Identification of a Novel *Ly49* Promoter That Is Active in Bone Marrow and Fetal Thymus[1]

Ali Saleh,* Andrew P. Makrigiannis,* Deborah L. Hodge,* and Stephen K. Anderson[2†]

The analysis of several *Ly49* genes has identified a tissue-specific promoter adjacent to the previously defined first exon. The current study reveals the presence of an additional *Ly49* promoter (Pro-1) and two noncoding exons upstream of the previously defined promoter (Pro-2). DNA sequences homologous to Pro-1 are present 4–10 kb upstream of Pro-2 in all *Ly49* genes examined, and Pro-1 transcripts were detected from the *Ly49a*, *e*, *g*, *o*, and *v* genes. Pro-1 activity can be detected in bone marrow, embryonic thymus, freshly isolated liver NK cells, and the murine LNK cell line, but it does not function in adult thymus, sorted NK-T cells, spleen NK cells, or the EL-4 T cell line, even though these cells express Ly49 proteins. Luciferase reporter assays identified a Pro-1 core promoter region that functions in the LNK cell line but not EL-4 cells. The novel promoter is not active in mature NK cells, suggesting that Pro-1 represents an early *Ly49* promoter. *The Journal of Immunology,* 2002, 168: 5163–5169.

The murine Ly49 family of class I MHC receptors represents the functional analogs of the human killer cell Ig-related receptor gene family. Although they constitute structurally distinct protein families, the C-type lectin-related Ly49 proteins and the killer cell Ig-related receptor proteins associate with identical signaling molecules to achieve either activation or inhibition of NK cells in response to specific MHC class I ligands. The *Ly49* gene cluster has been mapped in the C57BL/6 (B6) mouse genome (1–3), and the genes are arranged in tandem with an identical transcriptional orientation. With the exception of the *Ly49 h*, *k*, and *n* cluster of highly related activators, the *Ly49* genes are not grouped with respect to activating/inhibitory function or gene homology. Of the 16 *Ly49* genes currently identified in the B6 genome, *Ly49a–j*, and *q* have been shown to produce mRNAs with a complete coding region, *Ly49k*, *m*, and *n* represent transcribed pseudo genes, and the remaining two genes (*Ly49l* and *v*) do not appear to produce transcripts in the B6 strain (4–8) (S. K. Anderson, unpublished observations).

The activation of *Ly49* genes during murine NK cell development presents an interesting system for the study of selective gene expression. It has been proposed that a stochastic process controls the activation of *Ly49* gene transcription, because the proportion of NK cells that express two Ly49 proteins is roughly equivalent to the product of the proportion of NK cells expressing the individual receptors (9). Single cell RT-PCR analysis of Ly49 expression has shown that the majority of NK cells express from one to four different receptors per cell and NK cells with five or more Ly49 proteins are extremely rare, supporting the theory that the *Ly49* genes are activated by a probabilistic mechanism (10).

The promoter region responsible for the expression of *Ly49a* in mature NK/NK-T cells has been cloned, and the role of *cis*-acting elements in gene activation and cell-specific transcription has been studied (11–13). The *trans*-acting T cell-specific factor-1 has been shown to be required for the acquisition of Ly49A expression during development (13). The activating transcription factor-2 binds to a 13-bp element adjacent to the predicted TATAA of *Ly49a*, and it is the major factor responsible for *Ly49a* transcription in EL-4 cells (12). A study of the 5′ region of the *Ly49i* gene identified a core promoter that is preferentially transcribed in a Ly49-expressing cell line, EL-4 (14). Surprisingly, sequence elements immediately upstream from the core promoter were found to inhibit activity, and no enhancers were detected in the upstream region analyzed. An inhibitory upstream element was also detected in a functional analysis of the *Ly49c* and *j* promoters (15). The lack of detectable enhancer elements immediately upstream of these promoters suggests that additional control elements may exist elsewhere. The current knowledge of *Ly49* promoter structure and function has not provided any insight with respect to the mechanisms responsible for the stochastic activation of *Ly49* genes.

Extensive screening of a 129/J liver NK cell cDNA library resulted in the identification of several novel *Ly49* genes, including an unusual cDNA clone of *Ly49g* containing an alternative first exon (16). This result indicated the presence of an alternative *Ly49* promoter. This study describes the isolation and characterization of the novel promoter from several *Ly49* genes.

Materials and Methods

Cell lines

EL-4 and P815 cell lines were cultured in RPMI 1640 supplemented with 10% FBS, 100 U/ml penicillin, and 100 U/ml streptomycin. The LNK cell line (17) was cultured in RPMI 1640 containing 2-ME, nonessential amino acids, 10% FBS, 100 U/ml penicillin, 100 U/ml streptomycin, sodium pyruvate, L-glutamine, HEPES, and IL-2 (8000 IU/ml).

Cell isolation

Mouse NK cells were purified from livers or spleens of 129/J or B6 mice. Liver NK cells were isolated as previously described (18). Routinely, 75–

*Laboratory of Experimental Immunology, Center for Cancer Research, and †Intramural Research Support Program, SAIC-Frederick, National Cancer Institute, Frederick, MD 21702

Received for publication November 15, 2001. Accepted for publication March 12, 2002.

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

[1] This project has been funded in whole or in part with Federal funds from the National Cancer Institute, National Institutes of Health, under contract no. NO1-CO-12400. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. By acceptance of this article, the publisher or recipient acknowledges the right of the U.S. Government to retain a nonexclusive, royalty-free license in and to any copyright covering the article.

[2] Address correspondence and reprint requests to Dr. Stephen K. Anderson, SAIC-Frederick, National Cancer Institute-Frederick Cancer Research and Development Center, Building 560, Room 31-93, Frederick, MD 21702-1201. E-mail address: andersn@ncifcrf.gov

0022-1767/02/$02.00

80% of the resulting cells were NK1.1$^+$. Spleen cells were sorted for DX5$^+$CD3$^-$ NK cells or DX5$^+$CD3$^+$ NK-T cells on a MoFlo Cell Sorter (Cytomation, Ft. Collins, CO). Bone marrow cells were isolated from the femurs and tibias of B6 mice. After lysis of RBCs, the remaining cells were washed with PBS and counted. Total liver and thymus cell suspensions were isolated from day-15 B6 embryos. Animal care was provided in accordance with the procedures outlined in "A Guide for the Care and Use of Laboratory Animals" (National Institutes of Health, publication no. 86-23, 1985).

RT-PCR analysis of promoter use

Cellular poly(A)$^+$ RNA was isolated using the QuickPrep mRNA kit (Amersham Pharmacia Biotech, Piscataway, NJ). Oligo(dT)-primed cDNA was synthesized using the Superscript cDNA synthesis kit (Invitrogen, Carlsbad, CA). One microliter of cDNA was amplified using 35 cycles of 94°C for 30 s, 58°C for 30 s, and 72°C for 30 s with specific primers using Platinum PCR Supermix (Invitrogen). The antisense primer used to detect promoter (Pro)[3]-1 activity was from a conserved region of exon 4 at position 261 of the *Ly49* coding sequence (5′-TCACTTTGCATGTTGCTG CAG). A degenerate antisense primer spanning the *Ly49* termination codon (5′-CTSRTTGGARARTYAATSAGG) was used to amplify and clone the complete coding region. The −1a exon forward primers were 5′-GCAA GTCCTCCCAGTGAAGC and 5′-TTGCAGCTGTGCAATGATAG, and the exon 1 forward primer was 5′-GAGGTTGAGTATCACTCAGTGG.

RNase protection assays

The RNase protection assay (RPA) was performed according to the manufacturer's protocol using an RPA kit (RPA II; Ambion, Austin, TX). To localize the transcription initiation site of the *Ly49*g alternative promoter in EL-4 cells, an antisense *Ly49g* RNA probe was generated from the *Ly49g* Pro-1 clone in pCR2.1-TOPO linearized with *Hin*dIII. In vitro transcription was performed according to the manufacturer's protocols using the T7 MAXIscript (Ambion) in vitro transcription kit. The [α-^{33}P]RNA was separated on a 6% denaturing polyacrylamide gel and the full-length 500-bp probe was excised and eluted by overnight incubation at 37°C in gel elution buffer (Ambion). Approximately 2×10^5 cpm of the gel-purified *Ly49*g-specific probe was added to 1 μg of poly(A)$^+$ mRNA from either LNK or EL-4 cells in each reaction mixture. The protected mRNA products were separated on a 6% denaturing polyacrylamide gel. The gel was dried and exposed for ~72 h in a PhosphorImager cassette (Molecular Dynamics, Sunnyvale, CA). The image was visualized using PhosphorImager SI analysis and ImageQuaNT (Molecular Dynamics).

Generation of luciferase reporter plasmids

The Ly49i-Pro1 construct was generated by PCR from a 5-kb *Pst*I fragment of the *Ly49i* promoter region (14) using a reverse primer within the predicted exon −1a (5′-AACCTGTCATAGTCACAGCC) and a T7 primer from the vector. The 330-bp PCR product was subcloned into the pCR-XL-TOPO vector (Invitrogen). A 400-bp *Xho*I/*Hin*dIII fragment was purified from pCR-XL-TOPO and inserted into the pGL3-basic vector (Promega, Madison, WI). To generate the Ly49i-3′-Pro1 construct, Ly49i-Pro1 was digested with *Sma*I and *Bst*Z17I to remove the 5′ region of the promoter, and the remaining vector was religated. The Ly49i-5′-Pro1 construct was generated by digestion of Ly49i-Pro1 with *Bst*Z17I and *Hin*dIII to remove the 3′ region, and the remaining vector was religated using Klenow to fill in the *Hin*dIII site. To generate a *Ly49j* construct similar to Ly49i-altpro, PCR was performed on a B6 BAC clone (RPCI-23 416H5) with primers corresponding to the 5′ and 3′ ends of Ly49i-altpro (forward, CCTGCTATCATGTTTATTACATTCC; reverse, GGGAGGACTTGCT TATCCTG). The PCR product was subcloned into pCR2.1-TOPO and then transferred to pGL3-basic using *Sac*I and *Xho*I, resulting in the Ly49j-altpro luciferase construct. The *Ly49g* alternative promoter region was isolated by PCR from a 129/J BAC clone (RPCI-22, 10L7) using primers (forward 5′-TTCCTTGCACTCAGTGCTTG; reverse 5′-CAGTCAGAC CCTTGGACTG) and cloned into pCR2.1-TOPO. The cloned insert was shown to contain the alternative promoter region of *Ly49g* (Fig. 1*B*). A fragment containing the core promoter region was generated using primers (forward 5′-CAGGGTAAATACAAGAACAGG; reverse 5′-CCAGTG TCCCACCTCTTGG) and cloned into pCR2.1-TOPO. The *Ly49g* Pro-1 core was transferred to pGL3-basic using *Hin*dIII and *Xho*I, resulting in the Ly49g-Pro1 luciferase construct. The Ly49a-Pro2 construct was generated by ligating a *Xba*I-*Eco*RV fragment of the *Ly49a* promoter region (12) into pGL3-basic. The Ly49a-2kb construct was generated by PCR from a B6 BAC clone (RPCI-23 416H5) using the following primers: forward, CAATCTATTAGGAGTCAGGC; reverse, CCACTGAGAGATACTCA CCTC. Ly49a-2kb+Pro-1 was generated by inserting the *Ly49g* Pro-1 core in front of the 2-kb region using *Sac*I and *Xho*I. All constructs were verified by sequencing with specific primers. Sequence analysis was performed with the SeqWeb package at the Frederick Cancer Reaearch and Development Center supercomputing center (Frederick, MD).

Cell transfection and luciferase assays

A total of 4×10^6 P815 cells in RPMI medium without serum were transfected with 10 μg of the individual reporter constructs by electroporation using a GenePulser (Bio-Rad, Richmond, CA) set at 240 mV, 960 μF. EL-4 or LNK cells were plated at 3×10^5 cells per well in a six-well plate and transfected with 1 μg of the individual reporter constructs plus 0.1 μg of the *Renilla* luciferase pRL-SV40 control DNA using Lipofectamine (Life Technologies, Gaithersburg, MD) according to the manufacturer's protocol. Luciferase activity was assayed using the Dual-Luciferase Reporter Assay System (Promega) according to the manufacturer's instructions. Cells were washed with PBS and lysed in 500 μl of passive lysis buffer. A clear lysate was obtained by centrifugation at 12,000 rpm for 30 s. A total of 20 μl of the supernatant was added to 100 μl of luciferase assay reagent and the firefly luciferase activity was read. To measure the activity of the *Renilla* luciferase control, 100 μl of Stop and Glow reagent (Promega) was added and the sample was read. The luciferase activity of the *Ly49* promoter constructs was normalized relative to the activity of the *Renilla* luciferase produced by the pRL-SV40 control vector.

Results

Identification of novel 5′ exons in the Ly49 *gene family*

The studies of *Ly49* promoter function reported to date have focused on the region preceding the 5′ end of cDNAs isolated from mature NK cells. The isolation of a variant of *Ly49g* from 129/J liver NK cells that contained a novel 5′ untranslated region suggested that an additional promoter existed. A search of GenBank with the novel sequence yielded significant matches to a region 5′ of the *Ly49a* promoter and two regions upstream of the *Ly49i* promoter. This result indicated the presence of an additional promoter as well as an additional noncoding exon upstream of the *Ly49g* promoter. Three B6 BAC clones were identified (GenBank accession nos. AC087336, AC090127, and AC090563) that contained the intergenic sequences preceding the *Ly49a*, *c–f*, *h–k*, *m*, and *n* genes. Sequences corresponding to the predicted novel exons were found in all *Ly49* genes examined; however, this sequence was less homologous in the activating *Ly49* genes. We have named the novel exons as exon −1a and exon −1b. For the purposes of this discussion, the novel promoter is designated as Pro-1. The promoter responsible for the majority of *Ly49* transcripts in mature NK cells is referred to as Pro-2, and the promoter located next to exon 2 recently described by McQueen et al. (15) is referred to as Pro-3. The relative position of each promoter is shown in Fig. 1*A*. Primers from conserved regions of the available Pro-1 sequences were used to isolate the Pro-1 region of the 129/J *Ly49g* and *j* genes by PCR. The nucleotide sequences of the novel promoter and downstream exon regions of *Ly49a*, *d*, *e*, *g*, *i*, and *j* are presented in Fig. 1, *B* and *C*. Exon −1b was flanked by consensus splice donor and acceptor sequences in the *Ly49a*, *d*, *g*, *h*, *k*, and *n* genes, but the splice acceptor sequence of the putative exon −1b was altered in the *Ly49c*, *e*, *f*, *i*, and *j* genes, suggesting that it is nonfunctional in the *Ly49e/c* subfamily (Fig. 1*C*). The genomic DNA sequence preceding the Pro-1 region of the activating *Ly49* genes was found to be significantly different from that of the inhibitory *Ly49*s, whereas the region upstream of the previously described Pro-2 is highly conserved among all *Ly49*s (19). The Pro-1 sequence was compared against the available rat genomic *Ly49* sequences (GenBank accession no. AC103500; Fig. 1*B*). The Pro-1 region was conserved to the same extent as the sequences surrounding exons (>80%), while intron sequences were less conserved (60%), suggesting that Pro-1 represents an important functional element that has been conserved between the two species.

[3] Abbreviations used in this paper: Pro, promoter; RPA, RNase protection assay.

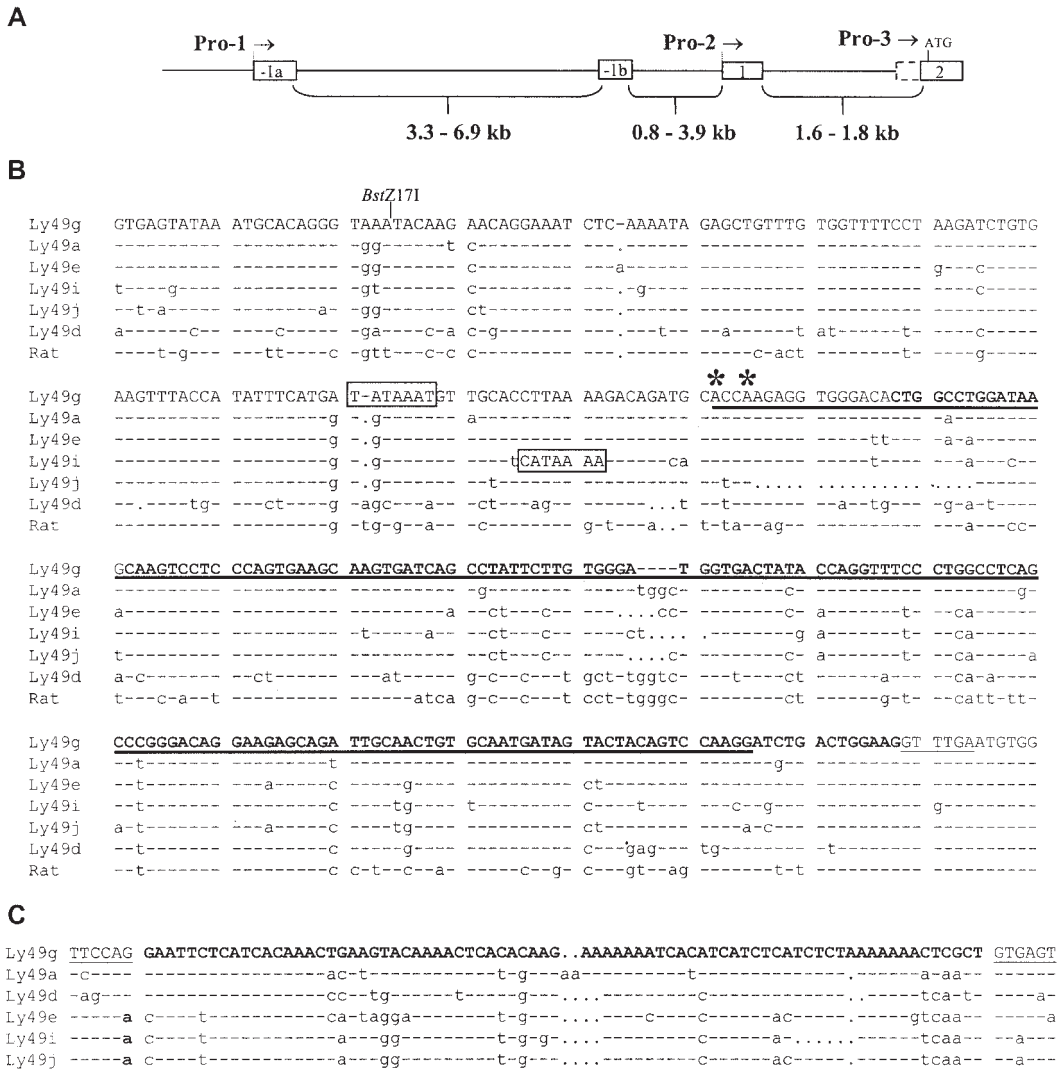

FIGURE 1. Identification of a novel promoter in the *Ly49* genes. *A*, Location of exons −1a, −1b, 1, and 2 in *Ly49* genes. The sequence of the *Ly49a*, *c–k*, *m*, and *n* genes was analyzed and the location of regions homologous to the novel *Ly49g* cDNA was determined. Predicted exons are shown as boxes labeled with their respective exon designation (−1a, −1b, 1, 2). The range of intron sizes among the *Ly49* genes is shown below each segment. The location of the three *Ly49* promoters identified (Pro-1, Pro-2, and Pro-3) is indicated. The solid arrow indicates the transcription initiation site of each promoter. *B*, A comparison of the nucleotide sequence of the Pro-1/exon −1a region of several *Ly49* genes. The *Ly49g* sequence is shown, with only nucleotide differences shown for the *Ly49a*, *d*, *e*, *i*, and *j* genes. Dashed lines indicate conserved residues, and gaps are indicated by dots. The *Ly49g* sequence has been deposited in GenBank (accession no. AF421168). The *Ly49a*, *d*, *e*, *i*, and *j* sequences are from B6 BAC sequences (GenBank accession nos. AC087336, AC090127, and AC090563). A Pro-1-related region from a rat *Ly49* (GenBank accession no. AC103500) is shown to demonstrate the evolutionary conservation of the Pro-1 element. The *Ly49g* TATAAA element and the CATAAA element of *Ly49i* are boxed. Bold nucleotides in the *Ly49g* sequence represent sequence found in the *Ly49g* cDNA (GenBank accession no. AF419249). The region of the *Ly49g* mRNA protected by the antisense RNA probe (Fig. 3) is indicated by the thick line. The predicted transcription start sites at positions 141 and 144 are marked by asterisks. *C*, Nucleotide sequence of the exon −1b region. The *Ly49g* cDNA sequence is shown in bold, with nucleotide differences shown for the *Ly49a*, *d*, *e*, *i*, and *j* genes. Predicted splice donor and acceptor sequences are underlined in the *Ly49g* gene. The G to A nucleotide substitution that disrupts the splice acceptor signal in *Ly49e*, *i*, and *j* is shown in bold.

Detection of Ly49 mRNAs containing exon −1a

To survey cell lines and tissues for Pro-1 activity, an oligonucleotide primer from a conserved region of exon −1a was used in conjunction with an antisense primer from a conserved region of exon 4 to perform RT-PCR on mRNA isolated from bone marrow, thymus, spleen NK, liver NK, NK-T cells, the LNK cell line, and the EL-4 cell line. Fig. 2*A* shows the results of an RT-PCR experiment performed with the exon −1a forward primer as well as a control exon 1 forward primer on mRNAs isolated from 6- to 8-wk-old B6 mice. Purified liver NK cell mRNA generated *Ly49* PCR products corresponding to spliced transcripts containing exon −1a or exon 1, whereas purified NK-T cell mRNA produced only Pro-2-derived PCR products. Transcripts derived from Pro-1 were detected in bone marrow and liver NK cells, but not in thymus or spleen NK cells. Several PCR products were cloned and sequenced from each reaction to verify that the bands observed represented Pro-1 or Pro-2 transcripts. In the experiment shown in Fig. 2*A*, all cloned Pro-1 PCR products represented *Ly49g* transcripts. Previous PCR cloning experiments have shown a bias toward *Ly49g* amplification (16); therefore, it is possible that Pro-1 transcripts for other inhibitory *Ly49* genes are present, but they are not readily isolated using the PCR approach. To demonstrate the existence of a Pro-1 transcript for the *Ly49a* and *o* genes, Pro-1 forward and exon 4 reverse primers specific for *Ly49a/o* were synthesized and

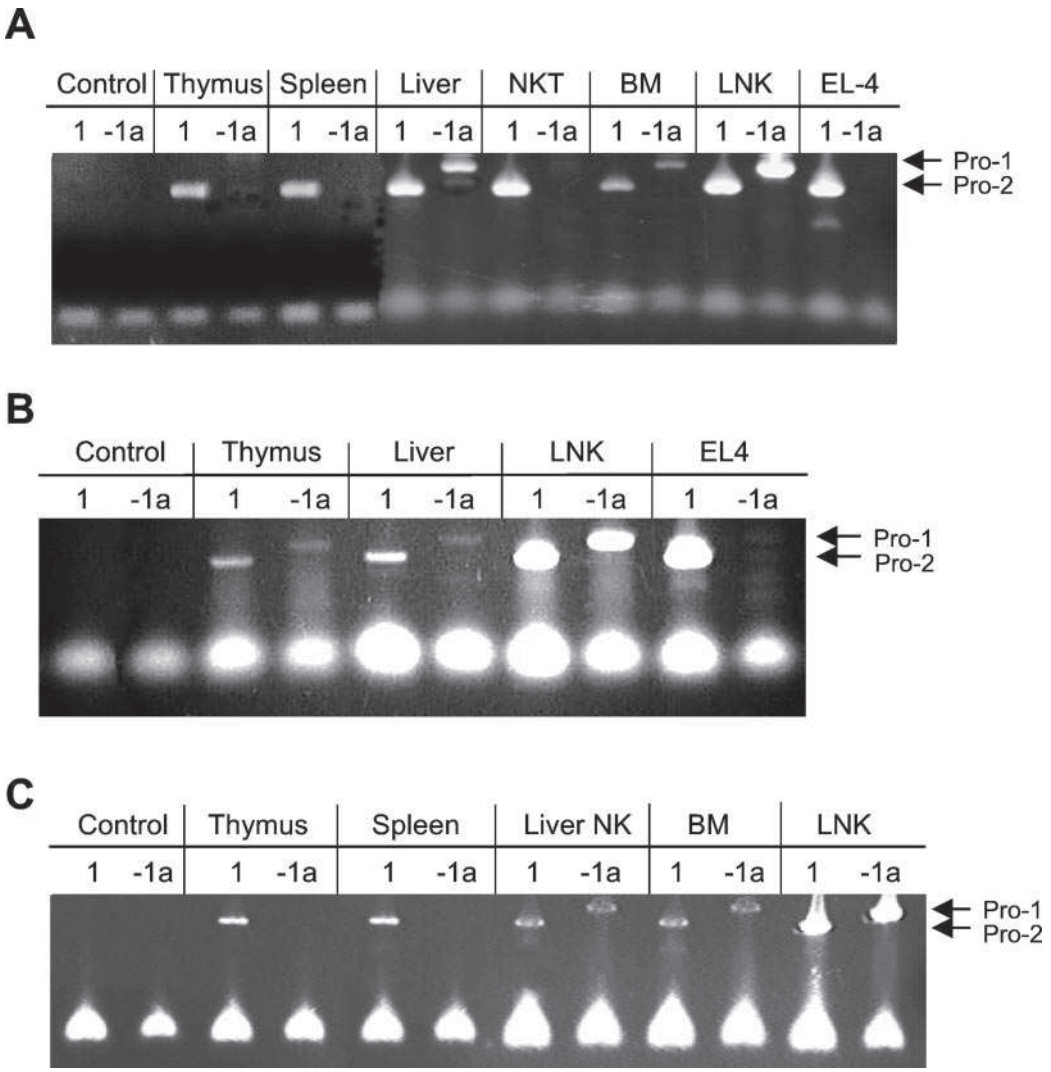

FIGURE 2. Detection of Pro-1 and Pro-2 transcripts by PCR. *A*, Purified mRNA from B6 mouse tissues and cell lines was subjected to RT-PCR with either an exon 1 or exon −1a primer. Arrows indicate the positions of the expected amplification products from either the Pro-1 or Pro-2 transcripts. *B*, Total thymus or liver mRNA from day-15 embryos was subjected to RT-PCR as described in *A*. *C*, Total thymus, liver NK, spleen NK, or bone marrow mRNA from 1-wk-old mice was subjected to RT-PCR as described in *A*.

used to clone the *Ly49a*/*o* Pro-1 transcripts from liver and LNK mRNA. Analysis of the EL-4 and LNK cell lines demonstrated that Pro-1 is used in LNK but not EL-4 cells. EL-4 represents a mature $Ly49^+NKRP1^+CD3^+$ NK-T cell tumor (20), whereas the LNK cell line represents a $Ly49^-CD3^-IL\text{-}2R\beta^+CD16^+CD94^+$ cell that lacks NK activity (17). The *Ly49g* Pro-1 PCR product from LNK (BALB/c derived) was sequenced, and it was identical to the 129/J *Ly49g* alternative transcript. The Pro-2 transcript amplified from LNK was from the *Ly49a* gene. However, the *Ly49a* Pro-1 transcript contains exon 1 and includes the Pro-2 primer site, indicating that Pro-2 may not be active in the LNK cell line because no Pro-2 *Ly49g* transcript was detected, and the *Ly49g* Pro-1 transcript does not contain exon 1. Fig. 2*B* shows the results of RT-PCR on mRNAs derived from day-15 embryonic liver and thymus. Products containing either exon 1 or exon −1a were detected; however, the exon −1a products were less intense. Exon −1a-containing PCR products generated from thymus were cloned into the pCR2.1 vector and sequenced. All of the clones analyzed corresponded to transcripts from the alternative promoter of *Ly49e*. This is in agreement with the previous report of Ly49E expression in fetal thymus (21). In addition, all clones of *Ly49e* contained exon −1a but not exon −1b, as predicted by the absence of a consensus splice acceptor preceding exon −1b in the *Ly49e* gene. To determine whether Pro-1 was active at a time when most Ly49 protein expression is initiated, mRNAs isolated from 1- to 2-wk-old B6 mice were analyzed. Fig. 2*C* demonstrates significant Pro-1 activity in the bone marrow of neonatal mice.

Characterization of Pro-1-derived transcripts

Pro-1 transcripts containing the complete Ly49 coding region were isolated by PCR with the exon −1a primer and a degenerate reverse primer located 3′ of the *Ly49* termination codon. Products from 129/J and B6 liver NK cell mRNA were cloned into the pCR2.1 vector and 10 clones were sequenced from each strain. A summary of the observed cDNAs derived from Pro-1 is shown in Fig. 3. Both strains produced several cDNAs corresponding to *Ly49g* that contained exons −1a and −1b, but skipped exon 1, splicing directly to exon 2. The B6- and 129/J-derived transcripts of *Ly49e* lacked exon −1b and used an exon 1 splice acceptor close to the Pro-2 start site. Pro-1-derived transcripts of *Ly49v* isolated from 129/J NK cells contained both exons −1a and −1b; however, exon 1 was also found in all *Ly49v* clones due to the use of a splice acceptor 175 bp upstream from the previously described Pro-2 initiation site in the *Ly49a* gene (11). Pro-1 transcripts of

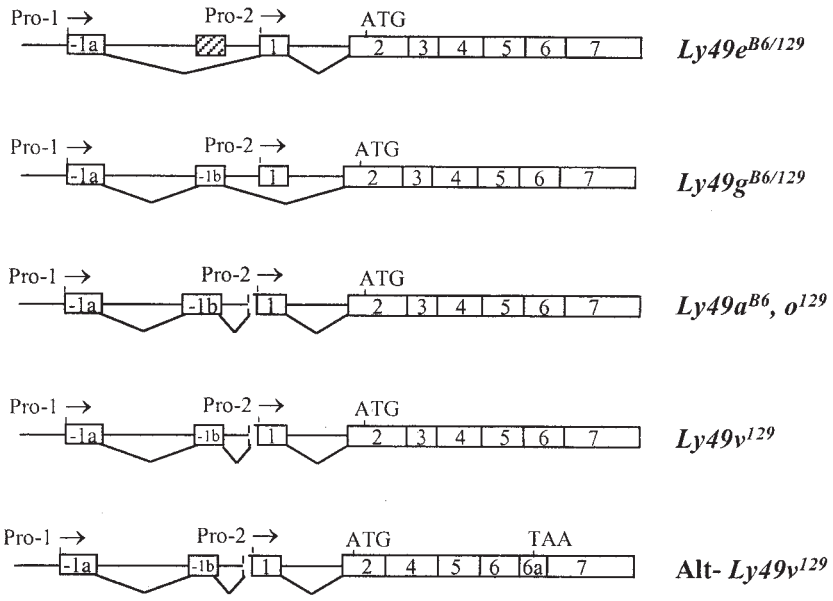

FIGURE 3. Spliced *Ly49* mRNAs generated by Pro-1 transcription. The exons used in the *Ly49a*, *e*, *g*, *o*, and *v* Pro-1 cDNAs detected by PCR are shown. Exons are shown as boxes labeled with their respective exon designation. The pseudo-exon −1b found in the *Ly49e* gene is shown as a hatched box. Lines shown underneath each gene join the exons found in individual cDNAs. The solid arrow indicates the transcription initiation site of each promoter.

Ly49a could not be isolated with the exon −1a primer used for PCR analysis, so an additional primer near the 3′ end of exon −1a was used to successfully amplify the Pro-1 *Ly49a* transcript. Exon −1b of *Ly49a* was 117 bp larger than the *Ly49g* and *v* exon −1b due to the use of an upstream splice acceptor site. The *Ly49a* exon 1 splice acceptor site was the same as that used by the *Ly49v* Pro-1 transcripts, and this provides an explanation for the previous identification of *Ly49a* cDNAs that extend 5′ of the reported Pro-2 start site (22, 23). The *Ly49a*, *o*, and *v* Pro-1 transcripts contain the Pro-2 region, preventing specific detection of Pro-2 transcripts if *Ly49a*, *o*, or *v* Pro-1 transcripts are present. Of the five *Ly49v* clones sequenced, three represented alternatively spliced products lacking a complete open reading frame. Three clones did not contain exon 3, and one of these also contained a novel alternative exon contained in intron 6, as determined by comparing the *Ly49v* sequence with the partial B6 *Ly49v* sequence contained in GenBank (accession no. AC090563). The novel exon (6b) contains an in-frame stop codon, resulting in a truncated Ly49 coding region. The Pro-1 transcripts of *Ly49a*, *e*, *g*, *v*, and alt-*v* have been deposited in GenBank and can be found under accession numbers AY078436, AF419251, AF419249, AF419250, and AF444273, respectively. To determine whether full-length transcripts from the promoter in intron 2 (Pro-3) could be detected, a forward primer from intron 2 was used in conjunction with an exon 7 antisense primer. Fully spliced *Ly49g* and *v* cDNAs derived from Pro-3 were detected in spleen NK cell mRNA (data not shown).

Identification of the transcription start site of the Ly49g *alternative promoter*

Cloning and sequencing of cDNAs produced from Pro-1 indicated that the *Ly49g* gene was actively transcribed from this alternative promoter in LNK cells. To determine the transcription initiation site, an RNA probe encompassing the Pro-1 region of *Ly49g* was synthesized and used to perform an RPA on mRNA purified from LNK cells. Fig. 4 shows the region of the *Ly49g* gene contained within the probe, as well as the protected fragments observed. Protected fragments of 158 and 161 bp were detected in LNK mRNA, but not in mRNA from EL-4 cells which lack detectable alternative promoter activity. This result indicates the presence of two transcription start sites 13 and 16 bp upstream of the 5′ end of the alternative *Ly49g* cDNA isolated from 129/J NK cells. The region protected by the antisense probe and the location of the predicted start sites are indicated in Fig. 1*B*. To rule out the possibility of additional upstream start sites, RT-PCR was performed

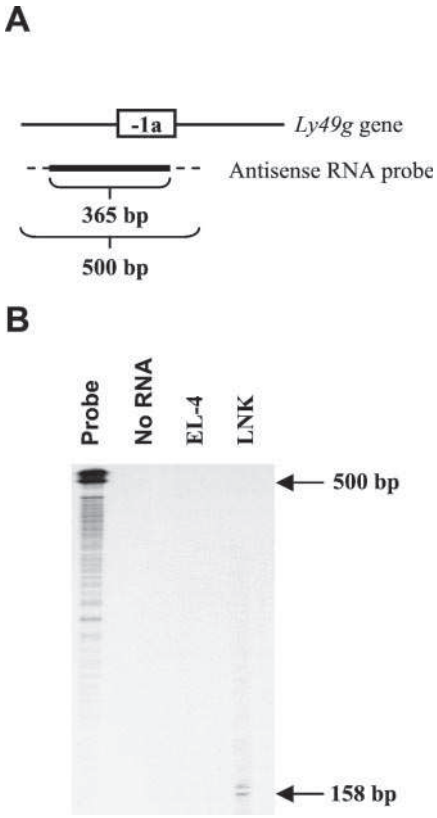

FIGURE 4. Transcription initiation site of *Ly49g* Pro-1. *A*, RNA probe used for RNase protection. The size and location of the *Ly49g* Pro-1 probe used is shown. The thick line indicates the region identical to the *Ly49g* gene, and the dotted line indicates vector-derived sequences. *B*, RPA. One microgram of poly(A)$^+$ mRNA from EL-4 or LNK cells was hybridized to the antisense *Ly49g* Pro-1 RNA probe, digested with RNase, and run on a 6% denaturing polyacrylamide gel. The arrow at 500 bp indicates the size of the undigested probe. The arrow at 158 bp indicates the position of a control 18S RNA probe (Ambion).

with a series of 5′ primers spanning the predicted start site. Primers upstream of the predicted initiation region were unable to amplify Pro-1 cDNA, indicating that there are no additional upstream start sites (data not shown).

Comparison of Pro-1 and Pro-2 activity in the EL-4 and LNK cell lines

To locate and characterize sequence elements required for Pro-1 activity in LNK cells, several luciferase reporter constructs were generated. The regions cloned and the reporter constructs generated are shown in Fig. 5*A*. Examination of the region preceding exon −1a in the *Ly49i* gene revealed a potential TATAA-related element (CATAAAA) 25 nt upstream from the 5′ end of the region with homology to the *Ly49g* cDNA. A (CATAAAA) sequence is also found 29 bp upstream from the Pro-2 initiation site of the *Ly49a* promoter (11). To test this region for promoter activity, a 250-bp fragment containing the predicted Pro-1 element of *Ly49i* was cloned into the pGL3 luciferase reporter vector. Transfection of this construct (Ly49i-Pro1) into EL-4 and LNK cells revealed the presence of a promoter that displayed strong activity in the LNK cell line and no detectable activity in EL-4 cells (Fig. 5*B*), consistent with the PCR results shown in Fig. 2. Removal of the 5′ half of the Ly49i-Pro1 construct yielded a 140-bp fragment that retained the same relative promoter activity, localizing the core promoter region. A *Ly49j* luciferase construct similar to Ly49i-Pro1 was tested and no activity was found. The CATAAAA element is not present in *Ly49j*, and the region encompassing the transcriptional start site has been deleted. Although the *Ly49g* Pro-1 region does not contain a CATAAAA element, a TATAAA element is located 25 nt upstream from the predicted transcription initiation sites. A luciferase construct containing the *Ly49g* alternative promoter region also possessed LNK-specific transcriptional activity. A luciferase construct containing the previously characterized *Ly49a* promoter (Ly49a-Pro2) was used as a control, and it was active in EL-4 but not LNK cells. An additional construct containing 2 kb upstream of the *Ly49a* Pro-2 core promoter (Ly49a-2kb) was tested and shown to contain the same inhibitory activity observed in the *Ly49c*, *i*, and *j* promoters. The lack of Pro-2 activity in LNK cells is contrary to the observed Pro-2 transcripts detected by PCR in Fig. 2; however, the *Ly49a* Pro-1 transcript contains the Pro-2 start site, resulting in a false positive result for Pro-2 because *Ly49a* Pro-1 transcripts are present. Cloning of the Pro-2 PCR band indicated that it corresponded to the *Ly49a* transcript. To determine whether Pro-1 was capable of activating Pro-2 transcription, the core Pro-1 element without the exon −1a splice donor was added to the extended *Ly49a* construct (Ly49a-2kb+Pro1). The exon −1a splice donor was not included in the Ly49a-2kb+Pro1 construct to avoid the production of spliced transcripts that could result in luciferase activity driven by Pro-1. Pro-1 transcripts produced from this construct would contain a long noncoding region that would presumably inhibit luciferase translation. Transfection of this construct into LNK cells did not result in significant luciferase activity, which may be due to the inability of the Pro-2 core to function in LNK cells coupled with inefficient translation of the unspliced Pro-1 transcript. To reliably test for the ability of Pro-1 transcription to overcome the inhibitory effects of the region preceding Pro-2 it will be necessary to identify a cell line that is permissive for both Pro-1 and Pro-2.

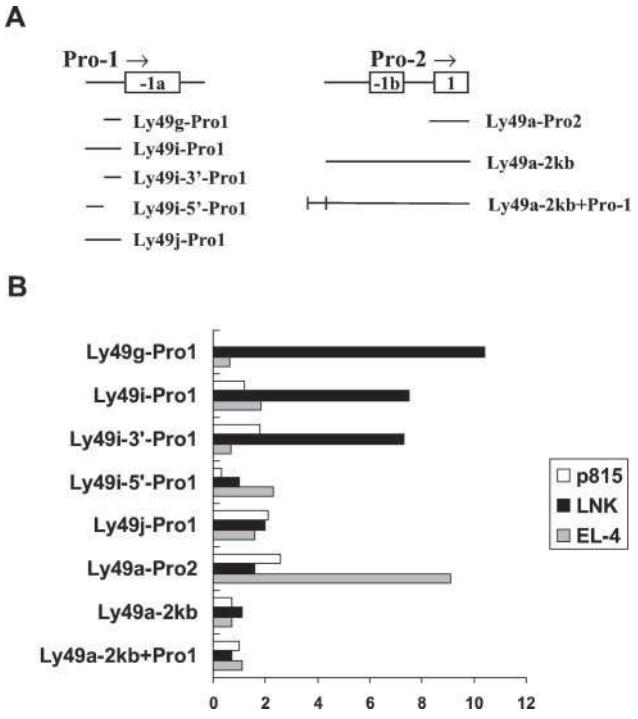

FIGURE 5. Activity of luciferase reporter constructs. *A*, Size and position of Pro-1 and Pro-2 constructs generated. Schematic representations of the Pro-1 and Pro-2 regions are shown with the size and position of the fragments used to generate luciferase reporter vector listed underneath. *B*, Luciferase activity of *Ly49* constructs relative to empty pGL3 vector transfectants. The scale indicates the fold increase of relative light units over cells transfected with empty pGL3 vector. Representative results from a minimum of five independent transfections for each construct are shown.

Discussion

The *Ly49* promoter regions previously described (Pro-2) are highly conserved among all *Ly49* family members, and several have been shown to contain a core promoter preceded by an inhibitory region, suggesting that additional control regions might exist. The current study has identified a novel promoter (Pro-1) that is preferentially expressed in bone marrow and fetal thymus, both sites of initiation of Ly49 expression (24, 25). Therefore, Pro-1 is a good candidate for a regulatory element involved in the process of selective *Ly49* gene activation. Pro-1 activity might be restricted to a small subset within the populations tested because PCR was the only method that efficiently detected Pro-1-derived transcripts, and the identification of the Pro-1 start site in LNK cells required long exposures of products generated by the highly sensitive RPA assay. Also, LNK cells do not express detectable levels of Ly49G protein, even though Pro-1 *Ly49g* transcripts containing a complete open reading frame were detected by PCR. In addition, the scarcity of Pro-1 transcripts was suggested by the identification of only one Pro-1-derived clone of >100 independent *Ly49* cDNAs analyzed from 129/J liver NK cells (16). This suggests that either Pro-1 is a very weak promoter or only a small percentage of cells express Pro-1 transcripts. The latter possibility is supported by the detection of similar Pro-1 and Pro-2 strength in the reporter assays performed in LNK and EL-4 cells. An extensive study of the start sites of *Ly49* genes (19) did not detect any Pro-1 transcripts in adult spleen-derived NK cells, in agreement with our inability to detect Pro-1 activity in these cells. The observed rarity of Pro-1 transcripts together with the detection of *Ly49g* Pro-1 activity in bone marrow and *Ly49e* Pro-1 activity in fetal thymus suggests that Pro-1 is transiently activated in cell populations that are initiating Ly49 expression. It is interesting to note that the Pro-1 region of *Ly49j* does not have significant activity in luciferase reporter assays, and the percentage of NK cells expressing Ly49J (5%) is significantly lower than the percentage of NK cells expressing Ly49I (37%) (10). If Pro-1 is involved in the initial activation of

Ly49 gene expression, the difference in Pro-1 activity between *Ly49i* and *j* may explain the different probability of activation of these two highly related genes. Pro-1 activity would be expected to suppress transcription of Pro-2, due to transcriptional interference (26), because the Pro-1 transcript reads through the Pro-2 region. However, there may not be concurrent activity of Pro-1 and Pro-2 because luciferase assays indicated that Pro-2 was specific for EL-4 cells and Pro-1 was specific for LNK cells. The positive PCR result with the Pro-2 primer in LNK cells was due to the presence of *Ly49a* Pro-1 transcripts in LNK, because we were unable to detect *Ly49g* Pro-2 transcripts and the *Ly49a* Pro-1 transcript encompasses the Pro-2 start site, unlike the *Ly49g* Pro-1 transcript, which skips exon 1. Therefore, the Pro-2 PCR primer set only reliably indicates Pro-2 activity in tissues that do not have any Pro-1 activity. The role of Pro-1 transcripts may be to open the Pro-2 region, making it accessible to transcription factors required for subsequent activation. This would provide an explanation for the presence of an inhibitory region preceding Pro-2. Pro-2 may lack elements required to make the promoter accessible; therefore, it would be dependent on transcription from Pro-1 for its initial activation. The stochastic nature of *Ly49* gene activation may be related to the window of opportunity between the loss of Pro-1-specific transcription factors and the appearance of Pro-2-specific transcription factors. Once activated, Pro-2 appears to function as a constitutive NK cell promoter, because there is no evidence for modulation of Pro-2 activity in mature NK cells. Whether or not Pro-1 transcription is controlled by a probabilistic mechanism is currently under investigation.

The 5′ untranslated region of Pro-1 transcripts is significantly longer than that of Pro-2 transcripts, suggesting possible differences in their translation efficiencies. Transfection of full-length *Ly49g* Pro-1 or Pro-2 cDNAs into 293T cells failed to show any differences in the surface expression of Ly49G (data not shown). One potential effect of the Pro-1-derived 5′ region may relate to the production of alternatively spliced products. Three of five *Ly49v* clones isolated were missing exon 3, and one clone contained a novel alternative exon (6b) that disrupted the open reading frame. Perhaps the presence of the Pro-1 noncoding region enhances the production of alternatively spliced *Ly49* mRNAs. Our laboratory has been unable to identify any Ly49 protein products resulting from cDNAs lacking the third exon (27), suggesting that the primary role of Pro-1 transcription may be to activate the gene, and Ly49 protein expression from this promoter is down-regulated by alternative splicing.

Studies of the *Ly49* promoter used by mature NK cells (Pro-2) have not provided an explanation for the stochastic activation of *Ly49* genes. The discovery of an upstream *Ly49* promoter (Pro-1) that is primarily active in tissues where Ly49 expression is initiated suggests that further study of this novel promoter may lead to an understanding of the mechanisms underlying the seemingly random activation of *Ly49* genes. Generation of *Ly49* transgenic mice with mutations in the Pro-1 region will enable a direct evaluation of the role of Pro-1 in the stochastic process of *Ly49* gene activation.

Acknowledgments

We thank John Wine, Robin Winkler-Pickett, and Bill Bere for their excellent technical assistance in the preparation of mouse tissues and NK cell separation. We also thank Jonathan Summers for assistance with illustrations.

References

1. Brown, M. G., S. Fulmek, K. Matsumoto, R. Cho, P. A. Lyons, E. R. Levy, A. A. Scalzo, and W. M. Yokoyama. 1997. A 2-Mb YAC contig and physical map of the natural killer gene complex on mouse chromosome 6. *Genomics 42:16.*
2. McQueen, K. L., J. D. Freeman, F. Takei, and D. L. Mager. 1998. Localization of five new *Ly49* genes, including three closely related to *Ly49c*. *Immunogenetics 48:174.*
3. Depatie, C., S. H. Lee, A. Stafford, P. Avner, A. Belouchi, P. Gros, and S. M. Vidal. 2000. Sequence-ready BAC contig, physical, and transcriptional map of a 2-Mb region overlapping the mouse chromosome 6 host-resistance locus Cmv1. *Genomics 66:16.*
4. Smith, H. R. C., F. M. Karlhofer, and W. M. Yokoyama. 1994. Ly-49 multigene family expressed by IL-2-activated NK cells. *J. Immunol. 153:1068.*
5. Brennan, J., D. Mager, W. Jefferies, and F. Takei. 1994. Expression of different members of the Ly-49 gene family defines distinct natural killer cell subsets and cell adhesion properties. *J. Exp. Med. 180:2287.*
6. McQueen, K. L., S. Lohwasser, F. Takei, and D. L. Mager. 1999. Expression analysis of new *Ly49* genes: most transcripts of *Ly49j* lack the transmembrane domain. *Immunogenetics 49:685.*
7. Makrigiannis, A. P., J. Etzler, R. Winkler-Pickett, A. Mason, J. R. Ortaldo, and S. K. Anderson. 2000. Identification of the Ly49L protein: evidence for activating counterparts to inhibitory Ly49 proteins. *J. Leukocyte Biol. 68:765.*
8. Silver, E. T., D. Gong, B. Hazes, and K. P. Kane. 2001. Ly-49W, an activating receptor of nonobese diabetic mice with close homology to the inhibitory receptor Ly-49G, recognizes H-2D^k and H-2D^d. *J. Immunol. 166:2333.*
9. Held, W., and B. Kunz. 1998. An allele-specific, stochastic gene expression process controls the expression of multiple Ly49 family genes and generates a diverse, MHC-specific NK cell repertoire. *Eur. J. Immunol. 28:2407.*
10. Kubota, A., S. Kubota, S. Lohwasser, D. L. Mager, and F. Takei. 1999. Diversity of NK cell receptor repertoire in adult and neonatal mice. *J. Immunol. 163:212.*
11. Kubo, S., Y. Itoh, N. Ishikawa, R. Nagasawa, T. Mitarai, and N. Maruyama. 1993. The gene encoding mouse lymphocyte antigen Ly-49: structural analysis and the 5′-flanking sequence. *Gene 136:329.*
12. Kubo, S., R. Nagasawa, H. Nishimura, K. Shigemoto, and N. Maruyama. 1999. ATF-2 binding regulatory element is responsible for Ly49A expression in murine T lymphoid line, EL-4. *Biochim. Biophys. Acta 1444:191.*
13. Held, W., B. Kunz, B. Lowin-Kropf, M. van de Wetering, and H. Clevers. 1999. Clonal acquisition of the Ly49A NK cell receptor is dependent on the *trans*-acting factor TCF-1. *Immunity 11:433.*
14. Gosselin, P., A. P. Makrigiannis, R. Nalewaik, and S. K. Anderson. 2000. Characterization of the Ly49*i* promoter. *Immunogenetics 51:326.*
15. McQueen, K. L., B. T. Wilhelm, F. Takei, and D. L. Mager. 2001. Functional analysis of 5′ and 3′ regions of the closely related Ly49*c* and *j* genes. *Immunogenetics 52:212.*
16. Makrigiannis, A. P., A. T. Pau, A. Saleh, R. Winkler-Pickett, J. R. Ortaldo, and S. K. Anderson. 2001. Class I MHC binding characteristics of the 129/J Ly49 repertoire. *J. Immunol. 166:5034.*
17. Tsutsui, H., K. Nakanishi, K. Matsui, K. Higashino, H. Okamura, Y. Miyazawa, and K. Kaneda. 1996. IFN-γ-inducing factor up-regulates Fas ligand-mediated cytotoxic activity of murine NK cell clones. *J. Immunol. 157:3967.*
18. Watanabe, M., K. L. McCormick, K. Volker, J. R. Ortaldo, J. M. Wigginton, M. J. Brunda, R. H. Wiltrout, and W. E. Fogler. 1997. Regulation of local host-mediated anti-tumor mechanisms by cytokines: direct and indirect effects on leukocyte recruitment and angiogenesis. *Am. J. Pathol. 150:1869.*
19. Wilhelm, B. T., K. L. McQueen, J. D. Freeman, F. Takei, and D. L. Mager. 2001. Comparative analysis of the promoter regions and transcriptional start sites of mouse *Ly49* genes. *Immunogenetics 53:215.*
20. Gays, F., M. Unnikrishnan, S. Shrestha, K. P. Fraser, A. R. Brown, C. M. G. Tristram, Z. M. A. Chrzanowska-Lightowlers, and C. G. Brooks. 2000. The mouse tumor cell lines EL4 and RMA display mosaic expression of NK-related and certain other surface molecules and appear to have a common origin. *J. Immunol. 164:5094.*
21. Van Beneden, K., F. Stevenaert, A. De Creus, V. Debacker, J. De Boever, J. Plum, and G. Leclercq. 2001. Expression of Ly49E and CD94/NKG2 on fetal and adult NK cells. *J. Immunol. 166:4302.*
22. Chan, P. Y., and F. Takei. 1989. Molecular cloning and characterization of a novel murine T cell surface antigen, YE1/48. *J. Immunol. 142:1727.*
23. Silver, E. T., D. E. Gong, C. S. Chang, A. Amrani, P. Santamaria, and K. P. Kane. 2000. Ly-49P activates NK-mediated lysis by recognizing H-2D^d. *J. Immunol. 165:1771.*
24. Williams, N. S., A. Kubota, M. Bennett, V. Kumar, and F. Takei. 2000. Clonal analysis of NK cell development from bone marrow progenitors in vitro: orderly acquisition of receptor gene expression. *Eur. J. Immunol. 30:2074.*
25. Fraser, K. P., F. Gays, J. H. Robinson, K. van Beneden, G. Leclercq, R. Vance, D. Raulet, and C. G. Brooks. 2002. NK cells developing in vitro from fetal mouse progenitors express at least one member of the Ly49 family that is acquired in a time-dependent and stochastic manner independently of CD94 and NKG2. *Eur. J. Immunol. 32:868.*
26. Proudfoot, N. J. 1986. Transcriptional interference and termination between duplicated α-globin gene constructs suggests a novel mechanism for gene regulation. *Nature 322:562.*
27. Makrigiannis, A. P., and S. K. Anderson. 2001. Mechanisms controlling the expression and diversity of the Ly49 gene family. *Mod. Asp. Immunobiol. 1:228.*