Evidence for high bi-allelic expression of activating Ly49 receptors

Arefeh Rouhi
British Columbia Cancer Agency

C. B. Lai
University of British Columbia

Tammy P. Cheng
Washington University School of Medicine in St. Louis

Fumio Takei
British Columbia Cancer Agency

Wayne M. Yokoyama
Washington University School of Medicine in St. Louis

See next page for additional authors

Follow this and additional works at: https://digitalcommons.wustl.edu/open_access_pubs

Part of the Medicine and Health Sciences Commons

Recommended Citation
https://digitalcommons.wustl.edu/open_access_pubs/78

This Open Access Publication is brought to you for free and open access by Digital Commons@Becker. It has been accepted for inclusion in Open Access Publications by an authorized administrator of Digital Commons@Becker. For more information, please contact vanam@wustl.edu.
Authors
Arefeh Rouhi, C. B. Lai, Tammy P. Cheng, Fumio Takei, Wayne M. Yokoyama, and Dixie L. Mager
Evidence for high bi-allelic expression of activating Ly49 receptors

Arefeh Rouhi1, C. Benjamin Lai1,2, Tammy P. Cheng3, Fumio Takei1,4, Wayne M. Yokoyama3 and Dixie L. Mager1,2,*

1The Terry Fox laboratory, British Columbia Cancer Agency, 2Department of Medical Genetics, University of British Columbia, Vancouver, BC, Canada, 3Howard Hughes Medical Institute, Rheumatology Division, Washington University School of Medicine, St Louis, MO 63110, USA and 4Pathology and Laboratory Medicine, University of British Columbia, Vancouver, BC, Canada

Received April 4, 2009; Revised May 28, 2009; Accepted June 27, 2009

ABSTRACT

Stochastic expression is a hallmark of the Ly49 family that encode the main MHC class-I-recognizing receptors of mouse natural killer (NK) cells. This highly polygenic and polymorphic family includes both activating and inhibitory receptor genes and is one of genome’s fastest evolving loci. The inhibitory Ly49 genes are expressed in a stochastic mono-allelic manner, possibly under the control of an upstream bi-directional early promoter and show mono-allelic DNA methylation patterns. To date, no studies have directly addressed the transcriptional regulation of the activating Ly49 receptors. Our study shows differences in DNA methylation pattern between activating and inhibitory genes in C57BL/6 and F1 hybrid mouse strains. We also show a bias towards bi-allelic expression of the activating receptors based on allele-specific single-cell RT–PCR in F1 hybrid NK cells for Ly49d and Ly49H expression in Ly49h+/− mice. Furthermore, we have identified a region of high sequence identity with possible transcriptional regulatory capacity for the activating Ly49 genes. Our results also point to a likely difference between NK and T-cells in their ability to transcribe the activating Ly49 genes. These studies highlight the complex regulation of this rapidly evolving gene family of central importance in mouse NK cell function.

INTRODUCTION

Natural killer (NK) cells constitute an important part of the body’s defence both as centurions of innate immunity and as communicators and collaborators of the adaptive immune system. They exert their function via the interpretation of signals received from their surface receptors. Normal cells display major histocompatibility complex (MHC) class I molecules that are recognized by a number of inhibitory receptors on the surface of NK cells. The inhibitory receptors prevent activation of NK cells and the destruction of normal, MHC class I-possessing cells (1). NK cells also possess stimulatory (activating) receptors that recognize other molecules on the surface of potential target cells which are mostly pathogen-encoded or stress-induced self proteins (2–4).

NK activation and the killing of target cells therefore depend on the balance between stimulatory and inhibitory signals received from these surface receptors. In primates and some other mammals, the killer cell immunoglobulin-like (KIR) family of genes code for the main MHC class-I recognizing receptors (5). However, in rodents, the structurally unrelated Ly49 family provides this function. Both KIR and Ly49 families are polygenic and polymorphic among different individuals and mouse strains (6,7).

The KIR gene cluster of human is ~150 kb containing approximately 14 genes and pseudogenes with very high coding and regulatory sequence similarity (8) indicating that they arose through recent gene duplication events (9). There is great diversity in both gene number and sequence polymorphisms for the KIRs among different people (6). The Ly49 gene cluster has also arisen from recent duplications and gene conversions of ancestral genes (10–12). Hence, there is high sequence similarity both in the coding and non-coding regions among the majority of the genes. The Ly49 cluster includes 16 genes and pseudogenes spanning over 600 kb in the C57BL/6 (B6) mouse strain and is located in the natural killer gene complex (NKC) region of chromosome 6 (8,13). There are two functional Ly49 activating receptors, Ly49D and H, coded by the B6 genome. Ly49D binds to the MHC-class-I allele H2-Dd (14,15) and Ly49H binds to the m157 protein of mouse cytomegalovirus (MCMV) conferring resistance towards virus infection in the strains of mice that express Ly49H (2,3,16,17). Unlike the

*To whom correspondence should be addressed. Tel: +604 675 8139; Fax: +604 877 0712; Email: dmager@bccrc.ca

© 2009 The Author(s)
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.0/uk/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
inhibitory Ly49 receptors that are expressed on T and NKT cells as well as NK cells, the stimulatory Ly49 receptors are only expressed on NK cells (18). Lack of a well-defined upstream promoter-1 (Pro-1) region, which likely acts as a stochastic switch for inhibitory receptor genes in immature NK cells (19,20) and reports of the higher acts as a stochastic switch for inhibitory receptor genes defined upstream promoter-1 (Pro-1) region, which likely subject to distinct regulatory mechanisms.

DNA methylation of the 5′-region of the inhibitory Ly49 genes correlates with their expression. In the case of Ly49a and Ly49c, where stochastic mono-allelic expression has been demonstrated, we have previously shown that DNA methylation and mono-allelic receptor expression correlate (22). The existence of stochastic mono-allelic expression and/or correlation with DNA methylation of the 5′-region of activating receptors and the maintenance of their expression. Furthermore, we show evidence for a mode of transcriptional regulation for the activating Ly49 genes that differs from that of the inhibitory genes.

**MATERIALS AND METHODS**

**Mice**

All C57BL/6 mice were bred and maintained in the animal facility of the Columbia Cancer Research Centre (Vancouver, British Columbia, Canada). 129SvEvTac and 129SvEvTac/C57BL6 F1 hybrids were ordered from Taconic farms. NOD/ShiLtJ and C57BL6/NOD-ShiLtJ F1 hybrids were ordered from the Jackson Laboratory (Bar Harbor, Maine, USA). All mice used in this study were more than 6-weeks old. The B6.BXD8/B6 (23) and age-matched B6 mice were 6.5-weeks old. All experiments were according to a protocol approved by the Committee on Animal Care of the University of British Columbia.

**Antibodies, cell separation and flow cytometry**

The monoclonal antibody (mAb) anti-FcRγ (2.4G2) (24,25) and 3D10-FITC (21) have been described before. Anti-CD3ε-PerCP-Cy5.5, anti-NK1.1-PE, anti-NK1.1-APC, anti-DX5-PE (alpha-2 integrin, CD49b), 1F8-FITC (anti-Ly49C/I/H), 5E6-biotin (anti-Ly49C/I), 4E5-FITC (anti-Ly49D), A1-biotin (anti-Ly49A/B) and fluorochrome-conjugated streptavidin were purchased from BD Biosciences (Mississauga, Ontario, Canada). The anti-mouse NKp46-PE antibody was purchased from eBioscience (San Diego, CA, USA). The 12A8 purified antibody was generously provided by Dr Stephen Anderson (NCI, Frederick, MD, USA) and was conjugated to FITC via Thermo Scientific Pierce EZ-label fluorescein isothiocyanate protein labelling kit (Rockford, IL, USA) per manufacturer’s protocol. Flow cytometry for cell sorting was performed on Cytopeia Influx cell sorter and Becton Dickinson FACSaria Cell Sorting System (Mississauga, Ontario, Canada). All sorted samples were >95% pure or resorted for high purity. For the single-cell sorting, NK cells were sorted into wells containing JEG-3 (human choriocarcinoma) as carrier cells. Doublet discrimination was applied to avoid more than one cell per well.

**Primary cell and tissue genomic DNA extraction**

Genomic DNA (gDNA) was obtained from FACS sorted cells as described before (22). gDNA was extracted from fresh B6 mouse liver using DNAzol reagent (Invitrogen) per manufacturer’s instructions. Further proteinase K digestion and phenol–chloroform extraction was performed.

**Sodium bisulfite conversion and PCR**

Bisulfite conversion was performed as described previously (22). The conversion rate was >98%. First-round PCR amplification of Ly49h 5′ region was performed using Ly49h forward and reverse flanking bisulfite primers. Converted DNA was used as template in a 45 µL reaction volume, containing 30 pmol of each primer, 1 mM dNTPs, 3 mM MgCl2 and 0.5 U Taq Platinum DNA polymerase (Qiagen). After initial denaturation for 7 min at 95°C, 30–40 cycles were performed, each consisting of 90 s at 95°C, 55 s at 50°C and 40 s at 72°C with a final extension of 7 min at 72°C. Two microliters of the first PCR was used for nested amplification using Ly49h forward and reverse nested bisulfite primers. The same amplification conditions were chosen as for first-round PCR with the exception that the annealing temperature was changed to 48 + 0.1°C/cycle. All bisulfite primer sequences are shown in Table 1.

For the Ly49d 5′ region of the B6 strain the following primers were used to amplify two regions for COBRA analysis of two CpG dinucleotides: first round (flanking) PCR was performed with Ly49d forward and reverse flanking bisulfite primers. After initial denaturation for 7 min at 95°C, 35 cycles were performed, each consisting of 90 s at 95°C, 55 s at 51°C and 50 s at 72°C with a final extension of 7 min at 72°C. Two separate nested PCRs were performed on 3 µl of the product of the flanking PCR to amplify two regions containing an upstream and a downstream (in relation to transcription start site) CpG dinucleotide. For the upstream CpG-containing fragment the Ly49d forward flanking and reverse nested bisulfite primers were used. After initial denaturation for 7 min at 95°C, 35 cycles were performed, each consisting of 90 s at 95°C, 55 s at 51°C and 17 s at 72°C with a final extension of 7 min at 72°C. The downstream CpG-containing fragment was amplified via the Ly49d forward nested and reverse flanking bisulfite primers. After initial denaturation for 7 min at 95°C, 35 cycles were performed, each consisting of 90 s at 95°C, 55 s at 53°C and 40 s at 72°C with a final extension of 7 min at 72°C. The resulting products were subjected to combined bisulfite and restriction enzyme analysis (COBRA).

For Ly49d and Ly49r 5′ region amplification from the 129SvEvTac/C57BL6 F1 hybrid cells, first round PCR was performed with the same flanking primers as Ly49d as their sequence is identical to Ly49r with the same PCR conditions as that used for Ly49d of B6. Nested PCR was
performed on 2 µl of the flanking PCR product with the forward nested Ly49d primer (which is identical to Ly49r sequence) and Ly49d/r reverse nested bisulfite primer. After initial denaturation for 7 min at 95°C, 35 cycles were performed, each consisting of 90 s at 95°C, 55 s at 50 + 0.1°C/cycle and 45 s at 72°C with a final extension of 7 min at 72°C. The PCR products were electrophoresed on 1% agarose gels and correct size bands were extracted using the MinElute gel extraction kit (Qiagen). The purified products were cloned into the T-vector using the pGEM T-vector kit (Promega). Sequencing was performed using the SP6 primer by McGill University and Genome Québec Innovation Centre sequencing facility. All clones included in the figures are unique as per criteria stated before (22).

Combined bisulfite and restriction enzyme analysis

For COBRA, gel purified fragments of Ly49d upstream and downstream regions were digested with TaqI and BmgBl restriction enzyme (NEB), respectively to distinguish between methylated (CpG) and unmethylated (TpG). Only fragments that were originally methylated in the gDNA and therefore not converted by sodium bisulfite treatment are cut. Gel purified Ly49h fragments were digested with TaqI (NEB) and BsaAI (NEB) restriction enzymes.

RNA extraction, RT–PCR and 5’ amplification of cDNA ends

cDNA was generated from total B6 and NOD/ShiLtJ spleens and FACS sorted B6 T-cell RNA per SuperScript III protocol. PCR was performed on the generated cDNAs using primers specific for Ly49d, g, h, Nkg2d and actin. Ly49g and actin primers have been described before (22). All RT–PCR primer sequences are shown in Table 1.

5’-Rapid amplification of cDNA end (5’-RACE) was performed on B6 spleen RNA using the FirstChoice RML-RACE kit (Ambion) per manufacturer’s protocol with Ly49h-specific primers. For the outer PCR, the Ly49h Exon 3–4 reverse primer used for RT–PCR was used as the outer 5’-RACE primer in combination with the kit’s outer forward primer. For the inner PCR, Ly49h Exon 2 inner reverse primer in combination with the kit’s inner forward primer. The products were cloned into the T-vector using the pGEM T-vector kit (Promega). Sequencing was performed using the T7 primer by McGill University and Genome Québec Innovation Centre sequencing facility.

Electrophoretic mobility shift assay

Nuclear extract from FACS sorted B6 T-cells and IL-2-expanded FACS-sorted B6 NK cells were prepared as described before (26). Electrophoretic mobility shift assay (EMSA) was performed with the following double-stranded probes located downstream of the transcriptional start site within exon1 of various Ly49 genes (only the forward strand is shown in Table 1): to control for length differences among the probes, poly-T tails were added to the shorter Ly49a and c probes. The YY1 antibody (sc-281) was purchased from Santa Cruz Biotechnology and supershift experiments were performed as described before (26).

Single-cell RT–PCR

Single-cell RT–PCR and Southern blotting was performed as described before (27) with minor modifications. RNA from JEG-3, a human choriocarcinoma cell line (28), was used as carrier. Ly49d forward exon 2 and reverse exon 4 primers were used to specifically amplify Ly49d cDNA of both B6 and NOD/ShiLtJ. Ly49dNOD-specific oligo probe and Ly49dNOD-specific oligo probe were used to
specifically detect cDNA from Ly49d<sup>B6</sup> or Ly49d<sup>NOD</sup> in Southern blot analysis. For Nkg2d, the same primers used for RT–PCR (previous section) were used and the products were identified after Southern blotting with an Nkg2d-specific oligo probe. Up to five sets of independent PCRs were performed on each single cell.

RESULTS

Lack of detectable activating Ly49 transcripts in T-cells

Unlike the inhibitory Ly49 receptors, the activating Ly49 receptors of the B6 strain (Ly49D and H) are not expressed on the surface of T and NKT cells (29,30). This absence of surface expression on T and NKT is most likely due to the negligible amounts of the activating adaptor protein DAP12 that seems to be required for cell surface expression of the B6 activating Ly49 receptors (31,32). However, there are reports of the association of Ly49D with the activating adaptor protein CD3ξ and DAP10 (33,34). We also did not detect Ly49R (an activating receptor in 129 mice) or Ly49D and R surface expression on 129SvEvTac or 129SvS6/B6 F1 hybrid fresh ex vivo splenic T-cells (DX<sup>5</sup>, CD3ε<sup>+</sup>) respectively by flow cytometry (data not shown). To determine if transcripts for the activating receptors exist in T-cells, fresh ex vivo splenic T-cells were sorted with high purity and RT–PCR with Ly49d and Ly49h-specific primers was performed on the isolated RNA (Figure 1A). Transcripts from both Ly49d and Ly49h were detected in whole spleen but very little to none in splenic T-cells, which correlates with the lack of surface expression of these receptors on T-cells. We detected similar amounts of the inhibitory Ly49g and Nkg2d cDNA in both splenic lymphocytes and sorted T-cells.

5′-region DNA methylation of Ly49d and Ly49h correlates with state of expression

Ly49 genes have multiple promoters but most are transcribed from a region called Pro-2, which is thought to be the main promoter in mature NK cells (19,35). The transcriptional start site of Ly49d has been mapped to the Pro-2 region (35). To determine the promoter of origin for Ly49h, 5′RACE was performed on whole spleen RNA with gene-specific primers. The Ly49h transcripts detected originated from Pro-2. However, unlike other Ly49 genes (35) we did not detect transcriptional start site variability for Ly49h (Figure 1B).

Based on the 5′RACE results, the region equivalent to the Pro-2 of the inhibitory Ly49 genes is the main area of transcriptional start for Ly49h. We examined the DNA methylation status of the CpG dinucleotides of this region that we shall refer to as the 5′-region of Ly49h from here onwards. There is a cluster of six CpGs within the proximal 5′-region of Ly49h with three CpGs located upstream and three located downstream of the transcription start site (Figure 2A).

Sodium bisulfite sequencing revealed heavy methylation of this region in FACSc-sorted Ly49H-negative (1F8<sup>−</sup>) splenic NK and T-cells but not in Ly49H-positive (1F8<sup>+</sup>, 5E6<sup>−</sup>) NK cells (Figure 2B). As with the inhibitory Ly49a gene (22), the CpG sites downstream of the transcriptional start site are most heavily methylated in Ly49H-negative cells. However, unlike for Ly49a and Ly49c, which display a “half-and-half” mono-allelic methylation pattern, the Pro2 region was hypo-methylated in all clones sequenced in the Ly49H<sup>+</sup> population. The bisulfite sequencing results were also confirmed by COBRA (Figure 2C).

There are very few CpG dinucleotides in the 5′-region of Ly49d. Hence, we only used COBRA to analyze two CpG dinucleotides, one upstream and the other downstream of the transcription start site (Figure 2D). Taq<sup>II</sup> and BmgBI restriction endonucleases were used to assay the upstream and downstream CpGs, respectively. Ly49D expressing and non-expressing NK cells of fresh ex vivo spleen were FACS sorted and analyzed for DNA methylation at these two CpGs. The results are the combination of four individual PCRs per each sorted population. Ly49D negative NK cells show moderate amounts of DNA methylation at both assayed CpGs where as the Ly49D positive NK cells show almost no DNA methylation at either site (Figure 2E) reflecting a DNA methylation pattern similar to that observed for the Ly49h 5′-region in Ly49H positive NK cells (Figure 2B and C).

5′-region DNA methylation of Ly49d and Ly49r does not correlate with stochastic mono-allelic receptor expression in 129SvS6/B6 F1 hybrid

Mono-allelic gene expression has not been shown for the activating Ly49 genes and we did not observe the bimodal
DNA methylation pattern typical of the mono-allelically expressed inhibitory Ly49a and c (22) for Ly49d and h in receptor-expressing NK cells. In order to verify the DNA methylation status of both alleles of an activating receptor in receptor-expressing NK cells, we investigated the DNA methylation status of Ly49d in the 129S6/B6 F1 hybrid. Antibody binding and specificities have been examined for the 129S6 strain Ly49 receptors allowing for sorting and analysis of specific Ly49-expressing subpopulations (36). Ly49d<sup>B6</sup> and Ly49<sup>129</sup> are considered alleles based on the criteria presented by Makrigiannis et al. (36,37) such as coding region homology, intron homology and gene order. We therefore chose to assay DNA methylation of the 5'-regions of Ly49d and r in the F1 hybrid.
There are no antibodies available that bind Ly49D but not Ly49R or vice-versa, making receptor/allele-specific sorting impossible because the 4E5 and 12A8 antibodies (14) detect both receptors. However, the 4E5 antibody also detects Ly49O129 and Ly49V129 with low binding affinity (36).

Assuming Ly49d and r are allelic, we hypothesized that, based on the lack of a bimodal DNA methylation pattern observed for Ly49D and H expressing NK cells in B6, both Ly49d and r 5’-regions should be hypo-methylated. We FACS sorted 12A8-positive/A1-negative NK cells (Figure 3A). This population should include Ly49D-single positive, Ly49R single-positive and Ly49D and R double positive cells. A few polymorphisms in the 5’-region of Ly49d and r, assayed by sodium bisulfite sequencing, allow for their distinction (Figure 3B).

We also sorted 12A8 negative (Ly49D/R non-expressing) NK cells and analyzed the 5’-region of Ly49d and r for this population as well (Figure 3B). As expected, the 5’-region of Ly49d and r in the Ly49D/R-negative NK cells was hyper-methylated for all the sequenced clones. With the exception of two clones for Ly49d and one clone for Ly49r, all other clones (16 clones) are hypo-methylated in Ly49D/R-positive NK cells. The most 5’ CpG of all of the nine independent Ly49R-positive clones is methylated in Ly49D/R-positive NK cells. However, the rest of the methylation pattern does not resemble that of the Ly49D/R-positive cells. We also sequenced the Ly49R Pro2 region from Ly49R-positive NK cells of the 129/SvEvTac mouse (the parent of the F1 hybrid) to gauge the methylation of the most 5’ CpG. We observed the same methylation pattern as in the F1 (Supplementary Figure 1) possibly indicating that this pattern of methylation is specific to the Ly49r locus. The bimodal DNA methylation observed for the inhibitory Ly49a and c is not observed for Ly49d and r, rather the pattern of DNA methylation resembles that observed for Ly49h.

Detection of bi-allelic expression of Ly49d

The recent sequencing and assembly of the NOD Ly49 cluster (7) provided an opportunity to test allelic expression of activating Ly49 receptors. While there may be some question as to the allelic relationship of Ly49d and Ly49r129, the NOD strain has a definite Ly49d allele with ~98.5% identity to the B6 Ly49d sequence at the level of cDNA (7,38). In order to directly test the quality and quantity of allelic expression of Ly49d, we single-cell FACS sorted NK cells (NKp46+, DX5+, NK1.1+, CD3ε-) from B6/NOD F1 hybrid spleen and performed single-cell RT–PCR (27) followed by Southern blot and hybridization with allele-specific probes (see Figure 4 for example).

We assayed 88 NK cells and detected cDNA for Ly49d129 and/or Ly49r129 in 40 cells where cDNA from both Ly49d129 and Ly49r129 was detected in 22 of these 40 cells indicating 55% bi-allelic expression (Table 2). We never detected any products from wells that only
Table 2. Single-cell RT–PCR results

<table>
<thead>
<tr>
<th>cDNA type</th>
<th>Ly49D&lt;sup&gt;B6&lt;/sup&gt;</th>
<th>Ly49D&lt;sup&gt;NOD&lt;/sup&gt;</th>
<th>Ly49D&lt;sup&gt;B6&lt;/sup&gt; and Ly49D&lt;sup&gt;NOD&lt;/sup&gt;</th>
<th>Total Ly49D&lt;sup&gt;+&lt;/sup&gt;</th>
<th>Ly49D&lt;sup&gt;+&lt;/sup&gt; NKG2D&lt;sup&gt;+&lt;/sup&gt;</th>
<th>Ly49D&lt;sup&gt;+&lt;/sup&gt; NKG2D&lt;sup&gt;-&lt;/sup&gt;</th>
<th>Total NKG2D&lt;sup&gt;+&lt;/sup&gt;</th>
<th>Ly49D&lt;sup&gt;+&lt;/sup&gt; NKG2D&lt;sup&gt;-&lt;/sup&gt;</th>
<th>Total cells with cDNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of transcribing single cells (n)</td>
<td>5</td>
<td>13</td>
<td>22</td>
<td>40</td>
<td>40</td>
<td>33</td>
<td>73</td>
<td>7</td>
<td>80</td>
</tr>
<tr>
<td>Percentage positive</td>
<td>5/40 = 12.5%</td>
<td>13/40 = 32.5%</td>
<td>22/40 = 55%</td>
<td>40/73 = 55%</td>
<td>33/73 = 45%</td>
<td>7/80 = 8.8%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

contained carrier cells. The single-cell RT–PCR technique will underestimate bi-allelic expression due to the fact that one allele may not be detected through inefficient cDNA generation during the RT reaction, multiple sample purification steps and random sampling error (27,39).

As a control for the efficiency of our method, we used Nkg2d, an activating receptor gene that is transcribed in all NK cells (40) most probably at levels similar to the Ly49 genes, to gauge the efficiency of cell sorting, cDNA generation and detection. We only detected 73 Nkg2d<sup>+</sup> cells from the total cells assayed after multiple sets of independent PCRs (Table 2). Nkg2d transcripts went undetected in seven NK cells indicating that possibly 9% of the time the transcript of a gene and/or allele is not detected by our method. In total, we detected transcript for at least one of Ly49d<sup>B6</sup>, Ly49d<sup>NOD</sup> and Nkg2d in 80 cells from the total of 88 NK cells assayed.

Interestingly, the difference between the fraction of cells positive for Ly49d<sup>NOD</sup> versus Ly49d<sup>B6</sup> is nearly 3-fold. The allelic expression of Ly49G in various F1 hybrid mice using allele-specific antibodies shows that different alleles of the same gene can be expressed at different levels in various mouse strains (41). This might be due to the MHC class-I background of the mice, the Ly49 cluster gene composition as well as polymorphisms in the promoter regions.

Expression of Ly49H in a single genomic allele mouse model

Ly49H is expressed on ~50% of splenic NK cells in the B6 strain. For a completely bi-allelically transcribed gene expressed in 50% of NK cells, each allele would be expressed in all expressing cells and would contribute half of the total transcript in each cell. In a situation with only one genomic allele present, the fraction of cells expressing the gene should, however, remain the same at 50% but the amount of total transcript should be halved. For a gene expressed mono-allelically through allelic exclusion, one allele is chosen for expression and the other allele is repressed. If this gene is expressed in 50% of cells, each allele is transcribed in 25% of cells only. In contrast, for a stochastic mono-allelic expression pattern, as described for the inhibitory Ly49 receptors, every allele is equally, randomly and independently expressed (42). Based on this model, if 50% of NK cells are positive for a receptor, each allele should be expressed in 29.2% of NK cells with 8.4% of all NK cells expressing this receptor bi-allelically (see model of this situation in Figure 5A).

We attempted to distinguish between the bi-allelic versus stochastic mono-allelic model using the B6.BXD8 (B6.Ly49h<sup>−/−</sup>) mouse, generated recently by two independent groups (23,43), which carries the B6 Ly49 cluster with a ~26 kb deletion spanning the length of the Ly49h gene. Both groups have independently shown that the expression of other Ly49 receptors does not change in the B6.BXD8 mouse compared to the wild-type B6 mouse indicating that the lack of Ly49H does not affect the expression of other Ly49 receptors (23,43). Unfortunately, an inhibitory Ly49 gene deletion mouse model does not exist for comparison. The B6.BXD8 is unique because the deleted Ly49 genomic allele is in the same Ly49 gene cluster as the other (non-deleted) chromosome. The variable Ly49 clusters as well as possible differences in MHC class-I background can produce confounding effects on expression of Ly49 receptors in F1 hybrids (41).

To investigate the expression pattern of Ly49H on NK cells carrying only one genomic allele of Ly49h on the B6 background, we FACS analyzed peripheral blood NK cells from 6-week old Ly49h<sup>+/−</sup> heterozygous (B6.BXD8/B6) mice by staining with the 3D10 antibody. The percentage of NK cells expressing Ly49H in the heterozygous mice (Ly49h<sup>+/−</sup>) was reduced to 35% from that of age-matched B6 (Ly49h<sup>++</sup>) that have 50% Ly49h<sup>+</sup> NK cells (P = 0.0005) (Figure 5B). This fraction of Ly49H expressing cells in the heterozygotes is lower than expected for a completely bi-allelically expressed gene but higher than the expected 29.2% based on the stochastic mono-allelic model (Figure 5A). Thus, based on these data, the situation for Ly49h<sup>+</sup> likely lies between these two extremes. Although we cannot predict the possible effects of the deletion of one Ly49h allele on the expression of the other allele, we can assume that in the wild-type B6 background both alleles should have the same probability of expression. Since Ly49H is expressed on 50% of NK cells and assuming that in the Ly49h<sup>+/−</sup> B6 mouse the second allele would also be expressed on 35% of NK cells, the bi-allelic expression of Ly49H would be ~20% (Figure 5C). This value is strikingly similar to the percentage of Ly49d<sup>B6</sup>/NOD (bi-allelic) NK cells (assuming 50% Ly49D<sup>+</sup> NK) of the B6/NOD F1 hybrid, as assayed with single-cell RT–PCR (Figure 5D).

Low sequence identity in the 5′-UTR of inhibitory and activating Ly49 genes

The regulatory elements governing the transcription of activating Ly49 genes have not been analyzed to date. It is unknown if inhibitory and activating genes share common transcription factors. Multiple alignment of a number of inhibitory and activating Ly49 genes from various sequenced mouse strains (performed with Clustal W)
shows a region in exon 1 that is present for the activating Ly49 genes but is either absent or different for the inhibitory ones (Figure 6A). This region contains a number of potential binding sites for the transcription factor Yin Yang-1 (YY1) in the activating but not the inhibitory genes.

Since more than half of the assayed CpG dinucleotides of the 5'-region of Ly49 genes fall downstream of the transcriptional start site, it is possible that differential methylation of this region might affect transcription of the Ly49 genes. Furthermore, investigation of the possible differential regulatory role of this region might explain some of the differences between inhibitory and activating receptor expression. We therefore probed the possible existence of activating gene specific transcription factor binding sites by gel-shift assay with nuclear extracts from FACS sorted IL-2 cultured NK cells and T-cells (Figure 6B and Supplementary Figure 2).

Fragments spanning the region of interest from Ly49a, c, d and h were used as probe. In both NK and T-cells a differential pattern of protein binding was observed between the activating and inhibitory Ly49 probes. A large band (indicated by a bracket in Figure 6B and Supplementary Figure 2), likely consisting of multiple protein complexes, was most prominent for the activating genes. In addition, a protein complex only bound to the activating Ly49 probes (indicated by arrow in Figure 6B and Supplementary Figure 2) disappeared (supershifted) in the presence of YY1 antibody.

DISCUSSION

Here, we have analyzed the DNA methylation of the 5'-region of B6 activating NK receptor genes, Ly49d and h. We have shown that DNA methylation of this region correlates with expression patterns of these receptors as with the inhibitory Ly49A, Ly49C (22) and NKG2A (44). Furthermore, as in the case of Ly49a, the CpG dinucleotides downstream of the transcriptional start site of Ly49h seem to have higher levels of methylation compared...
Figure 6. Identification of a region of high homology among activating Ly49 genes. (A) Alignment of 5' regions of inhibitory and activating Ly49 genes. A region, located in exon 1, with high homology among activating receptors only, is boxed (dashed line). The region of transcriptional start site is indicated by the solid black line above the alignment. Promoter region sequences of Ly49d and h of NOD, Ly49p and r of 129/S6 as well as Ly49a, c, d, e, f, g and h of the B6 strain are shown. Ly49a, c, e, f and g are inhibitory genes and all other sequences are from activating genes. CpG dinucleotides are also highlighted. The sequence used as probe in the supershift experiment is indicated for Ly49d with box (solid line). Potential YY1 binding sites (CCAT and ACAT) are shown in bold font. (B) Double-stranded probes harbouring the region of interest in exon 1 for the activating Ly49d and h of the B6 strain and the corresponding region for the inhibitory Ly49a and c were used in gel-shift experiments with nuclear extract from FACS-sorted T-cells and IL-2 cultured NK cells. A YY1 antibody was used for supershift assays. The bracket indicates the protein complexes bound most significantly by the activating Ly49d and h probes but not the inhibitory Ly49a and c in both NK and T-cells. The arrow indicates another protein complex specific to the activating Ly49 probes that disappears (supershifts) in the presence of anti-YY1 in both NK and T-cells.
to those located upstream in Ly49H-negative cells. This find again raises the possibility of the existence of a proximal downstream regulatory element for the Ly49 genes whose function might be affected by DNA methylation. Indeed for the KIR genes, a region spanning the transcriptional start site and the 5′-UTR provides the core promoter activity in T-cells but not NK cells (45). This region also binds different protein complexes in NK and T-cells as shown with gel shift experiments (45). The multiple alignment of the 5′-region of inhibitory and activating Ly49 genes revealed a region of high homology in exon 1 among the activating genes that is not present in the inhibitory Ly49 sequences (Figure 6A). Further investigation of this region, for the activating Ly49d and h versus the corresponding region for the inhibitory Ly49a and c, in gel shift experiments, showed different patterns of protein complex binding between inhibitory and activating gene regions (Figure 6B). YY1 was also confirmed as a candidate transcription factor binding to this region. YY1 is known to have both transcriptional activating and inhibiting properties (46). In addition, the tested region showed a different protein-binding pattern with nuclear extract from NK and T-cells (Figure 6B). These results combined with the lack of detectable activating Ly49 transcripts in T-cells support a varied mode of transcriptional regulation for the two types of Ly49 genes. Further analysis of this region is needed to confirm the extent of its role in the transcriptional regulation of the activating Ly49 genes.

In contrast to the inhibitory receptors, the DNA methylation pattern of the 5′-region in the Ly49D and H expressing NK populations does not follow the bimodal (half-and-half) methylation pattern. The half-and-half pattern of DNA methylation correlates with the monoallelic expression of Ly49A, C and NKG2A as was subsequently shown using F1 hybrids for Ly49a and Nkg2a (22,44). If DNA methylation does not correlate with stochastic mono-allelic expression; either stochastic gene expression is maintained at the level of histone modifications only or this pattern is possibly an indication of high, if not complete, bi-allelic expression and even the lack of stochastic expression of the activating receptors.

We have previously shown that the transcription of some Ly49 genes and Nkg2a correlates with histone acetylation levels (22,44,47). Ly49g transcription in the EL4 cell line is activated mostly in response to histone deacetylase inhibitors but is mostly unaffected by DNA methyltransferase inhibitors. Histone acetylation levels of the Pro-2 region as assayed via chromatin immunoprecipitation (ChIP) also correlate with state of expression in EL4-derived subclones (47). It is possible that the stochastic mono-allelic expression of Ly49G (42) is controlled at the level of histones only. Based on these results, it is also possible that the maintenance of the activating receptor expression patterns is through differential histone modifications.

However, an alternative hypothesis to explain the lack of a half-and-half DNA methylation pattern could be that the expression of the activating receptors is predominantly bi-allelic. Based on this hypothesis, in the Ly49D/R expressing NK cells of the 129/B6 F1 hybrid, the 5′-regions of both Ly49d and r should always be hypomethylated to reflect absolute bi-allelic expression. Our results show that most Ly49d and r clones are hypo-methylated (16/19) in the receptor-expressing NK cells indicating a strong deviation away from the monoallelic DNA methylation pattern observed for the inhibitory genes (Figure 3B). If Ly49d and r are not alleles but are independent loci, based on the product rule, assuming no bias for co-expression given 50% surface expression of each receptor by NK cells, we would expect 25% of all NK cells to be co-expressing the two receptors. Since the total of NK cells expressing one or both of Ly49D and R would be 75% (assuming independent loci), we would expect that 1/3 of all clones sequenced for each gene (when sorting for D and/or R expressing NK cells) to show hyper-methylation patterns similar to the negative population. Our data indicates a much lower methylation frequency than would be expected with stochastic mono-allelic expression and hence we can conclude that neither Ly49d nor r show bimodal DNA methylation patterns. Our experiment does not prove or disprove the allelic relationship of Ly49d and r.

The bias towards bi-allelic expression of the activating Ly49 receptors is further supported by the results of the single-cell Ly49d allele-specific RT–PCR performed on NK cells of the B6/NOD F1 hybrid (Table 2) as well as the higher than expected expression of Ly49H from a single genomic allele (Figure 5). We detected bi-allelic transcription of Ly49d in 55% (22/40) of Ly49d-transcribing NK cells. If Ly49D is expressed on 50% of NK cells, the percentage of NK cells expressing this receptor bi-allelically would be ~28% (Figure 5D). The single-cell RT–PCR method tends to be somewhat inefficient in detecting bi-allelic gene expression. The transcription factor, Pax-5, was statistically deemed to be bi-allelically expressed by this method even though only close to 65% of the cells analysed showed bi-allelic expression (39). Transcripts for our control gene Nkg2d, that is expressed by all NK cells (40), went undetected in ~9% (7/80) of NK cells (Table 2), indicating a likely underestimation of the bi-allelic expression percentage calculated for Ly49d. However, based on the few hyper-methylated clones detected from the Ly49D/R-expressing NK cells as well as the detection of only one Ly49d allele in a few single cells (from the B6/NOD F1 hybrid) even after five sets of single-cell RT–PCR, we believe that a real albeit very small population of NK cells express Ly49D mono-allelically.

It is not inconceivable that the activating receptors are controlled and expressed differently from the inhibitory receptors. The lack of a well-defined Pro-1 element and its transcripts (19,20) in addition to deviation of co-expression percentages from the product rule (21) suggests a different mode of transcriptional regulation from that of the inhibitory genes. Also, in the Ly49h genomic transgenic mouse, Ly49H expression was restricted to NK cells as is the case with the endogenous receptor (48). This might be due to the lack of the adapter protein DAP12 in T and B cells. However, in mice carrying only an Ly49d cDNA transgene (driven by H-2Kb promoter and IgH enhancer) (32) some surface expression of
Ly49D was observed on T cells (49) indicating possible pairing of Ly49D with other adaptor proteins in these cells. The ability of T-cells to express some Ly49D on their surface in the absence of DAP12 in this transgenic mouse (49) supports the notion that the endogenous Ly49d promoter might be inactive in T-cells due to lack of necessary transcription factors and/or the presence of repressive epigenetic marks at regulatory sequences. Our results support both possibilities. The Ly49h 5’ region is hypermethylated in T-cells (Figure 2B) indicating a possible epigenetic hindrance to transcription of this gene in these cells. The lack of detectable Ly49d and h transcripts in T-cells (Figure 1A) also supports this notion. Furthermore, the 5’-UTRs of both Ly49d and h show different protein complex binding pattern compared to that of the inhibitory Ly49 genes (Figure 6B).

In contrast to the endogenous receptor, in the Ly49a genomic transgenic mouse, expression of Ly49A was seen on the majority of splenic B-cells for all transgenic lines regardless of copy number of the transgene (50). The 30Kb Ly49a genomic transgene contained the Pro-1 region and it was subsequently shown that the deletion of this region in the same transgene abrogated the expression of Ly49A in NK, T and B cells (50). Mice carrying a 79kb Ly49h genomic transgene spanning the complete length of the gene displayed Ly49H expression patterns similar to the B6 Ly49h gene indicating that the transcriptional regulatory sequences necessary for the ‘wild-type’ expression of this gene are contained within the 79kb region (48). Ly49h transcript and percentage of NK cells expressing Ly49H correlated in the transgenic mice and increased according to the transgene copy number, but only up to a certain threshold (48). This phenomenon indicates a possible control mechanism for an upper limit of expression likely at the level of transcription and/or at the level of NK selection.

The activating Ly49 genes evolve faster than their inhibitory family members (51). This is evident from the large variation in the number and sequence of the activating Ly49 genes and the high number of activating psuedogenes in nearly all the sequenced mouse strains (7,11). It is also possible that the regulatory sequences of the activating receptors evolved differently from that of the inhibitory receptors in order to allow a tighter control in NK cells and as a side effect of this evolution, mouse T-cells lost the ability to express these receptors. Deviation from stochastic expression might also allow more control on the number of activating receptors on the surface of NK cells because mono-allelic expression might lead to a lower number of receptors on the cell surface compared to bi-allelic expression (41). Hence, in a non-stochastic system, the expression level of these receptors is homogeneous among different NK cells. In conclusion, our results suggest that distinct modes of transcriptional regulation govern the expression of the activating and inhibitory Ly49 genes.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We sincerely thank Liane Gagnier, Hyun-Jung Goo, Vivian Lam and Dr Maura Gasparetto for technical assistance; Dr Florian Kuchenbauer, Dr Sally Rogers and Dr Nooshin Tabatabaei for sharing expertise; Dr Claudia Luther, Evette Haddad and Tim Halim for sharing reagents. We are very grateful to the staff of the TFL cell sorting facility and the BCCRC animal facility. We are thankful to Dr Stephen K. Anderson for the gift of 12A8 antibody.

FUNDING

Canadian Institutes of Health Research with core support from the British Columbia Cancer Agency. Funding for open access charge: Canadian Institutes of Health Research.

Conflict of interest statement. None declared.

REFERENCES


