Molecular Biology of the Cell
Vol. 3, 761–773, July 1992

# A New Approach to Understanding T Cell Development: The Isolation and Characterization of Immature CD4$^-$, CD8$^-$, CD3$^-$ T Cell cDNAs by Subtraction Cloning

**Sally L. Orr, Elaine Gese, and Leroy Hood**

Division of Biology, California Institute of Technology, Pasadena, California 91125

Submitted March 26, 1992; Accepted May 28, 1992

During T cell development in the mammalian thymus, immature T cells are observed that lack the cell surface markers CD4, CD8, and CD3. A subtracted cDNA library was constructed to isolate cDNAs that are specific for these immature T cells. Tissue-specific expression of 97 individual cDNAs were examined using different cell types by Northern blot analysis, and six cDNAs were analyzed by reverse transcriptase (RT) polymerase chain reaction (PCR) detection of RNA. Approximately 50% of the clones could not be detected on Northern blots, and 40% of the clones were expressed by at least one other cell-type including monocytes, mature T cells, and B cells. Eight cDNA clones appear to be specific for the CD4$^-$, CD8$^-$, CD3$^-$ T cell line, used to construct the library, as determined by Northern blot analysis. In addition, 330 cDNA clones were subjected to partial automated DNA sequence determination. Database searches, with both nucleotide and protein translations, revealed cDNAs that exhibit interesting similarities to human cell-cycle gene 1, platelet-derived growth factor receptor, c-*fms* oncogene (CSF-1) receptor, and members of the immunogobulin gene superfamily. This approach of employing subtraction coupled with large scale partial cDNA sequence determination can be useful to identify genes that may be involved in early T cell growth, cellular recognition or differentiation.

## INTRODUCTION

Murine T cell precursors originate in the fetal liver and bone marrow, then migrate into the thymus where they develop into mature T cells (Moore *et al.*, 1967; Scollay *et al.*, 1986; Adkins *et al.*, 1987). The first precursor cells to colonize the thymus are CD4$^-$, CD8$^-$, CD3$^-$ and are referred to as triple-negatives (TN).[1] During development in the thymus, TN T cells undergo rapid proliferation and differentiate into double positives (DP) T cells that express both the CD4 and CD8 proteins. After a complex program of intrathymic maturation and selection, a small number of DP cells mature into immunologically competent T cells that express either the CD4 protein (helper T cells) or the CD8 protein (cytotoxic/suppressor T cells). These mature single positive (SP) T cells exit the thymus into the peripheral lymphoid organs and lymph nodes. After birth, the level of TN T cells decreases to ~1–3% of the total thymocyte population (Fowlkes and Pardoll, 1989). Currently there are 11 described phenotypes of thymic CD4$^-$, CD8$^-$ T cells, and the lineage relationships among all the different phenotypes has not been determined (Wilson *et al.*, 1988). Both fetal and adult TN T cells are clearly important because they have self-renewal capacity and precursor potential for the mature single positive T cells (Fowlkes *et al.*, 1985; Scollay *et al.*, 1988). Thus, while triple negative T cells constitute the major form of intrathymic stem cells, little is known about the molecular aspects of these cells.

Recent progress in understanding cellular functions in the later stages of T cell development has arisen from the analysis of differentially expressed cell surface proteins including the T cell receptor (TCR), CD4, and CD8 (Adkins *et al.*, 1987; Richie *et al.*, 1988). Murine triple-negative T cells can also be separated into subpopulations depending on the differential expression of cell surface proteins including the IL-2 receptor, Pgp-1/

[1] Abbreviations used: TN, triple negative

CD44, and heat-stable antigen (HSA) (Scollay *et al.*, 1988). Although some differentially expressed T cell-specific proteins have been identified with monoclonal antibodies, alternate approaches such as differential hybridization and cDNA subtraction have been successful in isolating a number of T cell-specific genes such as the TCR (Hedrick *et al.*, 1984; Yanagi *et al.*, 1985), CD4 (Maddon *et al.*, 1985), and CD8 (Kavathas *et al.*, 1984). Additional cDNAs isolated by subtraction methods include, CTL-specific cDNAs (Brunet *et al.*, 1986; Schall *et al.*, 1988), NK-specific cDNAs (Houchins *et al.*, 1990), and cDNAs that encode proteins that may be involved in T cell activation (Burd *et al.*, 1987; Jongstra *et al.*, 1987; Zipfel *et al.*, 1989). These hybridization-based methods have proved useful in isolating cDNAs encoding proteins that were initially unknown, but subsequent analysis revealed new insights into the functions of the cells from which they were isolated. For example, the identification of CTL-specific serine proteases lead to new understanding of CTL-induced lysis (Gershenfeld and Weissman, 1986).

The kinetics of cDNA:RNA hybridizations between a T cell hybridoma and a B cell line suggest that there are ~200–300 different mRNAs that encode the genes that are T cell-specific (Hedrick *et al.*, 1984). No kinetic solution hybridization experiments have been done between precursor T cells and their more mature progeny because of the difficulty in isolating sufficient amounts of homogenous precursor T cells. Indeed the phenotypic complexity and rarity of immature T cells in the thymus have hampered conventional cellular analyses. As a new approach to study the developmental processes that may be unique to immature T cells, we have employed a sensitive technique to construct a subtracted immature but not mature cDNA library. The library contains TN T cell cDNAs, obtained from an AKR thymoma, minus sequences shared by mature peripheral T cells. Individual cDNAs in the subtracted library were tested for differential expression and 330 clones were subjected to limited DNA sequence determination. Results from the expression analysis and DNA sequence determination of TN T cell cDNAs will be used to identify the function(s) of their corresponding gene products and help elucidate the cellular processes unique to T cell development in the thymus.

## MATERIALS AND METHODS

### *Cell Lines and Flow Cytometry*

The thymic lymphoma 705 was a gift from Jim Allison (Richie *et al.*, 1988). For flow cytometry, 3–5 × $10^5$ cells were incubated at 25°C in phosphate-buffered saline (PBS), 5.0% fetal calf serum (FCS), 0.1% $NaN_3$, and a dilution of the mAb tested. Cells were washed, and if necessary, incubated with a second-stage fluorescent reagent. After washing, cells were resuspended in PBS, 1.0% formaldehyde and stored in the dark at 4°C before analysis. Flow cytometric analysis was performed on 10 000 cells with an Ortho System 50H cell sorter with a 5 W argon laser. Antibodies were obtained from the following sources: anti-Thy-1, anti-CD3 epsilon (Boehringer Mannheim, Indianapolis, IN); anti-mouse Lyt-2, anti-mouse L3T4 (Becton Dickinson, Mountain View, CA) and the following antibodies were kind gifts of Dr. Ellen Rothenberg, IM7.8.1 (anti-Pgp-1) and M1.69 (anti-HSA).

### *Construction of the $CD4^-$, $CD8^-$, $CD3^-$ T Cell Subtractive cDNA Library*

The oligonucleotide primers used were the following; oligo-dT/*Xba* I primer (5′ GCAGGTCGACTCTAGATTTTTTTTTT 3′), the *Eco*RI primer (5′ GGGAGACCGGAATTC 3′), and the vector multiple cloning site oligonucleotide (vector-MCS) (5′ GAATTCGAGCTCGCCC-GGGCATCCTCTAGA 3′). Total RNA was isolated by the method of Chomczynski and Sacchi (1987), and poly(A)+ RNA was isolated by oligo dT chromatography. The subtracted cDNA library was constructed with the lambda GEM-2 vector (Promega Biotec, Madison, WI), as previously described, (Palazzolo and Meyerowitz, 1987) with an Amersham cDNA synthesis kit (Amersham, Arlington Heights, IL). Five micrograms of 705 poly(A)+ RNA was used to synthesize double-strand cDNA with the oligo-dT/*Xba* I primer. The cDNA was methylated with *Eco*RI methylase, ligated to *Eco*RI linkers with T4 DNA ligase and digested with *Eco*RI and *Xba* I. The cDNA was ligated into the lambda GEM-2 vector, packaged, and the entire library amplified. Lambda DNA from the 705 library was prepared with formamide (Davis *et al.*, 1980), digested with *Xba* I, and T7 RNA polymerase was used to synthesize sense cRNA (Amersham). Poly(A)+ sense cRNA was isolated by oligo dT chromatography and used to synthesize antisense cDNA with the oligo-dT/*Xba* I primer. Remaining RNA was removed by hydrolysis with NaOH (0.1 M NaOH, 70°C for 20 min).

Subtractive hybridizations were carried out according to Davis (1986). The first hybridization contained 180 ng cDNA and 5 μg L10A (B cell) poly(A)+ RNA in 0.5 M phosphate buffer, 0.25% sodium dodecyl sulfate (SDS), 1.25 mM EDTA at 68°C to a $C_ot$ of 1050 (M × s). Hybridizations were performed in siliconized sealed capillary tubes containing two to three glass beads to facilitate the hybridizations by mixing (Van Ness and Hahn, 1982). The nonhybidizing single-strand cDNA was separated from double-stranded cDNA-RNA hybrids by hydroxylapatite (HAP) chromatography (Bio-Rad, Cambridge, MA) at 60°C. Single-strand cDNA containing fractions were concentrated and phosphates removed by the use of Centricon 30 columns prespun with 10 μg bovine serum albumin (BSA) and 10 μg tRNA. The subtractive hybridizations were repeated with the use of a 50-fold excess of AKR lymph node poly(A)+ RNA to a $C_ot$ of 960. For subtractive hybridizations poly(A)+ RNA was isolated from total lymph node cells excised from 2- to 3-mo-old mice. After HAP chromatography the remaining cDNA was made double-stranded with the *Eco*RI primer, DNA polymerase I and cloned into the lambda GEM-2 vector. The library was packaged with Gigapack gold (Stratagene, La Jolla, CA) and was not amplified.

### *Screening the Subtractive cDNA Library*

For all probes and library construction, poly(A)+ RNA was purified twice by oligo dT chromatography. $^{32}$P-cDNA probes were synthesized with AMV reverse transcriptase and 250 μCi $^{32}$P-dCTP to a specific activity of >1.0 × $10^7$ cpm/μg (Gerard, 1988). Membrane-bound polysomal RNA was isolated by the method of Mechler and Rabbitts (1981). One microgram of poly(A)+ RNA was used to synthesize a single-strand cDNA probe with a specific activity of 2.17 × $10^7$ cpm/μg. Hybridizations were carried out overnight in 50% formamide, 5× SSPE, 10% (wt/vol) dextran sulfate, 5× Denhardt's, 100 μg/ml salmon sperm DNA, 1.0% dried nonfat milk, and 0.1% SDS at 37°C. Blots were washed twice in 2× SSC, 0.1% SDS at room temperature and twice in 0.2× SSC, 0.1% SDS at 65°C.

### *Northern and RNA Analysis*

Total RNA was isolated from cell lines or whole tissues, and poly(A)+ RNA was isolated by oligo dT chromatography. For Northern blots,

3–10 μg of poly(A)+ RNA was separated on 0.66 M formaldehyde gels as described (Davies *et al.*, 1986). Probes were labeled by random priming (Feinberg and Vogelstein, 1983) and were added to hybridization solutions containing 5× SSPE, 5× Denhardt's, 100 μg/ml salmon sperm DNA, 10% dextran sulfate, 0.1% SDS and hybridized 16–24 h at 55°C. Blots were washed once in 2× SSC, 0.1% SDS at 25°C and three to four times in 0.2× SSC, 0.1% SDS at 65°C. Blots were air-dried and exposed to X-ray film (XAR, Eastman Kodak, Rochester, NY) for 24 h to 2 wk. Blots were stripped between hybridizations by boiling in 0.01% SSC, 0.1% SDS for 10 min. All blots were tested by hybridization to murine beta actin. For specific reverse transcriptase (RT)/polymerase chain reaction (PCR) amplification (Rupp and Weintraub, 1991), 1 μg poly(A)+ RNA was obtained from the tissues indicated and cDNA synthesized with random hexamer primers and AMV reverse transcriptase for 1 h at 42°C. Input cDNA was diluted and the amount was calibrated with actin primers in a 50-μl reaction containing PCR buffer (10 mM tris(hydroxymethyl)aminomethane (Tris)-HCl pH 8.5, 50 mM KCl, 3 mM MgCl2, 0.1 mg/ml gelatin), 0.2 mM of each dNTP, 1 pmol actin primers, and 5 U AmpliTaq DNA polymerase (Perkin Elmer Cetus, Emeryville, CA). Once the amount of input cDNA was calibrated, PCR reactions were run as before except for the addition of 25 pmol of primers specific for each clone. PCR cycles were performed in an automated DNA thermal cycler (Perkin Elmer Cetus) with the following temperature profile: denaturation at 94°C for 30 s, annealing at 55°C for 45 s and primer extension at 72°C for 1 min for 24–28 cycles. One-fifth to one-half of each sample was electrophoresed on a 2% agarose, 1% nusieve gels.

### *DNA Sequence Analysis*

DNA sequence determination was performed with Sequenase I and M13 fluorescent primers (United States Biochemical, Cleveland, OH). Reactions were run on an ABI 373 automated DNA sequencer (Applied Biosystems, Foster City, CA). Approximately 400 bp were compared with the EMBL and GenBank databases (version 69) with both the FASTA (Pearson and Lipman, 1988) and BLAST (Altschul *et al.*, 1990) sequence analysis programs run on a VAXstation 3100 M38 at the Caltech Division of Biology Sequence Analysis Facility. DNA and protein matches were analyzed with the sequence analysis software from the Genetics Computer Group (GCG) (Devereux *et al.*, 1984). The sequences were also translated, searched for possible coding regions (Gribskow *et al.*, 1984; Staden, 1984), and the resulting protein sequences were also used to search the Protein Information Resource (PIR) (release 20) and National Biomedical Research Foundation (NBRF) (release 27) databases.

## RESULTS

### *Characterization of the CD4⁻, CD8⁻, CD3⁻ Triple Negative T Cell Line*

The aim of this study was to isolate the genes that are solely expressed by $CD4^-$, $CD8^-$, $CD3^-$ (triple-negative) T cells. To construct the initial library, we used an AKR thymic lymphoma (705) previously characterized at the molecular level and was determined to be negative for CD4, CD8, and CD3 cell-surface expression (Richie *et al.*, 1988). The phenotype of this cell line was further examined by flow cytometry (Figure 1). The 705 cell line was determined to be $CD4^-$, $CD8^-$, $CD3^-$, $Thy\text{-}1^+$, $Pgp\text{-}1^+$, $HSA^+$. In addition, this cell line is $H\text{-}2K^{hi}$ and IL-2 receptor negative. This thymic lymphoma is negative for expression of the TCR beta genes, however, it does express the TCR alpha chain transcript that may represent immature J-$C_{alpha}$ transcripts.

### *Construction of a CD4⁻, CD8⁻, CD3⁻ T Cell Minus Lymph Node cDNA Library*

The strategy used to construct the subtractive cDNA library is diagrammed in Figure 2 (Palazzolo and Meyerowitz, 1987). First a cDNA library was made with TN T cell cDNA and was directionally cloned into the lambda GEM-2 vector that contains both the promoters for the SP6 and T7 RNA polymerases. These promoters were used to synthesize large quantities of antisense cDNA used in subtractive hybridizations. The first subtractive hybridization contained single-strand antisense cDNA prepared from the TN T cell library (705) that was hybridized to a fivefold excess of B-cell (L10A) poly(A)+ mRNA to a $C_o t$ of 1050. This initial subtraction eliminated abundant common lymphocyte sequences and sequences shared by growing cell lines. Another advantage of the initial subtraction is that it reduced the amount of input TN cDNA and therefore less lymph node poly(A)+ mRNA was needed to achieve a high $C_o t$ value in the second subtraction. The second subtractive hybridization was performed with a 50-fold excess of lymph node poly(A)+ mRNA, as a source of mature peripheral T cell sequences, to an estimated $C_o t$ of 960. The double-subtracted TN T cell (minus B cell, minus lymph node) cDNA library (705-B-LN), contained 10 500 individual cDNA clones. To assess the effectiveness of the subtractive hybridizations, filters containing cDNAs from both the primary 705 library and the subtracted 705-B-LN library were screened with $^{32}P$-labeled actin cDNA. Filters from the unsubtracted 705 library revealed 3% actin clones, whereas the 705-B-LN library revealed 0.2% actin clones out of 480 705-B-LN phage tested. Therefore two rounds of subtraction resulted in >10-fold depletion of abundant unwanted sequences.

### *Screening the Triple Negative T Cell-Specific Library*

The strategy used to screen the 705-B-LN subtractive cDNA library is diagrammed in Figure 3. One-third (3500) of the 705-B-LN subtractive library was plated and analyzed by lambda plaque filter hybridizations. First, to identify clones that contained no cDNA inserts (background), filters were screened with the $^{32}P$-end-labeled vector-MCS to identify the multiple cloning site of the vector. In addition, filters were screened with a $^{32}P$-labeled cDNA probe derived from L10A B-cells to identify abundant, common lymphocyte sequences that escaped subtraction. Background estimates revealed that a total of ~33% of the cDNAs corresponded to background vector sequences and 6% were unsubtracted B-cell sequences. The library was

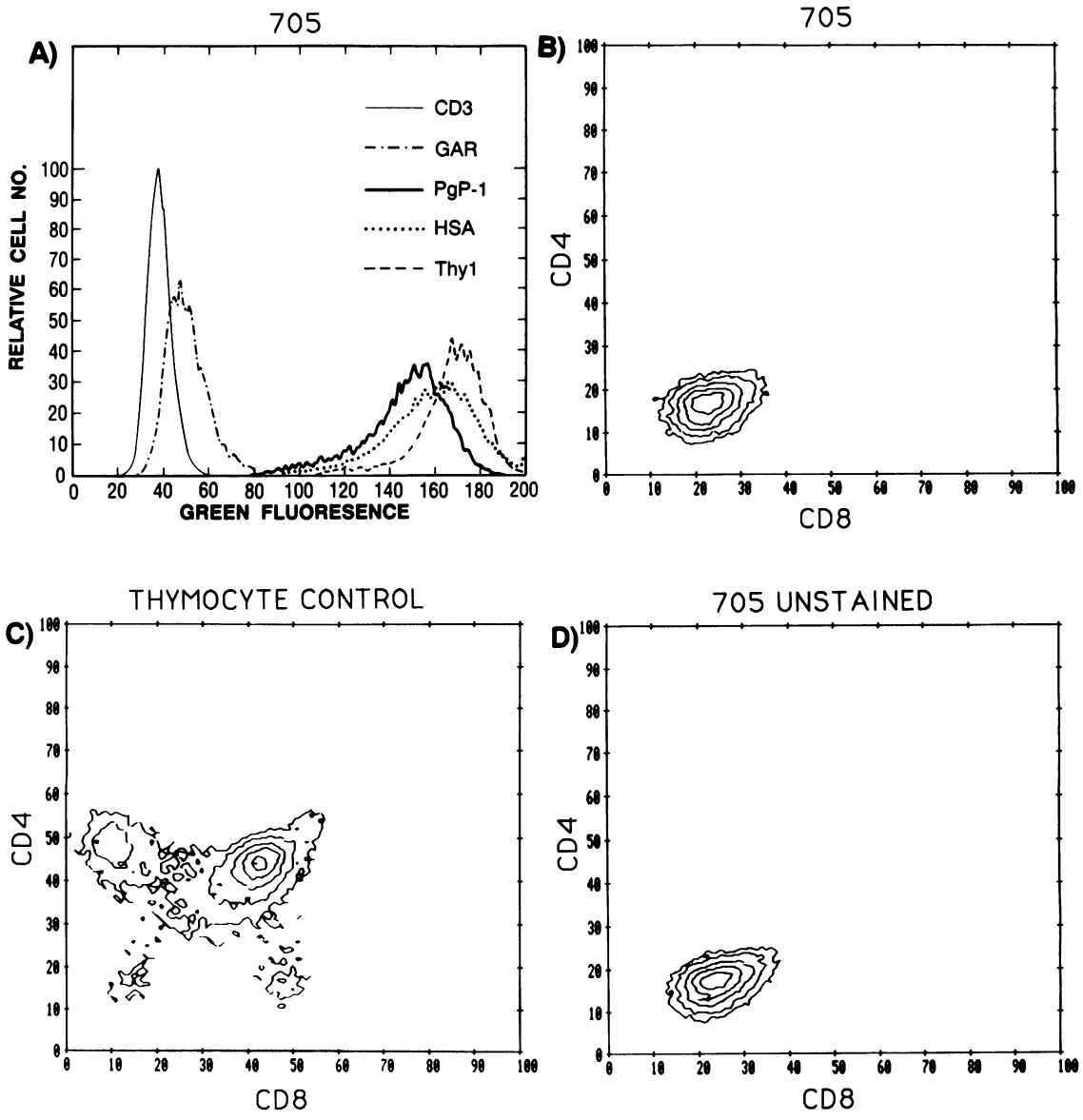


**Figure 1.** Flow cytometry analysis of the 705 cell line. One- and two-color immunofluorescence analysis of the 705 T cell line and normal AKR thymocytes. (A) 705 cells reacted with FITC-conjugated anti-CD3; anti-Pgp-1, anti-HSA, and anti-Thy-1 and visualized with FITC-conjugated goat anti-rat. (B) 705 T cell line stained with PE-conjugated anti-CD4 and FITC-conjugated anti-CD8. (C) AKR thymocytes from an adult mouse stained with PE-conjugated anti-CD4 and FITC-conjugated anti-CD8. (D) Unstained 705 cells.

also subjected to a differential screen to identify those sequences that are differentially expressed at a high level by TN T cells. Labeled probe was synthesized from the 705 T cell line and hybridized to the cDNAs that were negative with the L10A B cell probe. This differential screen revealed 106 clones (D-clones) that were positive with the TN T cell probe. Because the goal of this study was to isolate rare low-abundance mRNAs that could not be identified by hybridization with $^{32}$P-labeled cDNA probes of high sequence complexity, 2083 phage were individually picked that were negative with the vector, 705 and B-cell probes. These cDNA clones that failed to hybridize to any probe, may contain some of the rare single-copy sequences (S-clones).

Of the 2083 S-clones, 1300 phage cDNA inserts were amplified by PCR, separated by gel electrophoresis, and immobilized on nylon filters. This strategy was employed to increase the amount of cDNA present on the filters and thereby increase the signal intensity after hybridizations. These filters were then screened with a variety of $^{32}$P-labeled single-strand cDNA probes. First, a $^{32}$P-labeled first-strand cDNA synthesized from B-cell mRNA (L10A) was used to identify and therefore elim-

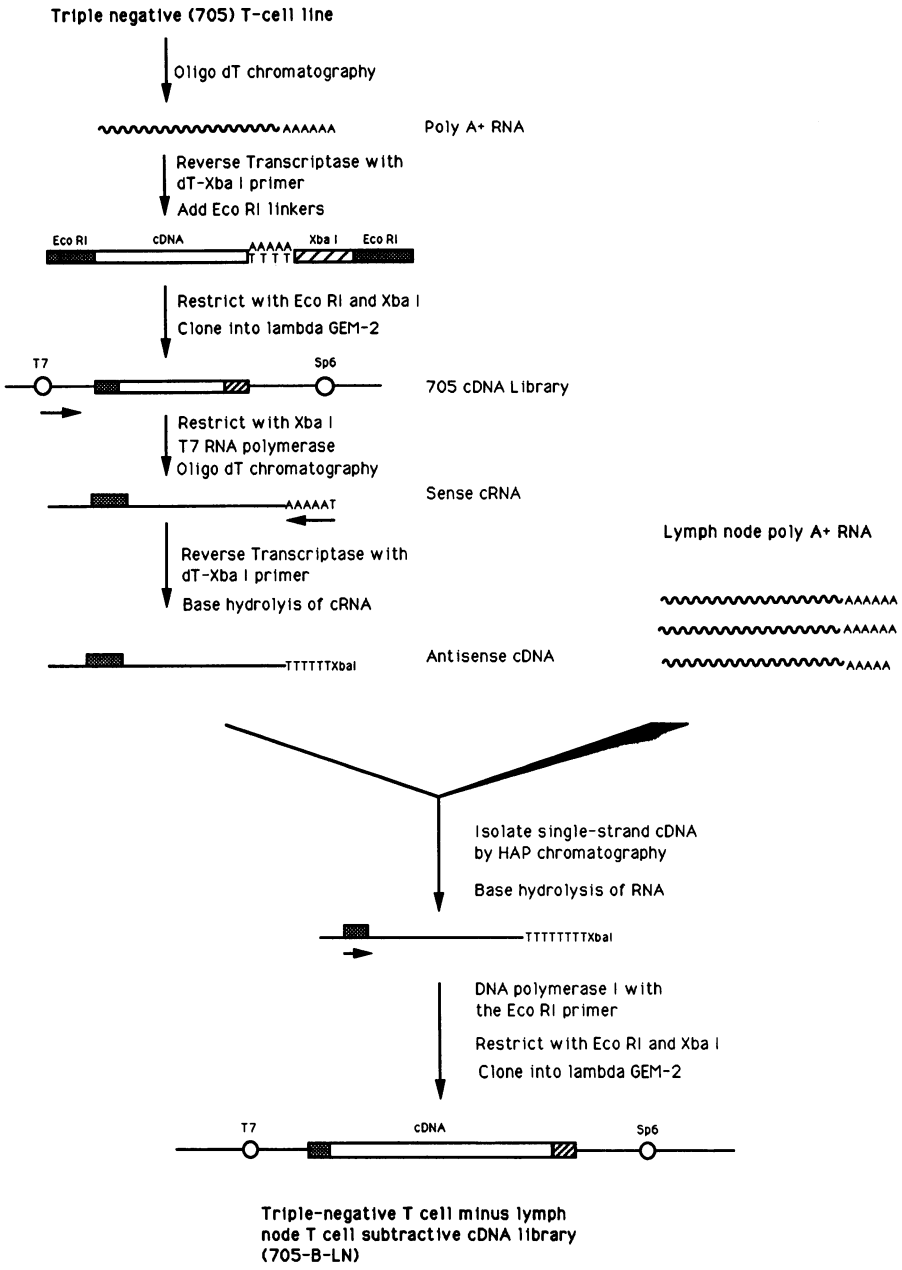


**Figure 2.** Construction of the subtractive triple-negative T cell cDNA library. To construct the library, 5.0 μg of 705 poly(A)+ RNA were used to synthesize a directionally-cloned cDNA library with the use of the oligodT-*Xba* I primer. Double-stranded cDNA was methylated, ligated with *Eco*RI linkers, restricted with both *Eco*RI and *Xba* I, and directionally cloned into the lambda GEM-2 vector. The 705 library contained $2.15 \times 10^6$ independent cDNA clones. DNA from this library was digested with *Xba* I and used to synthesize 58 μg of sense-cRNA with T7 RNA polymerase. The resulting poly(A)+ cRNA (3.4%) was isolated and used to synthesize antisense cDNA. After two rounds of subtractive hybridizations, 60% of the input cDNA remained. Because neither single-strand cDNA nor unincorporated label bind to the HAP column, the amount of radioactivity in the first fractions may reflect recovery off the column rather than enrichment of cDNA sequences.

inate 127 clones that were not detected by the first lambda plaque screen.

The next screen was utilized to identify cDNA clones that correspond to mRNAs found on membrane-bound polysomes and thus enrich for cDNAs representing potential cell-surface proteins (Mechler and Rabbitts, 1981). Membrane-bound polysomal mRNA from the 705 T cell line was isolated and used to synthesize a $^{32}$P-labeled cDNA probe. This probe was hybridized to the PCR amplified inserts and 115 clones were identified as positive with the membrane-bound polysomal probe (M-clones). In addition, 12% of the M-clones also corresponded to cDNAs previously identified as being expressed by B cells and may represent common abundant lymphocyte sequences that escaped subtraction.

Next the complexity of the cDNA clones was determined by labeling random pools of clones and eliminating clones that were redundant. The first complexity screens contained 10 random clones that were labeled and hybridized to the library. The strategy was to continue these cross-hybridizations until a majority of the library had been analyzed. After a small number of complexity screens, it became apparent that some of the clones cross-hybridized and were abundant among the M-clones, D-clones, and the S-clones. DNA sequence analysis revealed that these over-represented

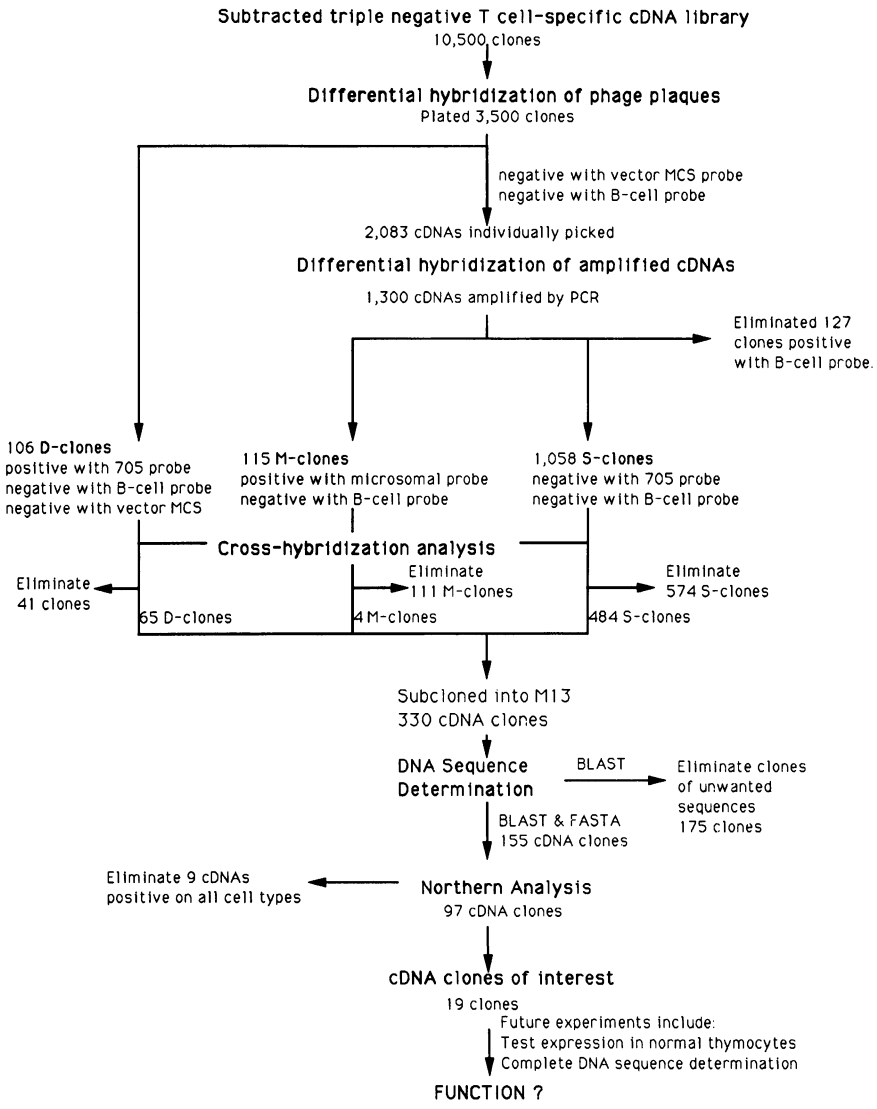


**Figure 3.** Protocol used to sort the subtracted triple-negative T cell cDNAs. This figure represents the screening procedures used to categorize the 705-B-LN cDNAs. After cross-hybridization analysis, cDNA clones were subjected to expression analysis on Northern blots and DNA sequence determination. Clones chosen for DNA sequence determination include 1 M-clone, 7 D-clones, and 322 S-clones. Not all clones subjected to DNA sequence determination have been subjected to Northern blot analysis. Clones chosen for Northern analysis include 1 M-clone, 28 D-clones, and 68 S-clones. A small number of clones will be subjected to further analysis to determine their corresponding cellular functions.

sequences were endogenous MuLV retroviral sequences. The D-clones contained the highest percentage of retroviral sequences with 40% of the clones representing one of two common retroviral sequences. Cross-hybridization analysis and limited DNA sequence determination of abundant clones also revealed that some of the clones were *Escherichia coli*-derived. These *E. coli* contaminants were not present in the subtracted cDNA library and came from the use of PCR products (cross-hybridizing PCR primers) on filters and in the subcloning procedure. After eliminating the retroviral and *E. coli* contaminants from the complexity analysis, cross-hybridization demonstrated that 19% of the M-clones were unique. Cross hybridization with pools of S-clones demonstrated that these clones were the most complex and were determined to be between 95 and 97% unique. After the complexity screens and the removal of all retroviral and *E. coli* sequences, 65 D-clones, 4 M-clones, and 484 S-clones remained to be analyzed.

### *Expression Analysis of Triple-Negative T Cell-Specific cDNAs by Northern Blots and PCR Amplification of RNA*

Inserts of each expression group were subcloned into M13mp19 to determine a limited (300–500 bp) DNA sequence, and the RNA expression of individual clones was examined by Northern blot analysis. A single-run DNA sequence determination was used to eliminate any unwanted cDNA clones such as mitochondrial genes. The remaining cDNA clones that encoded previously unidentified DNA sequences were then tested for TN T cell-specific expression and/or expression in other cell types, employing RNA isolated from continuous cell lines on Northern blots. Continuous cell lines were used

because this initial round of screening required large amounts of poly(A)+ RNA to identify possible rare messages. Northern blots contained 3–10 μg of poly(A)+ RNA from a minimum of four cell lines. These cell lines included TN T cells (705), B cells (L10), pro-B cells (NS70), monocytes (Wehi-3), fibroblasts (L-929), mast cells (P815), and T cells (EL-4 and YAC-1). Northern blots revealed widely divergent expression patterns including, TN T cell-specific expression (D51, S39), the expression of multiple transcripts of different sizes (S183), and transcripts that appear to be expressed at an increased level by TN T cells (M517) (Figure 4). The major RNA expression patterns observed are represented in Table 1. Approximately 67% of the S-clones, i.e., clones that failed to hybridize with any of the labeled cDNA probes, were negative on Northern blots against all of the represented cell types. Over 50% of the cDNAs in the library, when tested for expression on Northern blots, failed to identify a transcript. This suggests that these cDNAs represent very low abundance transcripts. The majority of the S-clones that did hybridize to poly(A)+ RNA were either expressed only in the 705 TN T cell line or were widely expressed in all cell types examined. Approximately 30% of all the cDNA clones tested on Northern blots were expressed in other cell types in addition to the 705 T cell used to construct the library. These clones may have either escaped the subtraction procedure (expression pattern 2, 4, 5) or were not expressed by the lymph node T cells or B cells used for the subtraction (expression pattern 3). Eight cDNAs (4 S-clones, 2 M-clones, and 2 D-clones) were identified that were expressed solely by the 705 TN T cell line and not the other cell types tested.

To further examine the expression of individual cDNA clones, the more sensitive assay of multiplex RT/

**Table 1.** Expression patterns of immature cDNAs determined by Northern blot analysis

| | RNA expression pattern[a] | | | | | |
|---|---|---|---|---|---|---|
| Cell line RNA/cell type | 1 | 2 | 3 | 4 | 5 | 6 |
| 705/triple negative T cell | + | + | + | + | + | − |
| L10/B cell | − | + | − | − | + | − |
| Wehi-3/monocyte | − | − | + | − | + | − |
| EL-4/T cell | − | − | − | + | + | − |
| Single copy clones (S clones) | 4 | 1 | 0 | 1 | 3 | 19 |
| Membrane-bound clones (M clones) | 2 | 1 | 0 | 0 | 0 | 3 |
| Differential clones (D clones) | 2 | 2 | 2 | 2 | 4 | 7 |
| Total number/expression pattern | 8 | 4 | 2 | 3 | 7 | 29 |

This table presents data from the major expression patterns observed using Northern blots with 10 μg/lane poly(A)+ RNA from the indicated cell types. Other Northern blots produced contained poly(A)+ RNA from other cell types than those listed here.

[a] The pattern of expression is read downward with + representing observed hybridization and − repesenting no hybridization as detected by Northern blot analysis.

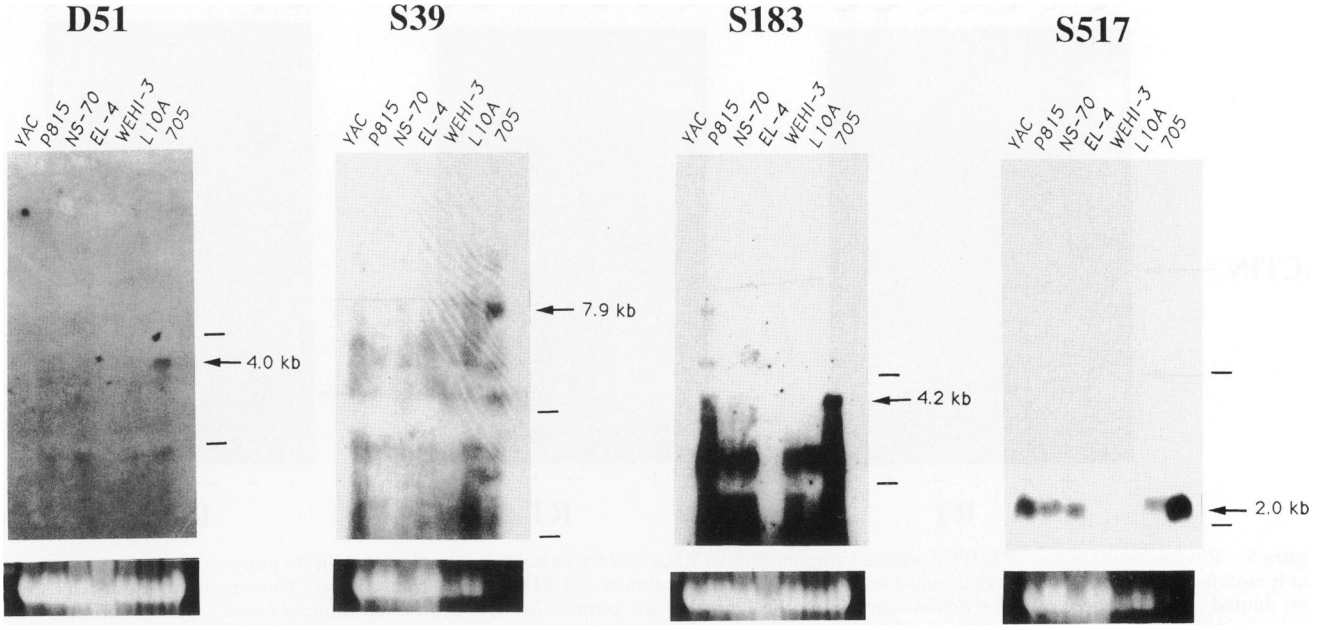


**Figure 4.** Northern analysis of cDNA clones tested upon various cell types. cDNA inserts were isolated from indicated lambda clones and purified by gel electrophoresis. The insert sizes of the cDNAs ranged from 200 bp to >4 kb, with the average insert size ~1.0 kb. cDNAs were labeled with $^{32}P$ by random priming and hybridized as described in METHODS. All blots shown contain 10 μg/lane poly(A)+ RNA from the various cell types indicated. The position of the 28S and 18S rRNA are indicated and amount of RNA loaded is shown by the 18S band.

PCR amplification was used to reveal the presence of RNA transcripts in the small amounts of RNA obtainable from normal thymocytes. Individual primer pairs were synthesized for six cDNAs. Four pairs were synthesized for cDNAs that demonstrated TN-specific expression on Northern blots and two pairs synthesized for clones with interesting sequence similarities to known genes. Developmental expression was examined employing RNA from day 16, fetal thymocytes, newborn thymocytes, and lymph node cells (Figure 5). One cDNA chosen for its similarity to known growth factors (S334), revealed equal expression in all cell types tested and demonstrated that this technique will be useful for further testing of Thymocyte-specific expression. TN-specific cDNA D12 was observed to be expressed in both day 16 thymocytes and newborn thymocytes, but not lymph node cells. Two other clones were expressed only in newborn thymocytes and not by day 16 thymocytes or lymph node cells (S183, D66). This approach was unsuccessful for two cDNAs that may reflect altered expression by the 705 thymic lymphoma as a result of transformation or incorrect synthesis of primers due to single-run DNA sequence determination.

### *Limited DNA Sequence Determination of Triple-Negative T Cell-Specific cDNAs*

Partial nucleotide sequence was obtained from the 5′ end of 330 random cDNA clones > 2 kb. Limited sequence determination produced an average of 452 ± 114 (mean ± SD) bases, which were analyzed with the FASTA and BLAST sequence analysis programs. Moreover, sequences were translated in all six reading frames and subjected to analysis to determine putative coding regions. Peptide sequences generated from large open reading frames or stretches of 20 or more amino acids from the coding analysis data were also used to search the GenBank and protein databases. Sequence determination indicated that the library contained ~10–20% random yeast sequences. This contaminant may have been a consequence of continuous use of tissue culture cells for the initial cDNA library. If any yeast sequences were cloned in the initial library, they would not have been eliminated by the subtractive hybridizations with normal lymph node tissues. Of the 330 cDNA clones sequenced, 175 were unwanted sequences, including *E. coli* genes (subcloning artifacts) (32%), yeast genes (7%), and known genes such as viral genes, mi-

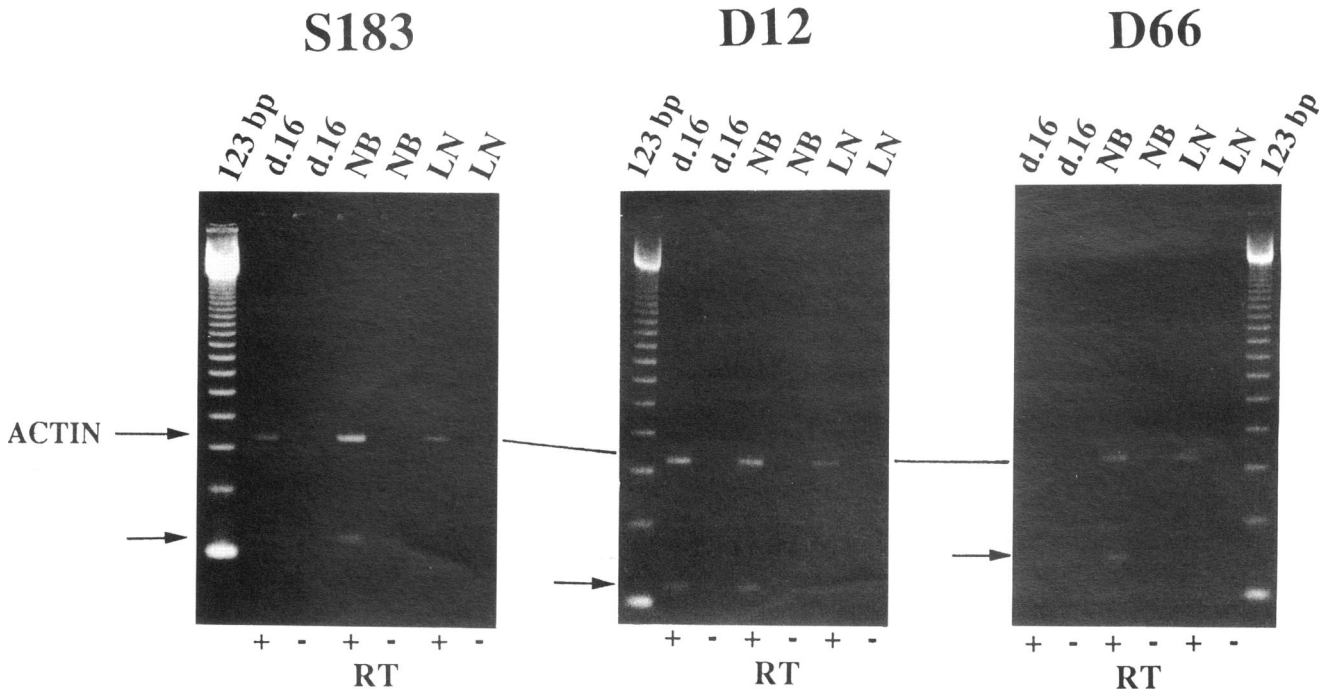


**Figure 5.** Reverse transcriptase (RT)/PCR-assisted amplification of RNA. RNA was isolated from day 16 AKR thymocytes, newborn thymocytes, and lymph node cells. cDNA was synthesized with random hexamer primers and AMV reverse transcriptase. The amount of input cDNA was then diluted and calibrated in PCR reactions containing the 1 pmol of the primers for beta-actin as an internal control. Specific primers were synthesized for individual cDNA clones and tested in multiplex RT/PCR reactions with variable cycle numbers to determine exponential amplification. Bands were identified by size in comparison to 123 bp markers. Arrows indicated the band produce by each set of specific primers. PCR primers used: murine beta-actin, 5′-CTGAAGTACCCCATTGAACATGGC-3′ (bA-1), 5′-AAGCTGTAGCCACGCTCGGTCAGG-3′ (bA-3); D66; 5′-TCCTCTGTGCTGGACAG-3′ (d66A), 5′-TGACAGCTTTGGCTAGGAGA-3′ (d66B); 5′-GGATTTTGACTAGACCAC-3′ (183A), 5′-ATTTGGAACACAGGATTT-3′ (183B); 5′-AGGTGTGCTGCTTCCTCT-3′ (D12A), 5′-GACTGGGTGCCAATCCTT-3′ (D12B).

tochondrial genes, murine repeats, and vector sequences (14%).

Of the remaining 155 cDNA clones, 27 exhibited limited DNA or protein similarities with the BLAST and FASTA sequence analysis programs (Table 2). Sequences corresponding to repeat regions were easily identified by the large number of significant DNA sequence matches that were found by the BLAST program. Protein similarity searches revealed the majority of contaminants, including a *E. coli* protein identified only by its protein sequence (97% identity/71 amino acids). A potentially significant and interesting nucleic acid similarity is that of TN-specific S19 to the rat calcium binding protein calbindin D28 (68%/274n.t.). Similarities to the translated protein sequences include the similarity of cDNA clone S596 to human cell cycle gene 1 (CCG1) (20%/116a.a.) and the similarity of TN-specific D66 to the fibroblast growth factor (FGF) receptor precursor (70%/54a.a.). The most interesting cDNAs are those that both sequence analysis programs

**Table 2.** Preliminary nucleic acid and protein similarities determined with the FASTA and BLAST sequence analysis programs

| | | | Database search programs | | |
|---|---|---|---|---|---|
| | | | BLAST | | FASTA |
| Clone number | Accession number | Description | p(N)[a] | Similarities[b] | Similarities |
| M517 | M94290 | h. lambda Ig | | | 64%/50n.t. |
| | | m. T-cell receptor alpha chain | | | 87%/23a.a. |
| S19 | M94291 | r. calbindin D28 | 6.3e-20 | 68%/158n.t. | 68%/274n.t. |
| S20 | M94292 | h. PDGF receptor precursor | 0.76 | 74%/27a.a. | |
| S65 | M94293 | m. leuk. common Ag (LCA) (T200) | | | 86%/22a.a. |
| S160 | M94294 | h. stem cell factor | 1.0 | 65%/23a.a. | 83%/23a.a. |
| S183 | M94295 | m. N-*ras* protein | 0.23 | 68%/60n.t. | |
| S191 | M94296 | m. LCA-related phosphatase | 0.49 | 75%/44n.t. | |
| | | transposase | 7.1e-10 | 69%/43a.a. | 86%/44a.a. |
| S267 | M94297 | r. profilaggrin | 0.27 | 68%/58n.t. | |
| | | m. astrocyte glial fib. acidic prot. | 0.0064 | 60%/35a.a. | |
| S334 | M94298 | m. eosinophil differentiation factor | 0.96 | 86%/30n.t. | |
| | | m. epidermal growth factor | | | 70%/94n.t. |
| S510 | M94299 | h. *flt* receptor-rel. tyrosine kinase | 0.39 | 64%/67n.t. | |
| S553 | M94300 | r. nucleolar protein B23 | 0.00026 | 75%/48n.t. | |
| | | r. nucleolar phosphoprotein | 0.75 | 94%/17a.a. | 16%/97a.a. |
| S596 | M94301 | h. cell cycle gene 1 (CCG1) | 0.45 | 76%/21a.a. | 20%/116a.a. |
| S618 | M94302 | h. growth hormone | 0.026 | 54%/125n.t. | |
| S711 | M94303 | h. erythrocyte ankyrin | 0.77 | 73%/26a.a. | |
| S843 | M94304 | m. C3 gene, 5′ end | | | 67%/58a.a. |
| S949 | M94289 | h. c-*fms* oncogene (CSF-1) receptor | 0.011 | 66%/74n.t. | |
| | | h. furin | | | 78%/32a.a. |
| S1006 | M94305 | dros. *shaggy* gene protein | 0.085 | 57%/33a.a. | 69%/49a.a. |
| S1014 | M94306 | electric eel sodium channel prot. | 0.0035 | 91%/12a.a. | |
| | | r. secretin receptor | | | 70%/40a.a. |
| S1059 | M94307 | von Willebrand factor precursor | 0.44 | 64%/31a.a. | |
| S1165 | M94308 | IMP dehydrogenase | 0.74 | 75%/16a.a. | |
| | | m. c.f. transmem. conductance reg. | | | 80%/49a.a. |
| D43 | M94379 | h. leukocyte adhesion alpha-chain | 0.58 | 76%/17a.a. | |
| D51 | M94309 | h. Na,K-ATPase | | | 57%/205n.t. |
| | | T-cell surface glycoprot. CD1.1 | | | 89%/18a.a. |
| D66 | M94310 | h. T-cell receptor beta-chain | | | 60%/73n.t. |
| | | T-cell receptor alpha-chain | 0.85 | 61%/31a.a. | |
| | | FGF-receptor precursor | 0.85 | 73%/19a.a. | 70%/54a.a. |
| D90 | M94311 | beta-adrenergic receptor | 0.23 | 66%/27a.a. | |

cDNA clones chosen for sequence analysis contained inserts >2 kb to avoid sequencing any cDNAs that may correspond to 3′ untranslated regions. For BLAST analysis, DNA analysis includes significant primate sequences p(N) < 0.50 and all amino acid similarities (identical plus conservative substitutions) p(N) < 1.0. For FASTA analysis, arbitrary clones with the best scores to primate sequences were chosen. Abbreviations used: n.t., nucleotide; a.a., amino acids; c.f. cystic fibrosis; h., human; r., rat, and m., mouse.

[a] The P-value is the probability of a score occurring by chance, given the size of the sequence searched and the database size using the BLAST program.

[b] Similarities include identical plus related substitutions.

(FASTA and BLAST) revealed like similarities between both the nucleic acid sequence and the protein translation. For example, M517 revealed nucleic acid similarities to human immunoglobulin (64%/50n.t.) and protein similarities to another member of the immunoglobulin gene superfamily, mouse T cell receptor alpha chain (87%/23a.a.). Of the remaining TN-specific cDNAs, no significant similarities were observed; however, intriguing similarities include D51 to CD1.1 (89%/18a.a.) and S39 to seven rodent Ig superfamily genes (59%/75n.t.). The remaining 80% of the cDNAs sequenced showed no significant similarities to known primate or murine genes and therefore represent unknown genes.

## DISCUSSION

### *Phenotype of the 705 Thymic Lymphoma*

The major aim of this study was to isolate cDNA clones encoding immature T cell-specific mRNAs. As a source of the TN T cell cDNA, the AKR thymic lymphoma 705 was used in subtractive hybridizations against lymph node mRNA. Flow cytometric analysis of the 705 thymic lymphoma reveals that this cell line is $CD4^-$, $CD8^-$, $CD3^-$, Thy-$1^+$, $HSA^+$, Pgp-$1^+$. The phenotype of this cell line is unusual in that Pgp-1 and HSA expression are generally considered to be mutually exclusive. However, minor populations of Pgp-$1^+$ and $HSA^+$ TN cells have been observed and may represent a transitional stage from precursor Pgp-$1^+$ cells to $HSA^+$ cells. Although the phenotypes of TN T cell populations are complex, $CD4^-$, $CD8^-$, $CD3^-$, Pgp-$1^+$ thymocytes have demonstrated progenitor activity in irradiated animals after intravenous injection of donor cells (Lesley, 1985), and $CD4^-$, $CD8^-$, $CD3^-$, $HSA^+$ thymocytes have also shown progenitor activity after intrathymic (it) injection (Crispe and Bevan, 1987). Indeed, all observed subpopulations of $CD4^-$, $CD8^-$, $CD3^-$ thymocytes have demonstrated progenitor activity in thymic reconstitution experiments and vary only by the kinetics and phenotypic stage achieved by their progeny (Scollay *et al.*, 1988; Petrie *et al.*, 1990). Although the exact stage or lineage of T cell the 705 thymic lymphoma represents is speculative, this lymphoma may represent adult prothymocytes because of Pgp-1 expression. The Pgp-1 antigen is normally found on most bone marrow cells and precursor T cells, but very little is found on mature lymph node T cells (Trowbridge *et al.*, 1982; Miyake *et al.*, 1990).

### *Subtractive Strategies used to Obtain Immature T Cell-Specific cDNAs*

A subtractive cDNA library was constructed to identify cDNAs encoding new proteins differentially expressed by TN T cells. Various methods to isolate differentially expressed proteins include the use of monoclonal antibodies, oligonucleotide probes based on partial amino acid sequence, and hybridization-based methods. Of the hybridization-based methods, the most popular are differential hybridization and subtractive hybridization. The limiting factor in these approaches is the sensitivity or the level of RNA that can be detected. Differential hybridization, or "plus-minus" screening, is based on the synthesis of complex labeled cDNA probes and their use to screen cDNA libraries. This technique is limited by the complexity of the cDNA probes and can only identify sequences present at 0.1–0.5% of the total mRNA population (Sargent, 1987). RNA-cDNA reassociation experiments suggest that the majority of T cell-specific mRNAs are present at levels lower than 0.1% (Hedrick *et al.*, 1984). Therefore, we chose to construct a subtractive cDNA library with a modification of a strategy that has proven successful in isolating differentially expressed genes present at 0.002% of the mRNA population (Palazzolo *et al.*, 1989). There are three advantages to the use of this cloning strategy and the lambda GEM (lambda SWAJ) vectors. First, the construction of the primary library enables the use of T7 RNA polymerase to synthesize antisense cDNA used in subtractive hybridizations. This in vitro generation of large amounts of antisense cDNA is helpful in situations, where the starting cells or tissue of interest can only be obtained in small amounts. Second, using the restriction sites to prime the synthesis of the subtracted cDNA facilitates the final cloning and recovery of the small amounts of cDNA remaining after subtraction. Finally, after the unsubtracted and contaminating clones were removed by the screening procedures, all of the remaining clones were considered for expression analysis and DNA sequence determination. Therefore the success of this strategy is based on choosing clones that fail to hybridize to labeled cDNA probes of high sequence complexity, thus retaining low abundance T cell-specific clones. The drawbacks of this cloning strategy are the complicated enzymatic reactions necessary to generate antisense cDNA and the use of oligo dT in library construction leads to a bias of 3′ untranslated sequences. An additional drawback is the obligate subcloning of the inserts into M13 before automated DNA sequence determination. In the future, cDNA subtraction coupled with large-scale cDNA sequencing will be a powerful tool in the isolation of tissue-specific or developmentally regulated genes.

### *Analysis of the Subtracted cDNAs*

After construction of the library, the cDNAs were analyzed by differential hybridization to identify cDNAs corresponding to abundant differentially expressed RNAs (D-clones). Over 100 differentially expressed cDNAs were isolated by this method, however, cross-hybridization studies revealed that approximately one-half of these cDNAs were MuLV retroviral sequences.

These retroviral sequences are a result of using an AKR thymic lymphoma as the source of TN T cells. Although unwanted, their presence in the library is expected and can be used as a gross estimation of the success of the subtraction procedures. Lymphotrophic MuLV have been shown to infect and replicate in immature cortical lymphocytes in the thymus (Cloyd, 1983; Owen *et al.*, 1986) and therefore may not be expressed at levels in mature T cells sufficient enough to eliminate them from the library.

What is the estimated diversity and complexity of the subtracted TN minus lymph node cDNA library? After two rounds of subtraction, the library contained 10 500 cDNA clones. With estimated percentages derived from the screening of the cDNAs, once 40% (4200) vector/unsubtracted B cell sequences and 10% (1500) MuLV sequences were removed, the library should contain 4 800 cDNA clones. DNA sequence analysis revealed that ≥30% (1400) of the remaining cDNAs corresponded to 3′ untranslated, yeast or mitochondrial sequences. Therefore a rough estimate of 3400 cDNAs should be present in the final library. Kinetic analysis by solution hybridization has revealed ~13 000 different mRNAs expressed by liver cells (Hastie and Bishop, 1976) and T hybridomas (Hedrick *et al.*, 1984). With 13 000 total mRNAs as an approximation for TN T cells, our results indicate that 26% (3 400/13 000) of the mRNA population is different between immature T cells and mature T cells. This result is higher than the 2% difference observed between T hybridomas and B cells and close to the 25% difference observed between lymphocytes and fibroblasts as suggested by solution hybridization experiments (Crampton *et al.*, 1980). Some of this difference may be explained by nuclear sequences or a possible redundancy once a larger number of clones are subjected to DNA sequence analysis. The rough estimate obtained may also reflect empirical differences between a clone by clone analysis versus solution hybridization experiments. Overall, if the lineage relationships between two cell types can be estimated by the percent difference in their expressed mRNAs, our results suggest that immature TN T cells may express a greater diversity of mRNAs than their more mature progeny (greater than the mRNA diversity previously observed between T cells and B cells).

After hybridizations to screen and sort the cDNA clones, 97 random cDNAs were tested for tissue-specific expression by Northern blot analysis with RNA from various lymphoid and nonlymphoid cell types. The results from the Northern analysis indicated that in >50% of the cDNAs tested, no hybridization could be observed. Moreover, the majority of the cDNA clones that were expressed in a TN T cell-specific expression pattern were S-clones or the clones that failed to hybridize to any of the probes used to screen the library. This data suggest that the S-clones may contain the TN T cell-specific low abundance genes we set out to isolate. Nevertheless, all of the known surface proteins expressed by TN T cells (e.g., Pgp-1, HSA) are also found on a variety of other cell types. Therefore cDNA clones that are additionally expressed on the other cell types tested, such as monocytes and pro-B cells, are also of interest. Although only a small number of cDNAs were used to identify RNA transcripts by RT/PCR detection, the results suggest that many of the sequences may be expressed by adult rather than fetal (day 16) TN T cells. This implies that fetal and adult populations of TN T cells may be different at the level of gene expression. Overall, a combination of RNA expression pattern and DNA sequence similarities will be used to determine which clones will be subjected to further study.

### *DNA Sequence Analysis of Subtracted cDNAs*

Three hundred thirty cDNA clones representing all of the observed expression patterns were subject to single-run automated DNA sequence determination and analysis. Both BLAST and FASTA sequence analysis programs were used to combine the advantages of each program and to benefit from the use of different algorithms to assess the significance of the DNA and protein sequences generated. Comparisons between these programs revealed some interesting features and limitations of each program. The advantages of the BLAST program were its speed (usually 10–20 s), its ability to find short stretches of similarity to known genes, and the estimation of a Poisson p-value for each match. However, the inability of the BLAST program to introduce gaps to improve the alignment made it practical for primarily finding matches to genes already in the databases. Thus the BLAST program was used to identify the 175 unwanted sequences, and FASTA was used to analyzed the remaining 155 unknown sequences. Overall, the FASTA program was more useful because of its ability to introduce gaps and find partial similarities over large stretches of sequences (e.g., S596). None of the cDNAs sequenced were similar to known TN T cell genes such as Pgp-1/CD44, TdT, or RAG-1 genes. Failure to identify these genes may be due to the statistical limitations of sequencing 155 clones, or they may be expressed at some level in the mature lymph node lymphocytes used for subtraction and would thus have been removed. Several cDNA sequences had potentially significant DNA similarities to various genes including, calbindin D28 (S19), epidermal growth factor (S334), and Na,K-ATPase (D51). Interestingly many of the observed similarities were to genes expressed in the brain (S19, S267, S1165, D90). This confirms other observed similarities between the nervous system and the immune system at the level of gene expression and in the use of common motifs for cell-cell recognition (Gilbert and Payan, 1991; Grenningloh, *et al.*, 1991). Overall, the significance of the similarities presented in Table 2 are unknown and will require further sequence information to confirm the relationships to other genes.

Because protein sequence database searches are more sensitive than DNA searches, cDNA clones chosen for further analysis will be clones with significant protein similarities. Interesting protein similarities include, T-cell receptor alpha chain (M517 and D66), PDGF receptor precursor (S20), stem cell factor (S160), and the Drosophila *shaggy* gene (S1006). The isolation of a cDNA with similarity to the Drosophila *shaggy* gene is potentially interesting because the product of *shaggy* is a serine/threonine protein kinase that functions in a signal transduction pathway necessary for cell-cell communication (Bourouis *et al.*, 1990). In addition, all of the cDNAs with partial similarities to members of the immunoglobulin superfamily will be tested by further DNA sequence determination. This large family contains many diverse members with important roles in cell-cell recognition, and the identification of a new member expressed by immature T cells may lead to new insights into the processes of cellular recognition during T cell development (Hood *et al.*, 1985; Williams and Barclay, 1988).

In this study, a combination of subtraction strategies and differential hybridizations were employed to enrich for cDNA clones that may be expressed during T cell development in the thymus. Eight cDNA clones were identified that are expressed by TN T cells on Northern blots and ~11 cDNAs have interesting similarities to known genes. Further experiments include complete DNA sequence determination and further RNA expression analysis to determine the kinetic parameters and tissue distribution of each clone. In addition, these cDNAs along with the production of antibody reagents will be used to explore the cellular locations of their corresponding proteins. Antibody reagents will also be employed to further delineate subpopulations of $CD4^-$, $CD8^-$, $CD3^-$ T cells. Thus DNA sequence data and cellular distribution data may suggest further experiments to determine their functional roles during T cell development and differentiation in the thymus.

## ACKNOWLEDGMENTS

We gratefully acknowledge Linda Sakimura and Julie Bars for expert technical help. We thank Dr. Joan Kobori and Dr. Carmie Puckett for critical reading of the manuscript, and Dr. Ellen Rothenberg for the kind gift of the IM7.8.1 and M1.69 antibodies. Also Dr. Ellen Rothenberg and Dr. Jim Allison for the gift of the 705 cell line. We thank Rochelle Diamond and Pat Koen for expert technical assistance for the cytofluorographic analysis and members of the Caltech Division of Biology Sequence Analysis Facility for assistance with the VAX computer and the GCG software.

This work was supported by National Institutes of Health grant HD-07257 and the Lucille P. Markey Charitable Trust (S. Orr).

## REFERENCES

Adkins, B., Mueller, C., Okada, C.Y., Reichert, R.A., Weissman, I.L., and Spangrude, G.J. (1987). Early events in T-cell maturation. Annu. Rev. Immunol. *5*, 325–365.

Altschul, S.F., Gish, W., Miller, W., Myers, W.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403–410.

Bourouis, M., Moore, P., Ruel, L., Grau, Y., Heitzler, P., and Simpson, P. (1990). An early embryonic product of the gene shaggy encodes a serine/threonine protein kinase related to the CDC28/cdc2+ subfamily. EMBO J. *9*, 2877–2884.

Brunet, J.F., Dosseto, M., Denizot, F., Mattei, M.G., Clark, W.R., Haqqi, T.M., Ferrier, P., Nabholz, M., Schmitt-Verhulst, A.M., Luciani, M., and Golstein, P. (1986). The inducible cytotoxic T-lymphocyte associated gene transcript CTLA-1: sequence and gene localization to mouse chromosome 14. Nature *322*, 268–270.

Burd, P.R., Freeman, G.J., Wilson, S.D., Berman, M., Dekruyff, R., Billings, P.R., and Dorf, M. (1987). Cloning and characterization of a novel T cell activation gene. J. Immunol. *139*, 3126–3131.

Chomczynski, P., and Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. *162*, 156–159.

Cloyd, M.W. (1983). Characterization of target cells for MCF viruses in AKR mice. Cell *32*, 217–225.

Crampton, J., Humphries, S., Woods, D., and Williamson, R. (1980). The isolation of cloned cDNA sequences which are differentially expressed in human lymphocytes and fibroblasts. Nucleic Acids Res. *8*, 6007–6017.

Crispe, I.N., and Bevan, M.J. (1987). Expression and functional significance of the J11d marker on mouse thymocytes. J. Immunol. *138*, 2013–2018.

Davies, L.G., Dibner, M.D., and Battey, J.F. (1986). Formaldehyde gel for electrophoretic separation of RNA and Northern blot. In: Basic Methods in Molecular Biology, New York: Elsevier, 143–146.

Davis, M.M. (1986). Subtractive cDNA hybridization and the T-cell receptor genes. In: Handbook of Experimental Immunology, Oxford, UK: Blackwell Scientific, 76.1–76.13.

Davis, R.W., Botstein, D., and Roth, J.R. (1980). In: Advanced Bacterial Genetics, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory, 106–107.

Devereux, J., Haeberli, P., and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. *12*, 387–395.

Feinberg, A.P., and Vogelstein, B. (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. *132*, 6–13.

Fowlkes, B.J., Edison, L., Mathieson, B.J., and Chused, T.M. (1985). Early T lymphocytes: differentiation in vivo of adult intrathymic precursor cells. J. Exp. Med. *162*, 802–823.

Fowlkes, B.J., and Pardoll, D.M. (1989). Molecular and cellular events of T cell development. Adv. Immunol. *44*, 207–264.

Gerard, G.F. (1988). Synthesis of high specific activity cDNA. Focus *10*, 12–14.

Gershenfeld, H.K., and Weissman, I.L. (1986). Cloning of a cDNA for a T cell-specific serine protease from a cytotoxic T lymphocyte. Science *232*, 854–858.

Gilbert, M.S., and Payan, D.G. (1991). Interactions between the nervous and immune systems. Front. Neuroendocrinol. *12*, 299–322.

Grenningloh, G., Rehm, E.J., and Goodman, C.S. (1991). Genetic analysis of growth cone guidance in drosophila: fasciclin II functions as a neuronal recognition molecule. Cell *67*, 45–57.

Gribskov, M., Devereux, J., and Burgess, R.R. (1984). The codon preference plot: graphic analysis of protein coding sequences and prediction of gene expression. Nucleic Acids Res. *12*, 539–549.

Hastie, N.D., and Bishop, J.O. (1976). The expression of three abundance classes of messenger RNA in mouse tissues. Cell *9*, 761–774.

Hedrick, S.M., Cohen, D., Nielsen, E.A., and Davis, M.M. (1984). Isolation of cDNA clones encoding T cell-specific membrane-associated proteins. Nature *308*, 149–153.

Hood, L., Kronenberg, M., and Hunkapillar, T. (1985). T cell antigen receptors and the immunoglobulin supergene family. Cell *40*, 225–229.

Houchins, J.P., Yabe, T., McSherry, C., Miyokawa, N., and Bach, F.H. (1990). Isolation and characterization of NK cell or NK/T cell-specific cDNA clones. J. Mol. Cell. Immunol. *4*, 295–306.

Jongstra, J., Schall, T.J., Dyer, B.J., Clayberger, C., Jorgensen, J., Davis, M.M., and Krensky, A. (1987). The isolation and sequence of a novel gene from a human functional T cell line. J. Exp. Med. *165*, 601–614.

Kavathas, P., Sukhatme, V.P., Herzenberg, L.A., and Parnes, J.R. (1984). Isolation of the gene encoding the human T-lymphocyte differentiation antigen Leu-2 (T8) by gene transfer and cDNA subtraction. Proc. Natl. Acad. Sci. USA *81*, 7688–7692.

Lesley, J., Hyman, R., Schulte, R. (1985). Evidence that Pgp-1 glycoprotein is expressed on thymus-homing progenitor cells of the thymus. Cell. Immunol. *91*, 397–403.

Maddon, P.J., Littmen, D.R., Godfrey, M., Maddon, D.E., Chess, L., and Axel, R. (1985). The isolation and nucleotide sequence of a cDNA encoding the T cell surface protein T4: a new member of the immunoglobulin gene family. Cell *42*, 93–104.

Mechler, B., and Rabbitts, T. (1981). Membrane-bound ribosomes of myeloma cells IV. mRNA complexity of free and membrane-bound polysomes. J. Cell Biol. *88*, 29–36.

Miyake, K., Median, K.I., Hayashi, S., Ono, S., Hamaoka, T., and Kincade, P.W. (1990). Monoclonal antibodies to Pgp-1/CD44 block lympho-hemopoiesis in long-term bone marrow cultures. J. Exp. Med. *171*, 477–488.

Moore, M.A.S., and Owen, J.J.T. (1967). Experimental studies on the development of the thymus. J. Exp. Med. *126*, 715–726.

Owen, F.L., Strauss, W.M., Murre, C., Duby, A.D., Hiai, H., and Seidman, J.G. (1986). AKR murine thymic leukemias are from a distinct thymic cell lineage and do not express the beta chain of the T-cell antigen receptor. Proc. Natl. Acad. Sci. USA *83*, 7434–7437.

Palazzolo, M.J., Hyde, D.R., Raghavan, K.V., Mecklenburg, K., Benzer, S., and Meyerowitz, E.M. (1989). Use of a new strategy to isolate and characterize 436 Drosophila cDNA clones corresponding to RNAs detected in adult heads but not in early embryos. Neuron *3*, 527–539.

Palazzolo, M.J., and Meyerowitz, E.M. (1987). A family of lambda phage cDNA cloning vectors, lambdaSWAJ, allowing the amplification of RNA sequences. Gene *52*, 197–206.

Pearson, W.R., and Lipman, D.J. (1988). Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA *85*, 2444–2448.

Petrie, H.T., Hugo, P., Scollay, R., and Shortman, K. (1990). Lineage relationships and developmental kinetics of immature thymocytes: CD3, CD4 and CD8 acquisition in vivo and in vitro. J. Exp. Med. *172*, 1583–1588.

Richie, E.R., McEntire, B., Crispe, N., Kimura, J., Lanier, L.L., and Allison, J.P. (1988). alpha/beta T-cell antigen receptor gene and protein expression occurs at early stages of thymocyte differentiation. Proc. Natl. Acad. Sci. USA *85*, 1174–1178.

Rupp, R.A., and Weintraub, H. (1991). Ubiquitous MyoD transcription at the midblastula transition precedes induction-dependent MyoD expression in presumptive mesoderm of *X. Laevis*. Cell *65*, 927–937.

Sargent, T. (1987). Isolation of differentially expressed genes. In: Guide to Molecular Cloning Techniques, ed. S.L. Berger, and A.R. Kimmel, San Diego, CA: Academic, 423–432.

Schall, T.J., Jongstra, J., Dyer, B.J., Jorgensen, J., Clayberger, C., Davis, M.M., and Krensky, A.M. (1988). A human T cell-specific molecule is a member of a new gene family. J. Immunol. *141*, 1018–1025.

Scollay, R., Smith, J., and Stauffer, V. (1986). Dynamics of early T cells: prothymocyte migration and proliferation in the adult mouse thymus. Immunol. Rev. *91*, 129–157.

Scollay, R., Wilson, A., D'Amico, A., Kelly, K., Egerton, M., Pearse, M., Wu, L., and Shortman, K. (1988). Developmental status and reconstitution potential of subpopulations of murine thymocytes. Immunol. Rev. *104*, 81–120.

Staden, R. (1984). Measurements of the effects that coding for a protein has on a DNA sequence and their use for finding genes. Nucleic Acids Res. *12*, 551–567.

Trowbridge, I.S., Lesley, J., Schulte, R., Hyman, R., and Trotter, J. (1982). Biochemical characterization and cellular distribution of a polymorphic, murine cell-surface glycoprotein expressed on lymphoid tissues. Immunogenetics *15*, 299–312.

Van Ness, J., and Hahn, W.E. (1982). Physical parameters affecting the rate and completion of RNA driven hybridization of DNA: new measurements relevant to quantitation based on kinetics. Nucleic Acids Res. *10*, 8061–8077.

Williams, A.F., and Barclay, A.N. (1988). The immunoglobulin superfamily-domains for cell surface recognition. Annu. Rev. Immunol. *6*, 381–405.

Wilson, A., Ewing, T., Owens, T., Scollay, R., and Shortman, K. (1988). T cell antigen receptor expression by subsets of Ly-2$^-$, L3T4$^-$ (CD4$^-$, CD8$^-$) thymocytes. J. Immunol. *140*, 1470–1476.

Yanagi, Y., Chan, A., Chin, B., Minden, M., and Mak, T. (1985). Analysis of cDNA clones specific for human T cells and the alpha and beta chains of the T-cell receptor heterodimer from a human T-cell line. Proc. Natl. Acad. Sci. USA *82*, 3430–3434.

Zipfel, P.F., Irving, S.G., Kelly, K., and Siebenlist, U. (1989). Complexity of the primary genetic response to mitogenic activation of human T-cells. Mol. Cell. Biol. *9*, 1041–1048.