1h). Calcium plays important roles in thymocytes, controlling differentiation, proliferation, and apoptosis, depending on the intensity and the duration of signaling events. Our result is to be linked to the recent findings suggesting that DN thymocytes exhibit a greater ability to support capacitative Ca2+ entry than do DP thymocytes (55). This calcium-related property may contribute to the ligand-independent signaling by pre-TCR complexes. In DN thymocytes, the potent Ca2+ mobilization induced by the pre-TCR may also lead to the transcriptional induction of survival molecules (56). DP and SP clusters, as well, are highly represented by molecular components of the signal transduction machinery. Because thymocytes are able to receive signals through their TCR, they may induce a transcriptional program leading to the assembly of a specialized transduction system composed of numerous molecular sensors at the cell surface, of coactivators, and of downstream transducers such as Sh2d1a/SAP, Slp76, Sla/SLAP, ITK, Sh2d2a/Ribp, and SHP-1. In addition, the SP stage is marked by the expression of proteins involved in cytoskeletal dynamics that may be required to improve the efficiency of TCR-mediated signals. In this regard, systematic transcriptional analysis of mouse models for each of these components may lead to the dissection of TCR-mediated signal transduction pathways and to a better understanding of αβ T cell selection processes. Such an approach may help to unravel conflicting results in this area.
In contrast to most of the microarray-based studies of thymocyte development, we considered thymocytes and microenvironment as two compartments displaying an indissociable and coordinated development. Of 96 genes enclosed in the stromal signature, several have already been described as playing key roles in thymus development (e.g., Whn/Foxn1, Pax1). In the same way, this cluster also contained specific genes that may be finely regulated. Spatial expression is thought to depend on the three-dimensional organization of the thymus, and Tsclot expression is lost in epithelial cell lines and may require extracellular signals to be maintained (57). The cluster corresponding to medullary stromal cells contains genes highly expressed in WT as compared with RelB- and TCRα- thymi. As expected, it includes several genes whose expression is known to depend on the RelB/NF-κB pathway in different cellular models (RANTES/Ccl5, Cxcl10/IP-10, Icam1/CD54, MHC class II). More surprisingly, like Cd54/Icam1 and the MHC components, the great majority of the genes belonging to this signature have been shown to be regulated by IFN-γ (Ifng-3, Cxcl11/IP-10, Krl-1/17K17, and Oas12) (10, 58–62). Altogether, these genes exhibit features of tightly regulated genes that reflect the activated phenotype of UEA1+ medullary thymic stromal cells, known to be absent in RelB- mice. Their expression is dependent on SP-derived signals, because they are not expressed in TCRα- thymi. Therefore, this cluster pinpoints crucial cross-talk events that exist in the medulla. Lymphotixin (L-Tc1(2)) and LIGHT are known to be produced by SP and to deliver activation signals to medullary stromal cells through the LTβ receptor. They are presumably required for the transcriptional induction of these genes, although a putative role of IFN-γ in medullary stromal cell activation cannot be excluded.

Our analysis both at the RNA and protein levels additionally shows for the first time that RANTES/Ccl5 expression is restricted to the medullary stroma. Despite the fact that RANTES has been shown to be induced in T lymphocytes under some circumstances, its expression in the thymus is predominantly found in the medullary stromal cells. In the same way, our data indicate that, as for human (63), Cxcl10 is expressed in medullary stromal cells. The low level of expression of Cxcl16, Cxcl10/IP-10, and RANTES/Ccl5 in TCRα- mice underscores the mandatory presence of SP thymocytes for the induction of an efficient chemokine network in the medulla.

The present work underlines the benefit of using models of perturbed development to describe complex transcriptional processes in the mouse. The combined analysis of a still-limited number of models of transcriptional perturbation is sufficient to summarize the major phenomena related to thymus function in an integrated way. The identification of functionally relevant signatures will bring functional clues to uncharacterized transcripts coexpressed with well-known genes. Most importantly, the systematic analysis of KO models should be a very promising method to precisely and extensively microdissect the complex events occurring during T cell development.

Acknowledgments

We are grateful to P. Naquet, N. Auphan, and B. Malissen for critical reading of the manuscript, and to R. Houlgatte for helpful discussions. We thank A. Carrier for providing RNA from purified CD25+ DN T cells, A. Bourgeois for technical help, and N. Brun and M. Barad for cell sorting. We also thank Dr. B. B. Knowles for kindly giving us thymic stromal cell lines.

References


