Lai, C Benjamin and Zhang, Ying and Rogers, Sally L and Mager, Dixie L (2009). Creation of the two isoforms of rodent NKG2D was driven by a B1 retrotransposon insertion. Nucleic acids research, 37 (9), 3032-3043.

Published in Nucleic acids research, and available online at:

http://nar.oxfordjournals.org/content/37/9/3032.long

We recommend you cite the published version.

The URL for the published version is http://dx.doi.org/10.1093/nar/gkp174
Creation of the two isoforms of rodent NKG2D was driven by a B1 retrotransposon insertion

C. Benjamin Lai, Ying Zhang, Sally L. Rogers and Dixie L. Mager*

Terry Fox Laboratory, British Columbia Cancer Agency and Department of Medical Genetics, University of British Columbia, Vancouver, BC, Canada

Received September 17, 2008; Revised February 26, 2009; Accepted March 3, 2009

ABSTRACT
The mouse gene for the natural killer (NK) cell-activating receptor Nkg2d produces two protein isoforms, NKG2D-S and NKG2D-L, which differ by 13 amino acids at the N-terminus and have different signalling capabilities. These two isoforms are produced through differential splicing, but their regulation has not been investigated. In this study, we show that rat Nkg2d has the same splicing pattern as that of the mouse, and we mapped transcriptional start sites in both species. We found that the splice forms arise from alternative promoters and that the NKG2D-L promoter is derived from a rodent B1 retrotransposon that inserted before mouse–rat divergence. This B1 insertion is associated with loss of a nearby splice acceptor site that subsequently allowed creation of the short NKG2D isoform found in mouse but not human. Transient reporter assays indicate that the B1 element is a strong promoter with no inherent lymphoid tissue-specificity. We have also identified different binding sites for the ETS family member GABP within both the mouse and rat B1 elements that are necessary for high-promoter activity and for full Nkg2d-L expression. These findings demonstrate that a retroelement insertion has led to gene-regulatory change and functional diversification of rodent NKG2D.

INTRODUCTION
The activity of natural killer (NK) cells against potential target cells is regulated by a balance of inhibitory and activating receptors (1–3). One of the most well-studied activating receptors is NKG2D, which recognizes ligands often upregulated in response to stress, such as during virus infection or malignant transformation (4). A recent study using NKG2D deficient mice has demonstrated that expression of this receptor inhibits some types of cancer in spontaneous tumour models (5), indicating a significant role for NKG2D in tumour surveillance. The importance of NKG2D is also indicated by the existence of specific viral evasion mechanisms directed against it (6).

In the mouse, NKG2D is expressed on NK cells, activated CD8+ T cells, NKT cells, γδT cells and activated macrophages (7). Alternative splicing produces two different isoforms of NKG2D in mouse, one of which contains 13 extra amino acids at its N terminus (NKG2D-L) compared to the shorter isoform (NKG2D-S) (8). NKG2D-S can bind and signal through the adapters DAP10 and DAP12, while NKG2D-L can only bind DAP10 (8). Due to differential downstream pathways, DAP12 is capable of initiating cytotoxic as well as cytokine release responses while DAP10 can only activate an NK cell to kill (9). CD8+ T cells express only DAP10, whereas NK cells express both DAP10 and DAP12 (8). Resting mouse NK cells express very little NKG2D-S, but expression of this isoform is induced upon NK cell activation, enabling signalling through NKG2D-L/DAP10, NKG2D-S/DAP10 and NKG2D-S/DAP12 (8). These expression patterns of DAP10/12 and of mouse NKG2D isoforms result in functional diversification of this receptor. In mouse NK cells, NKG2D is strongly activatory, stimulating NK cells to attack bound targets. However, in CD8+ T cells, NKG2D serves only a co-stimulatory function (10). The situation in human is slightly different. The single human NKG2D isoform resembles mouse NKG2D-L and also only binds DAP10 (11–13). Not surprisingly, NKG2D stimulation using monoclonal antibodies in human NK cells initiates a cytotoxic response but not cytokine release (13). As in the mouse, NKG2D/DAP10 only acts in a co-stimulatory manner in human CD8+ T cells (14).

While many studies have concentrated on the function of NKG2D, surprisingly little is known about its transcriptional regulation. Here we show that the spliced isoforms of mouse NKG2D arise from the use of...
alternative promoters. We focus on the promoter for mouse NKG2D-L and show it to be derived from a rodent-specific B1 retrotransposon with a functional binding site for GABP, a member of the ETS family. We propose that insertion of this retrotransposon, through unusual donation of a polymerase II promoter, led to the alternative splicing patterns maintained in both mouse and rat that, in turn, result in functionally distinct NKG2D isoforms. These results document an intriguing example of how retrotransposons can affect the evolution of host gene expression and, more specifically, lend insight into the transcriptional regulation of NKG2D.

MATERIALS AND METHODS

Cell culture

LNK (mouse NK cell line), EL4 (mouse T cell line) and RNK16 (rat NK cell line) were cultured in RPMI 1640 medium supplemented with 10% (v/v) fetal bovine serum, 100 U/ml penicillin, 100 U/ml streptomycin, 50 μM 2-ME and 2 mM L-glutamine. LNK cells were further supplemented with 1000 U/ml IL-2 (Peprotech). NH3 3T3 (mouse embryonic fibroblast cell line) was cultured in Dulbecco’s modified Eagle’s medium supplemented with 10% fetal calf serum.

Mice

C57BL/6 (B6) mice were purchased from The Jackson Laboratory and bred in our animal facility. The use of animals for this study was approved by the Animal Care Committee of the University of British Columbia, and animals were maintained in accordance with the guidelines of the Canadian Council on Animal Care.

Reverse-transcriptase polymerase chain (RT–PCR) reaction and sequencing

All RT–PCR was performed using Superscript III (Invitrogen) using random primers according to manufacturer’s instructions. For Sprague–Dawley rat total RNA (Clonetech), PCR was then carried out using gene-specific primers: rat exon 1a forward, 5'-GATTCCACAAGAAAC AGGACCTC; rat exon 1b forward, 5'-TGGCATGTTTCTGTGACATATCCAGTT; rat exon 4 reverse, 5'-CAACAAGGACTCG ACAACAGAAGTGTCACC; rat B1 reverse: 5'-AAAGGATCCCATGGGTTCGTGATC; rat exon 6 reverse flanking, 5'-GGTTCCAGGCTTTGTTCTCA; rat reverse exon 6 nested, 5'-GGTTCCAGGTCTTGTAGCATATCCAGTT; rat reverse exon 6 flanking, 5'-GGTTCCAGGCTTTGTTCTCATA; rat reverse exon 4 nested, 5'-CAACAAGGACTCGAAACAGC.

Sequence alignment

Nucleotides positioned from −75 to +75 from the splice acceptor site of human NKG2D exon 2 were used as the reference sequence in UCSC Genome browser (http://genome.ucsc.edu). From the conservation tract, the orthologous sequences of rhesus, mouse, rat, guinea pig, hedgehog, horse, cow and armadillo were obtained. These sequences were aligned using Multiple Alignment using Fast Fourier Transform and G-INS-i strategy (MAFFT, http://align.bmr.kyushu-u.ac.jp/mafft/online/server/index.html). The same was done for −50 to +50 of the ATG in human exon 3 and for aligning PB1D9 with B1Mur3.

5'-RACE

C57BL/6 mouse spleens were homogenized. Total RNA was extracted from LNK cells and mouse splenocytes using Rneasy minikit and QIAshredder homogenizer (QIAGEN). Sprague–Dawley rat total RNA was purchased from Clonetech. 5'-RACE was performed using FirstChoice RLM-RACE kit (Ambion) according to manufacturer’s instructions. The primers used were as follows: mouse exon 6 reverse flanking, 5'-GGTTCCAGGTTTCTCTCTATT; mouse exon 6 reverse nested, 5'-TT GTTCTGTGACATATCCAGTT; rat reverse exon 6 flanking, 5'-GGTTCCAGGCTTTGTTCTCATA; rat reverse exon 4 nested, 5'-CAACAAGGACTCGAAACAGC.

Generation of luciferase reporter constructs

Promoter fragments were generated by PCR using forward primers with KpnI site inserted at the 5' end and reverse primers with NheI site inserted at the 5' end. PCR products were digested with KpnI/NheI (NEB) and cloned into pLG4.10 firefly luciferase promoter vector (Promega). The only exception is the generation of a construct containing the mouse B1 and a 391-bp fragment at separate sites. For this construct, the mouse B1 was cloned into the multiple cloning region using KpnI/NheI while the 391 bp fragment was inserted into a BamHI site downstream of the luciferase reporter gene. The primers used are as follows: mouse B1 forward: 5'-AAAGG TACCCCGGAAGTGTGTCACATA; mouse B1 reverse: 5'-AAAGCTAGCTTTTTTTAAAGACAGTGATC; mouse B1 -391 of B1 forward: 5'-AAAGGTACCCTGACATATTAT TATCTT; mouse -391 of B1 reverse: 5'-AAAGCTAGCA ATCTGTGCTAAATTTTCTG; mouse 391-bp BamHI forward: 5'-AAAGGATCTCCTGACATTATTATTTATCTT; mouse 391-bp BamHI reverse: 5'-AAAGGATCCCTCTGACTTTATCTT; mouse 391-bp BamHI reverse: 5'-AAAGGATCCTAGTCAATTTCTGCTTCAATTTTCTT; rat B1 forward: 5'-AAAGGTACC AACAGAAGCAGTGCTACCC; rat B1 reverse: 5'-AA GTAAGCTTTTCTAAGACAGTGTGTC; B1 3' 32-bp truncation reverse: 5'-AAAGCTAGGCGGTACCTCTAGATTTACAG; B1 3' 59-bp truncation reverse: 5'-AA GCTAGCTGCTCTGCTTCTGTAATG. The PB1D9 fragment truncated at 20 bp at the 5' end and 59 bp at the
Cell transfection and luciferase assays

LNK, EL4 and NIH 3T3 cells were transfected using Lipofectamine 2000 (Invitrogen) according to manufacturer’s instructions. Briefly, 2.5 L Lipofectamine 2000 according to manufacturer’s instructions. Briefly, 2.5 L Lipofectamine 2000 reagent was mixed with 1 mL of media supplemented with 20% fetal bovine serum or fetal calf serum and IL-2 where necessary. Cells were then incubated at 37°C for 16 h before assaying.

Luciferase activity was assayed using the Dual-Luciferase Reporter Assay System (Promega) according to the manufacturer’s instructions. Firefly luciferase activity was normalized to Renilla luciferase activity for each transfection and calculated as fold increase over pGL4.10-BASIC (pGL4B).

Electrophoretic mobility shift assays (EMSAs)

Double stranded probes were generated by annealing the following oligomers to their complementary oligomers: mouse wild type: 5'-AGGACAGATTCGGAAGT GGTGTCACATATC; mouse GABP mutant: 5'-A GACAGATTCCGTAGGGTTGTCACATATCT; mouse GABP interface: 5'-AGGACAGATTCCGTAGGGTTGTCACATATCT; mouse AP1 mutant: 5'-AGGACAGATTCCGTAGGGTTGTCACATATCT.

Electrophoretic mobility shift assays were performed as described before (15). Protein concentration was determined using the Qubit Quantitation Platform (Invitrogen). GABPz antibody (H-2 X) and c-Jun antibody (G-4 X) were purchased from Santa Cruz Biotechnology.

RESULTS

A rodent SINE element is linked to alternative exon 1 usage

To begin to dissect the transcriptional regulation of mouse NKG2D, we first examined the mouse Nkg2d sequence via the UCSC Genome browser (http://genome.ucsc.edu) and noticed a repetitive B1 element, termed PB1D9, inserted at the 3' end of intron 1. An orthologous SINE element is found in the rat (termed B1Mur3) but not in the human, NKG2D gene. The two B1 elements differ by a 5' 3-bp deletion in the mouse B1 and an internal 30-bp insertion in the rat B1, but are otherwise 84.4% identical (aligned using MAFFT, http://align.bmr.kyushu-u.ac.jp/mafft/online/server/index.html). Although annotated differently in mouse and rat by Repbase (http://www.girinst.org/repbase/index.html), these B1 elements represent the same ancestral insertion. To avoid confusion, we will refer to the mouse element as mB1 and the rat element as rB1. While there are two mouse NKG2D isoforms arising from transcripts that alternatively encode exon 1a or 1b, there is only one human NKG2D protein from transcripts that always include exon 2, which is homologous to mouse exon 1b (Figure 1A). The situation in rat is less clear. A rat mRNA initiating near the region homologous to mouse exon 1b has been previously described (16). However, a rat genomic sequence that is highly similar to mouse exon 1a is found upstream of this initiation exon (Figure 1A). To see if rat Nkg2d splicing patterns resemble those of the mouse, we undertook RT–PCR on Sprague–Dawley rat spleen total RNA. Primers were designed in rat sequences orthologous to mouse exons 1a and 4. Sequencing of the amplified...
product shows that a rat transcript indeed includes the region orthologous to mouse exon 1a and splices out the previously reported exon (Figure 1B). Furthermore, the splice sites used are the same as that of the mouse. We herein designate the exon found by RT–PCR as rat exon 1a and the alternative exon as rat exon 1b. As was done with the mouse (8), we also designate the isoform arising from exon 1a as rNKG2D-S and the isoform arising from exon 1b as rNKG2D-L. These results suggest that the Nkg2d splicing pattern in rat is identical to that of the mouse and that both species harbour a SINE element at the 3' end of intron 1. In comparison, human does not have the SINE element and splicing of NKG2D always gives rise to one protein.

The SINE element has polymerase II promoter potential

To determine if the SINE element is contributing an alternative promoter element, we first looked for transcription start sites (TSS) for mouse and rat NKG2D-L transcripts using 5' rapid amplification of cDNA ends (5'RACE). Primers were designed in mouse exon 6 and rat exons 4 and 6. We found two start sites within the mB1 element using C57BL/6 splenocytes (Figure 2A). Two additional sites were found using LNK, a mouse NK cell line that expresses both NKG2D isoforms (Figure 2B). Surprisingly, while one was located in the SINE element, the other lies downstream of the ATG1 start codon. Theoretically, this transcript could still be translated into NKG2D-S using the ATG2 start codon in exon 2. Further experiments are needed to verify this possibility. With RNA from Sprague–Dawley rat spleen, we found a dispersed pattern of TSS (Figure 2C). Two start sites were within the rB1 element while seven were downstream of the SINE but upstream of ATG1.

The 5’RACE experiments suggest that the rodent SINE element contains a functional promoter. To examine this possibility, we tested the mouse and rat SINE elements and the upstream region in luciferase reporter assays. Various fragments including and excluding mB1 and rB1 were cloned into pGL4B. These constructs were transiently transfected into LNK cells and promoter activity was measured. As shown in Figure 2D, both mB1 and rB1 elements by themselves display promoter activity. A fragment 391-bp upstream of mB1 shows no appreciable activity. Paradoxically, when this upstream region combined with mB1 was tested, the promoter activity was abolished. However, if this fragment was separated from mB1 in the same construct, only a small decrease in promoter activity was observed compared to a construct with mB1 alone. These results demonstrate that transcription not only starts within the SINE elements, but also that they contain polymerase II promoters.

The mouse B1 element contributes a functional GABP-binding site

Transposable elements harbouring functional enhancer sites are not uncommon (17). Thus, we interrogated the mouse mB1 sequence for potential transcription factor-binding sites using TESS (http://www.cbil.upenn.edu/cgi-bin/tess/tess). Curated results containing only mouse-specific, hematopoietic or lymphocyte-related transcription factors are shown in Figure 3A. To determine which of these sites contribute functional regulatory elements, luciferase reporter assays were carried out using truncated fragments of mB1 in pGL4B, transiently transfected in LNK cells. Surprisingly, fragments that truncate the 3' end of mB1 display much higher activity than the entire 109-bp mB1 (Figures 2D and 3B). A 77-bp fragment that eliminates the predicted TCF1, CEBPβ and YY1-binding sites at the 3' end shows 40-fold activity over basal. Of note, this piece removes an ATG at the +78 position. When the 77-bp segment was further truncated to 50 bp, eliminating two more TCF1 and one GATA3 binding site, 97-fold activity over basal was observed. Luciferase activity only dropped to 3-fold over basal when 20 bp were deleted at the 5' end of the 50-bp segment. This construct deletes a GABP, an AP1 and a GATA3-binding site. These results suggest that the essential transcription factor-binding sites are within the first 20 bp of mB1.

To further narrow down the essential transcription factor-binding sites, we mutated the potential GABP and AP1-binding sites and repeated the reporter assays in LNK cells. We chose to mutate the 77-bp construct, because the ATG at position +78 may be interfering with luciferase expression. Promoter activity does not
decrease appreciably using an AP1 mutant construct (Figure 3B). On the other hand, the GABP mutant and double mutant only show 2- and 4-fold activity over basal, respectively. These results indicate that the GABP-binding site is essential to enhancing transcription from the mB1 polymerase II promoter.

GABP binds the identified mB1 enhancer site in vitro

The core GGAA nucleotides in the binding sites of ETS family members, of which GABP is a member, are highly conserved. To determine if GABP is indeed binding to this mB1 element, EMSA were performed. When a DNA oligomer spanning the 5' end of mB1 was incubated with LNK nuclear extracts, two complexes are clearly visible in EMSAs, labelled I and II (Figure 4). Complex I contains GABP, since it can be supershifted by anti-GABP antibody. This complex is decreased or abolished using wild-type unlabelled oligomers in a dose-dependent manner (10-, 25-, 50- and 100-fold excess). Complex I is still clearly visible when unlabelled oligomers containing a mutated GABP site or with mutations in both the GABP and AP1 sites were used as competitors. The fact that these signals are lower compared to complex I without competitors is likely the result of non-specific competition at 100-fold excess. In agreement with the previous experiments, the oligomer did not seem to bind AP1. A competition assay using unlabelled oligomers with a mutated AP1 site did not produce any complexes. In addition, anti-AP1 antibody did not induce a supershift. The identity of complex II is currently unknown. It is possible that this complex contains GATA3, since the oligomer spans a GATA3-binding site. Furthermore, all mutant competitors were equally effective at abolishing this complex and neither anti-GABP nor anti-AP1 antibodies supershifted the complex. Taken together, these results suggest that GABP, but not AP1, is an important factor in transcription from mB1.

GABP binds a separate site in rat B1 in vitro

We wondered if GABP regulation of the B1 promoter is conserved between mouse and rat. An alignment of mB1

Figure 2. Detection of polymerase II promoter in SINE elements. (A) RLM-5’RACE of mouse Nkg2d. Sequence shown corresponds to chr6:129,572,548–29,572,772 of the mouse genome from UCSC Genome Browser (mm9 assembly). Underlined sequence represents the mB1 element inserted in front of exon 1b. Single stars mark transcription start sites found in C57BL/6 mouse spleen. Stacked double stars mark transcription start sites found in LNK mouse cell line. ATG1 start codon for mNKG2D-L is in brackets. Splice donor site of exon 1b is denoted with an arrow. (B) Detection of Nkg2d splice forms in LNK, RNK16, EL4 and NIH 3T3 cell lines with RT–PCR. The mouse exon–intron structure of Nkg2d is shown with exons labelled in roman numerals. Primers used are shown as arrowheads. (C) 5’RACE of rat Nkg2d. Sequence shown corresponds to chr4:166,923,873–166,924,142 of the rat genome from UCSC Genome Browser (rn4 assembly). Single stars mark transcription start sites found in Sprague–Dawley rat spleen. Underlined sequence represents the rB1 element. ATG1 start codon for rNKG2D-L is in brackets. Splice donor site of exon 1b is denoted with an arrow. (D) Luciferase reporter assay of SINE elements. Fragments including the rodent SINEs and 391-bp upstream of the mouse SINE were cloned into pGL4B and assayed using a dual-luciferase system. Promoter activity was normalized to co-transfected Renilla luciferase activity and calculated as fold above pGL4B. Data indicate mean (±SD) of three independent experiments.
with rB1 shows that rB1 has an AGAA sequence instead of the core GGAA sequence at the mB1 GABP-binding site (Figure 5A). However, a separate potential GABP-binding site is found in the internal insertion of rB1 using TESS (http://www.cbil.upenn.edu/cgi-bin/tess). These sites in rB1 were labelled GABP1 and GABP2, respectively. To see if either of these sites bind rat GABP, we subjected them to EMSA. DNA oligomers containing GABP2, but not GABP1, formed DNA–protein interactions with RNK16 (an Nkg2d expressing rat NK cell line, Figure 2B) nuclear extract (Figure 5B). The identity of the protein in this sole complex, labelled III, was confirmed to be GABP in a supershift assay using an anti-GABP antibody. Complex III can be abolished using wild-type unlabelled oligomers at 100-fold molar excess (lanes 7–9). Supershift assays were carried out using 5μg of anti-GABPz (lane 10) or anti-c-Jun (lane 11) antibodies.

Figure 4. GABP forms protein–DNA complex within mB1 in EMSA. Sequences of probes and competitors used are shown at the top. Mutations are underlined. As control, lane 1 contains only γ32P-labelled WT probe. LNK nuclear extract was incubated with labelled WT probe at 4°C for 20 min (lane 2). For competition assays, unlabelled competitors were incubated with LNK nuclear extract at 4°C for 20 min before addition of labelled WT probe for another 20-min incubation. WT competitors were used at 10-, 25-, 50- and 100-fold molar excess (lanes 3–6). Competitors with GABP, AP1 or both binding sites mutated were used at 100-fold molar excess (lanes 7–9). Supershift assays were carried out using 5μg of anti-GABPz (lane 10) or anti-c-Jun (lane 11) antibodies.

Figure 3. Identification of transcription factor binding sites that contribute to Nkg2d-L regulation. (A) TESS search for transcription factor binding sites in mB1. The sequence of the mB1 inserted in front of mouse Nkg2d exon 1b is shown. Potential transcription factor binding sites are indicated with brackets. Only mouse-specific, hematopoietic or lymphocyte-related factors are shown. (B) Luciferase reporter assay of truncated and mutated SINE elements. mB1 sequence was truncated at the 5′ and 3′ ends or mutated at specific transcription factor binding sites, cloned into pGL4B and assayed using a dual-luciferase system. Mutated sites are indicated with an X. Promoter activities were normalized to co-transfected Renilla luciferase activity and calculated as fold above pGL4B. Data indicate mean (±SD) of two independent experiments.

GABP is required for full NKG2D-L expression in vivo

To show that GABP is involved in the transcription of the natural Nkg2d locus, we adopted an RNAi strategy. We examined the effects of transient siRNA knockdown
of Gabp in the LNK cell line. Transcript levels of Gabp/C11, Nkg2d-S and Nkg2d-L were assayed via real-time RT–PCR. As shown in Figure 6, Gabp/C11 transcripts decreased 70% with Gabp/C11 siRNA. The negative control siRNA did not decrease the expression of the three genes compared to untreated samples. Gabp/C11 knockdown had no effect on Nkg2d-S levels but decreased the expression of Nkg2d-L by 24% compared to the negative control. This suggests that the mB1 associated Nkg2d-S, but not Nkg2d-L, is regulated at least in part by GABP.

Prevalence of potential GABP sites derived from B1 elements

Neither the mB1 nor the rB1 GABP-binding sites are present in the consensus sequences of 15 B1 subfamily members found in RepBase, suggesting that these sites were created through mutation after insertion. Nonetheless, we wondered how many B1 elements in the mouse genome contain such sites and if there is evidence of these elements contributing to a regulatory network as has been revealed recently for other repetitive elements (see ‘Discussion’ section). A bioinformatics scan of all potential GABP-binding motifs to see how many are located in B1 elements throughout the genome and those near promoters. A survey of all B1 elements (252 832) in the mouse genome revealed that 4877 (1.93%) contain at least one GABP-binding site. For the elements (70 804) that lie in a 10 kb window centred around TSSs of all mouse RefSeq genes, we found 1240 (1.75%) to have at least one GABP-binding site. We then subjected these RefSeq genes to gene ontology (GO) analysis using BiNGO (18). A comparison of the genes with any B1 elements in the 10-kb window to genes with GABP motif-associated B1 elements showed no differences between these groups.

Mouse B1 promoter activity is not tissue specific

Since GABP is a ubiquitous transcription factor, it likely does not confer tissue specificity. To investigate this issue, we tested a 3’ truncated fragment of mB1 in pGL4B using the EL4 and NIH 3T3 cell lines. EL4 is a mouse T-lymphoblastoid cell line and NIH 3T3 is a mouse embryonic fibroblast cell line. Neither cell lines express NKG2D (Figure 2B). The promoter activity of this construct in EL4 is comparable to that of LNK cells (Figure 7). The activity in NIH 3T3 is ~3-fold higher than in LNK and EL4 cells. This data demonstrates that the promoter and enhancers within the truncated mB1 retrotransposon are not tissue specific.
Splice signals are conserved in mammals except rodents

We wondered if the splicing pattern seen in rodents and human correlated with divergence of splice signals. Alignment of nine mammalian species at the intron1–exon1b border shows the absence of splice acceptor signals in rodents: mouse, rat and guinea pig (Figure 8A). It is possible that the selection pressure to maintain this splice signal was alleviated after the inserted SINE became capable of initiating transcription of \( \text{Nkg2d} \). An added effect is the emergence of \( \text{Nkg2d-S} \), initiating at exon 1a and skipping exon 1b now that the splice acceptor is eliminated. Interestingly, out of the nine mammals, only mouse and rat have a SINE B1 element at this splice junction. Thus, the evolution of guinea pig \( \text{Nkg2d} \) is different from the mouse and rat subclade. Furthermore, we noted that guinea pig and cow were the only species out of the nine examined that have lost the ATG1 start codon (Figure 8B). Both retain the ATG2 start codon while hedgehog does not (Figure 8C). This data shows that the SINE is specific to a subclade of rodents, and the splice acceptor signal at the intron1–exon1b junction is not present in rodents.

DISCUSSION

In the present study, we show that the two splice forms of mouse and rat \( \text{Nkg2d} \) arise from usage of alternative promoters. The promoter for mouse and rat \( \text{Nkg2d-L} \) is a SINE retrotransposon with a functional GABP-binding site. Comparative genomics data suggest a complex evolutionary model. Since guinea pigs lack the orthologous SINE element, we believe the B1 inserted after the divergence of guinea pig from the mouse and rat subclade. The identified GABP-binding site in the mouse is not found in the consensus B1 sequence, likely emerging through mutation. This novel regulatory element may have allowed the SINE to mediate transcription of the long isoform of \( \text{Nkg2d} \) directly from exon 1b, lifting the selection pressure to maintain a splice acceptor signal at the intron1–exon1b junction. The mutation of this splice acceptor site forces the transcript starting from exon 1a to splice directly to exon 2, thus creating \( \text{NKG2D-S} \). The mutation and its consequences on splicing we have just described may apply to both mouse and rat. The only difference between the two species is the position of the acquired GABP-binding site. Thus, this study presents an interesting example of convergent evolution to achieve the same means in animals that share similar niches.

The 13 amino-acid difference at the N-terminus of the two isoforms results in differential adapter affinities: \( \text{NKG2D-L} \) binds only DAP10 while \( \text{NKG2D-S} \) binds both DAP10 and DAP12 (8). Signal cascades of these two adapters are reviewed elsewhere (19). In essence, DAP12-coupled Syk and ZAP70 signalling are strongly stimulatory and initiate cytotoxic and cytokine release pathways. On the other hand, DAP10 signalling resembles co-stimulatory pathways of CD28 in T cells. While DAP10 can initiate NK killing response through Grb2/Vav1, it cannot trigger cytokine release. This is shown in the inability of anti-NKG2D treated, DAP12 knockout mouse NK cells to produce IFN\( \gamma \) (9). Malarkannan et al. (20) also showed that DAP12/Syk/ ZAP70-dependent activation of PLC\( \gamma 2 \) leading to the Bcl10/NF\( \kappa \)B pathway is absolutely essential to mouse NK cytokine production. Diefenbach et al. reported that while the long form of NKG2D is always expressed,
the short form is only expressed after NK priming with IL-2 or polynosinic:polycytidylic acid (8). Thus, it is possible that the two differentially expressed, signalling divergent splice forms provide the mouse and rat with beneficial functional flexibility in a context-dependent manner. In our model, we also believe that the losses of ATG1 in guinea pig and in cow are separate events of convergent evolution. The \textit{Nkg2d} genes in these two animals likely produce only the short protein form. Indeed, Fikiri \textit{et al.} (21) has reported the characterization of a short NKG2D form in cow. Although the flexibility of signalling made possible to mouse and rat may not be present in guinea pigs and cows, NKG2D in the latter two can at least signal through DAP12 and may therefore be beneficial. Last, the loss of ATG1 in guinea pigs may also have lifted selection pressure for retention of the splice acceptor site, allowing it to mutate away.

Our results that repetitive elements aid in the regulation of an NK receptor gene resonate with findings that link an AluSx and a LINE L1M5 element to the regulation of a human NK inhibitory receptor: \textit{KIR3DL1} (22). Of particular interest is the fact that this AluSx SINE element also contains a binding site for an ETS member, ELK1. Removal of this SINE element abolished \textit{KIR3DL1} distal promoter activity. The enrichment of transposable elements within immune genes has been noted before as a possible mechanism of rapid evolution and functional diversification (23). Our findings fit this model. While an Alu element may have contributed to the evolution of \textit{KIR3DL1} by supplying the gene with regulatory elements, a B1 has managed to both provide regulatory signals and aid in the formation of a second, functionally distinct isoform of NKG2D in mouse and rat.

Recent ChIP-sequencing studies have shown that some classes of repetitive elements bind specific transcription factors and establish transcription regulatory networks on a genome-wide basis. Binding for the transcription factor p53 was shown to be enriched in LTR10 and MER61 elements of the endogenous retrovirus ERV1 family (24,25). Bourque \textit{et al.} (25) further showed that ESR1, POU5F1-SOX2 and CTCF-binding sites are pervasive in MIR, ERVK and B2 repeats, respectively. Finally, Roman \textit{et al.} (26) described a new subclass of B1 element that harbours functional AhR and Slug-binding sites. Thus some classes of transposable elements are postulated to disperse specific control elements and drive regulatory expansion. We therefore questioned if B1 elements in the mouse genome may also be involved in the regulatory network of GABP. Since only human GABP ChIP-sequencing data is available (27) and B1 is a rodent-specific SINE element, we took an \textit{in silico} approach. Genomic distribution and GO analysis did not reveal signs of selective retention of B1 elements with GABP sites near promoters of any gene classes. Thus, there is presently no evidence for this transposon family acting as the foundation of GABP-regulatory networks. However, it is possible that other B1 elements do bind GABP \textit{in vivo} and could regulate genes. A ChIP-sequencing experiment in mouse would be needed to address this possibility directly.

B1 elements evolved from the 7SL RNA gene, a small non-coding RNA species. As such, they are thought to be transcribed through an internal polymerase III promoter (28). The B1 element identified that regulates mouse and rat \textit{Nkg2d} represents a novel source of polymerase II promoter signals. B2 elements, a close relative of B1 that evolved from tRNA, have been found to provide a polymerase II promoter to the mouse gene \textit{lama3} (29). To our knowledge, this is the first instance that B1 elements are shown to also contain polymerase II promoters.
Throughout, we cannot explain why promoter activity decreases when the mB1 element is combined with 391 bp of upstream region. At least two non-mutually exclusive explanations are possible. First, the 391-bp fragment may contain an uncharacterized repressor. A search using TESS locates multiple transcription factor-binding sites. It remains to be seen if any of these are functional and bind negative regulators. However, the promoter activity of a construct containing mB1 decreases only slightly if the 391-bp fragment was inserted downstream of the reporter gene. This suggests that any repressor present on the fragment would be position dependent. The other possibility is that of chromatin effects. The reporter assays were conducted using a transiently transfected plasmid. This experimental system does not take into account the three-dimensional conformations or the degree of histone packaging of the DNA fragment in the context of the Nkg2d gene. It is possible that the suppressive effect of the 391-bp segment is not seen in the endogenous context.

The mechanisms that underlie the increased promoter activity with 3' truncated mB1 are also unclear. Again, it is possible that the truncations removed repressor-binding sites. At least four sites found via TESS in the truncated regions bind transcription factors known to be both activating and repressing. Two sites can bind TCF1, a factor that is activating when bound to β-catenin but repressive when interacting with Groucho (30). YY1 is a repressor when not interacting with E1A (31), while E12/E47 has been shown to act as a repressor to E-cadherin (32). Whether any of these sites are functional remains to be shown. An additional possibility is the presence of an ATG sequence that is removed in all 3' truncation constructs. Others have noted the possibility that ATG start codons can compete with reporter gene start codon and decrease reporter signal (33).

The ETS member GABP was first implicated to regulate a range of house-keeping genes such as ribosomal protein genes and a cytochrome c oxidase gene (34–36). GABP has also been linked to critical S-phase genes and cell-cycle progression (37). Recently, a string of reports have shown that GABP plays an important role in immune functions. In the myeloid lineage, GABP is known to regulate the genes for integrin CD18 as well as IgA Fc receptor (38,39). The effect of GABP in lymphocyte biology is even more prominent. GABP was recently shown to regulate Fc receptor γ using monocyte, basophil and T cell lines (40). In T cells, GABP bridges JNK/SAPK signalling and enhanced IL-2 expression when artificially stimulated (41). GABP is also known to enhance transcription of Fas, IL-16 and IL7 receptor α (42–44). Last, GABP deficiency leads to defective B cell development, presumably because GABP binds to regulatory regions of Pax5 and CD79a (45). In this report, we link GABP directly to NK cells by showing its positive regulatory effects on mouse and rat Nkg2d.

ETS family members share highly similar binding sites. In a genome wide study, Hollenhorst et al. (46) reported widespread co-occupancy of GABP, ETS1 and ELF1 at promoters of human genes. We cannot rule out the possibility that the binding site we discovered bind these factors as well. GABP may compete with ETS1 or ELF1 for binding while all three may act synergistically to enhance Nkg2d-L transcription. This would reflect the CD18 example, where GABP and PU.1 compete for binding sites but cooperatively activate the gene (38). Alternatively, these ETS members may not be competing but rather forming a complex in mB1. Physical interaction between the three GABP, ETS or ELF1 has not been described, but GABP has been shown to bind two other ETS members, ER71 and ER81 (47).

While EMSA experiments showed robust binding of GABP to the B1 elements, knocking down Gabpa showed only a moderate decrease in Nkg2d-L transcript levels. This could be due to an incomplete knockdown of the transcription factor. We only achieved ~70% silencing. The remaining GABP may be enough to sustain the observed Nkg2d-L expression. Another possibility is that GABP augments but is not absolutely required for Nkg2d-L expression. Other transcription factors that bind the B1 promoter may need to be silenced to see a more dramatic decrease to Nkg2d-L mRNA levels. Finally, similar ETS family binding motifs may allow other ETS members to bind the B1 element and compensate for the loss of GABP.

In mouse, NKG2D is known to be expressed only on NK cells, activated CD8+ T cells and some NK/T cells (48). The fact that the mB1 promoter is capable of inducing reporter gene expression in NK, T and embryonic fibroblast cell lines leaves the question of NKG2D tissue-specific expression open. This is perhaps not surprising since GABP is a factor expressed in a broad range of tissues (47). We can think of at least two ways that specificity of Nkg2d-L expression can be achieved. There could be additional enhancer sites that bind tissue-specific transcription factors we have not uncovered. Another possibility is epigenetic regulation through DNA methylation and histone modifications. Although the upstream regions of Nkg2d exons 1a and 1b are CpG poor, the expression of Ly49a, another NK receptor gene, was shown to correlate with methylation at a CpG poor promoter (49). The expression of other NK receptor genes within the NK complex, Ly49g and Nkg2a, have also been linked to CpG methylation or histone acetylation (50,51). In addition, Santourlidis et al. (52) suggests in a recent report that euchromatic histone marks open up the KIR loci early in development but use DNA methylation to silence specific KIR members and establish clonal expression later on.

As more studies emerge, the differences between human and mouse NK cells are starting to be appreciated. While mouse NKG2D-L is the isoform that more closely mimics human NKG2D, there are notable discrepancies. NKG2D is expressed on human CD8+ T cells regardless of activation status in human, but only on activated mouse CD8+ T cells. The receptor is also found on γδT cells of intestinal intraepithelial origin in human but not mouse (48). Since human NKG2D does not harbour a SINE element at the same position as the mouse, the mode of regulation is unknown and likely different from mouse. Furthermore, nothing is known about the regulation of the mouse short isoform, which responds to IL-2 and polyinosinic:polycytidylic acid treatment (8).
Dissecting the genomic and epigenetic factors that control these processes have therapeutic potential since it may allow the manipulation of an important NK-activating receptor involved in cancer immunosurveillance, viral clearance and autoimmune diseases.

ACKNOWLEDGEMENTS

We thank Drs Irina Masakova and Ping Xiang for valuable advice and Drs Arefeh Rouhi and Fumio Takei for advice and critical reading of the manuscript.

FUNDING

Canadian Institutes of Health Research [15082 and 10825] and the Cancer Research Society, BC Cancer Agency; University of British Columbia Graduate Fellowship (to C.B.L.); Alexander Graham Bell Graduate Scholarship, University of British Columbia Graduate Fellowship (to V.H.); and the Cancer Research Society, BC Cancer Agency; Canadian Institutes of Health Research [15082 and 10825]

REFERENCES

3043


