Conservation and Variation in Human and Common Chimpanzee CD94 and NKG2 Genes

Benny P. Shum, Laura R. Flodin, David G. Muir, Raja Rajalingam, Salim I. Khakoo, Sophia Cleland, Lisbeth A. Guethlein, Markus Uhrberg and Peter Parham

*J Immunol* 2002; 168:240-252;
doi: 10.4049/jimmunol.168.1.240
http://www.jimmunol.org/content/168/1/240

---

References
This article cites 66 articles, 18 of which you can access for free at:
http://www.jimmunol.org/content/168/1/240.full#ref-list-1

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

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

Email Alerts
Receive free email-alerts when new articles cite this article. Sign up at:
http://jimmunol.org/alerts
Conservation and Variation in Human and Common Chimpanzee CD94 and NKG2 Genes

Benny P. Shum, Laura R. Flodin, David G. Muir, Raja Rajalingam, Salim I. Khakoo, Sophia Cleland, Lisbeth A. Guethlein, Markus Uhberg, and Peter Parham

To assess polymorphism and variation in human and chimpanzee NK complex genes, we determined the coding-region sequences for CD94 and NKG2A, C, D, E, and F from several human (Homo sapiens) donors and common chimpanzees (Pan troglodytes). CD94 is highly conserved, while the NKG2 genes exhibit some polymorphism. For all the genes, alternative mRNA splicing variants were frequent among the clones obtained by RT-PCR. Alternative splicing acts similarly in human and chimpanzee to produce the CD94B variant from the CD94 gene and the NKG2B variant from the NKG2A gene. Whereas single chimpanzee orthologs for CD94, NKG2A, NKG2E, and NKG2F were identified, two chimpanzee paralogs of the human NKG2C gene were defined. The chimpanzee Pt-NKG2CII gene encodes a protein similar to human NKG2C, whereas in the chimpanzee Pt-NKG2CHII gene the translation frame changes near the beginning of the carbohydrate recognition domain, causing premature termination. Analysis of a panel of chimpanzee NK cell clones showed that Pt-NKG2CII and Pt-NKG2CHII are independently and clonally expressed. Pt-NKG2CII and Pt-NKG2CHII are equally diverged from human NKG2C, indicating that they arose by gene duplication subsequent to the divergence of chimpanzee and human ancestors. Genomic DNA from 80 individuals representing six primate species were typed for the presence of CD94 and NKG2. Each species gave distinctive typing patterns, with NKG2A and CD94 being most conserved. Seven different NK complex genotypes within the panel of 48 common chimpanzees were due to differences in Pt-NKG2C and Pt-NKG2D genes. The Journal of Immunology, 2002, 168: 240–252.

---

H
uman NK cell receptors specific for HLA class I molecules are encoded in two genetic complexes, both situated on different chromosomes from the MHC on human chromosome 6 (1, 2). The killer cell Ig-like receptors (KIRs) are encoded on the leukocyte receptor complex or cluster (LRC) on human chromosome 19q13.4 (3–6) and the lectin-like CD94 and NKG2 receptors are encoded within the NK complex (NKC) on human chromosome 12p12.3-p13.1 (7–12). In human, one CD94 and five NKG2 (A/B, C, D, E/H, F) genes are located within the NKC (7–12).

The CD94 and NKG2A polypeptides form a heterodimeric receptor (13, 14) that recognizes the MHC nonclassical class I HLA-E ligand complexed with peptides derived from the leader sequences of certain classical HLA-A, B, C, and nonclassical HLA-G H chains (15–19). Ligand recognition transduces inhibitory signals in NK cells via the immunomodulatory tyrosine-based inhibition motifs on the cytoplasmic tail of NKG2A (20–22). The CD94:NKG2C heterodimer is a receptor with ligand specificity similar to CD94:NKG2A, but which generates activating signals within NK cells (22, 23). Human NKG2D is a homodimeric activating receptor that interacts with MHC class I-like chain A (24, 25), a MHC class I-like protein with considerable polymorphism, and with unique long binding proteins (ULBPs), newly identified molecules distantly related to MHC class I (26). In NK cells, activating signals are conveyed via positively charged amino acid residues in the transmembrane domains of activating receptors (NKG2C2, D) that interact with signal-transducing adaptor molecules such as DAP-10 and DAP-12 (27, 28). The exact roles of human NKG2E and NKG2F are uncertain, but both encode a positively charged lysine residue in the transmembrane domain, as does NKG2C.

The KIR genes of the human LRC encode receptors that are specific for polymorphic determinants of HLA-A, B, and C, and also for HLA-G (1, 29). Previous work from our laboratory revealed diversity of KIR haplotypes in the human population, with respect to the number and content of KIR genes (30) and also allelic polymorphism at individual genes (31–34). Furthermore, comparison of human and chimpanzee KIR revealed many species-specific features to the KIR gene systems (35, 36). Given the diversity and rapid evolution of the KIR genes it became of particular interest to investigate intraspecies and interspecies diversity of NKC genes that encode NK cell receptors for MHC class I molecules. Therefore, we have investigated the diversity and variation in the CD94 and NKG2 genes of humans and apes.

---

Materials and Methods

Source materials from humans and apes

Healthy human donors recruited from our laboratory represent a variety of ethnic and racial backgrounds, as shown in Fig. 1. Tissue samples from non-human primates were purchased from the Yerkes Regional Primate Center at Emory University (Atlanta, GA) and from the Laboratory for Experimental Medicine and Surgery in Primates (New York University...
Medical Center, Tuxedo, NY). Blood was drawn from animal subjects as part of routine health examinations. Total cytoplasmic RNA was prepared from PMBC isolated from human and chimpanzee donors. In addition, RNA was extracted from NK cell clones NK1.3 and M1.1, established from chimpanzees Cathy and Mason, respectively (35); a splenic sample from the chimpanzee Edwina was also used as source material. Genomic DNA of higher primates used for PCR typing analysis was isolated from previously established EBV-transformed B cell lines (36, 37). These include samples from the family Hominidae: humans (Homo sapiens, n = 11), common chimpanzees (Pan troglodytes, n = 48), pygmy chimpanzees or bonobos (P. paniscus, n = 11), gorillas (Gorilla gorilla, n = 2), and orangutans (Pongo pygmaeus, n = 5); and from the family Hylobatidae: common gibbons (Hylobates lar, n = 3).

cDNA analysis of CD94 and NKG2 genes

Total cytoplasmic RNA was isolated with the reagent RNazol following manufacturer’s protocol (Tel-Test, Friendswood, TX). First-strand cDNA was prepared from total RNA using murine Moloney leukemia virus-reverse transcriptase (Life Technologies, Rockville, MD) by standard procedures (38).

cDNA was amplified by PCR in 1× AmpliTaq buffer (Applied Biosystems, Foster City, CA), 0.25 mM of each dNTP, 0.8 pM each of a sense (forward) and antisense (reverse) primer, and 1 U of AmpliTaq DNA polymerase (Applied Biosystems). The following PCR conditions were used: initial denaturation at 96°C for 1 min; then 30 cycles of denaturation at 94°C for 30 s, primer annealing at 52–60°C for 30 s, and extension at

Table I. Oligonucleotide primers

<table>
<thead>
<tr>
<th>Name</th>
<th>Orientation</th>
<th>Sequence (5’→3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>94-CyF1</td>
<td>Sense</td>
<td>TGG CAG TGT TTA AGA CCA CT</td>
</tr>
<tr>
<td>94-E6R1</td>
<td>Sense</td>
<td>TGA ACA GAA TCA ACC ACA GAA</td>
</tr>
<tr>
<td>94g-E5F1</td>
<td>Sense</td>
<td>TTT TAT GAG TCT CAG TCA ACA</td>
</tr>
<tr>
<td>94g-E5R1</td>
<td>Antisense</td>
<td>TCC TTA GGA TTA TAC GCT ATG</td>
</tr>
<tr>
<td>A-E2F1</td>
<td>Sense</td>
<td>TCA AAA AAC TTT GCA GGA TTT</td>
</tr>
<tr>
<td>A-E5R1</td>
<td>Antisense</td>
<td>TGA GGA TGG TGA AAT GAT GGA</td>
</tr>
<tr>
<td>Ag-E1F1a</td>
<td>Sense</td>
<td>GAA CAG GAA ATA ACC TAC GC</td>
</tr>
<tr>
<td>Ag-E2R1</td>
<td>Antisense</td>
<td>AGG CCA TTA AGA TAA GAC AG</td>
</tr>
<tr>
<td>βm-m-F1</td>
<td>Sense</td>
<td>TGT GTC TGG TTT TCA TCC ATC</td>
</tr>
<tr>
<td>βm-m-R1</td>
<td>Sense</td>
<td>CTA TCT TGG GTG TGT ACA AAG</td>
</tr>
<tr>
<td>CD94-E4F2</td>
<td>Sense</td>
<td>CCG GTG CAT TCA AAC TAT ATC</td>
</tr>
<tr>
<td>CD94-E5R2</td>
<td>Antisense</td>
<td>GGT GTG CTC TCT ACT GTA AGA</td>
</tr>
<tr>
<td>CD94-PCR1</td>
<td>Sense</td>
<td>CCT TCT TTA CTT TCC ATG TGT</td>
</tr>
<tr>
<td>Ge-E1F1</td>
<td>Sense</td>
<td>CCT CAA AAT CCT TCC</td>
</tr>
<tr>
<td>Ge-E4R1</td>
<td>Antisense</td>
<td>GGA GTT TTA CCA GGA TGA AGT</td>
</tr>
<tr>
<td>g2C21-F1</td>
<td>Sense</td>
<td>GTG GAT TTA CAA TAA TAT ATC</td>
</tr>
<tr>
<td>cG2C1-R1</td>
<td>Antisense</td>
<td>CAT CAC GAC ACA AAG CA AAG</td>
</tr>
<tr>
<td>cG2C2-F1</td>
<td>Sense</td>
<td>GGT GAT TTA CAA TGA TGA ATT</td>
</tr>
<tr>
<td>cG2C2-R1</td>
<td>Antisense</td>
<td>CAT CAC GAC ACA AAG TAA CG</td>
</tr>
<tr>
<td>CI-E1F1</td>
<td>Sense</td>
<td>GAA CAG GAA ATA TTA TCC CA A</td>
</tr>
<tr>
<td>CI-E6R1</td>
<td>Antisense</td>
<td>AGG CCA GAA TCA TTA AGA TAT</td>
</tr>
<tr>
<td>CII-E1F1</td>
<td>Sense</td>
<td>GAA CAG GAA ATA TTA TCC CA A</td>
</tr>
<tr>
<td>CII-E6R1</td>
<td>Antisense</td>
<td>AGG CCA GAA ATA TTA TCC CA A</td>
</tr>
<tr>
<td>D-CyF1</td>
<td>Sense</td>
<td>TGT GTC TGG GTT TCA TCC ATC</td>
</tr>
<tr>
<td>D-E10R1</td>
<td>Antisense</td>
<td>CTT TAA AGC TGG AGG CAT AGA</td>
</tr>
<tr>
<td>Dfor1</td>
<td>Sense</td>
<td>AGC AAT AGT GGG TGA TGT ATT</td>
</tr>
<tr>
<td>DFORa</td>
<td>Sense</td>
<td>TAT CCA GAA TCA AGA TCG TCT</td>
</tr>
<tr>
<td>Dg-E4F1</td>
<td>Sense</td>
<td>GTG GAT TCA GTG TGG GAG TGC</td>
</tr>
<tr>
<td>Dg-E5R1</td>
<td>Antisense</td>
<td>TGG CTT TGA CCA TCG TGT TGA</td>
</tr>
<tr>
<td>DREV</td>
<td>Antisense</td>
<td>GTC TGT AGT CTA AGA AAT</td>
</tr>
<tr>
<td>Drev1</td>
<td>Antisense</td>
<td>TTT TAG GAC ATG GGC CAC AGT A</td>
</tr>
<tr>
<td>E-E7R1</td>
<td>Antisense</td>
<td>ATG ATG AAA CCC GTT CTA ATG</td>
</tr>
<tr>
<td>E-R1</td>
<td>Antisense</td>
<td>GGG CGG TGG CTT GTG TCA GTA</td>
</tr>
<tr>
<td>E-R4</td>
<td>Antisense</td>
<td>GAG GCC AAG ATG TTG TTA T</td>
</tr>
<tr>
<td>ECF-E1F1</td>
<td>Sense</td>
<td>CGA ACA GAA ATT CTA CCA AGT</td>
</tr>
<tr>
<td>ECF-Fi</td>
<td>Sense</td>
<td>TTA TCA TAG AGC ACA GTC CCT</td>
</tr>
<tr>
<td>ECF-F4</td>
<td>Sense</td>
<td>CAG TCC TCT ACA TCA CAC AGC</td>
</tr>
<tr>
<td>hNKG2AC-R1</td>
<td>Antisense</td>
<td>GGC AAC CAC TAT TCT ACT TCC</td>
</tr>
<tr>
<td>hNKG2A-PCR1</td>
<td>Sense</td>
<td>CTG GGG ACA GAA GAG TAC AGT</td>
</tr>
<tr>
<td>hNKG2ABC-R3</td>
<td>Antisense</td>
<td>TAT AGA AAG CAG ACT GGA GTT</td>
</tr>
<tr>
<td>hNKG2AC-PR1</td>
<td>Antisense</td>
<td>GCC CGA CAC AAA TGC TAG GAT</td>
</tr>
<tr>
<td>NKG2ABC-R2</td>
<td>Sense</td>
<td>GTG GCC ATT GTC CTG AGG AGT</td>
</tr>
<tr>
<td>NKG2ABC-R3</td>
<td>Antisense</td>
<td>ATT AGA AAG CAG ACT GGA GTT</td>
</tr>
<tr>
<td>NKG2ABC-F2</td>
<td>Sense</td>
<td>ATC AGG CCA GTG TGG ATC TCC</td>
</tr>
<tr>
<td>NKG2ABC-F3</td>
<td>Sense</td>
<td>ATC AGG CCA GTG TGG ATC TCC</td>
</tr>
<tr>
<td>NKG2ABC-PCRF</td>
<td>Sense</td>
<td>ATC ACA CAG CAG GAG AGA</td>
</tr>
<tr>
<td>NKG2ABC-PCRR</td>
<td>Antisense</td>
<td>TGA GGA TGG TGA AAT GAT GGA</td>
</tr>
<tr>
<td>NKG2ABC-R2</td>
<td>Antisense</td>
<td>ATA AAA TGT ATC TGA TGC ACT</td>
</tr>
<tr>
<td>NKG2ABC-R3</td>
<td>Antisense</td>
<td>TAT AGA AAG CAG ACC GGA GTT</td>
</tr>
<tr>
<td>NKG2C-F1</td>
<td>Sense</td>
<td>ATC ACA CAG CAG GAG AGA</td>
</tr>
<tr>
<td>NKG2C-PRF</td>
<td>Sense</td>
<td>AGC AGC GTC CCT CAC ATC ACA</td>
</tr>
<tr>
<td>NKG2C1-R1</td>
<td>Antisense</td>
<td>CAC AAA GAA ACA TTA TAC AGA</td>
</tr>
<tr>
<td>NKG2CN-R1</td>
<td>Antisense</td>
<td>TGA GCA AAA TGA GCC CA</td>
</tr>
<tr>
<td>NKG2F1</td>
<td>Sense</td>
<td>CTG CAG TTT GCC TAT ACC AGG</td>
</tr>
<tr>
<td>NKG2F-for</td>
<td>Sense</td>
<td>GCA TGG TCC TGA TGG CCA CTC T</td>
</tr>
<tr>
<td>NKG2F-R1</td>
<td>Antisense</td>
<td>ACA TTT ATG GAA GCA TGG GTT</td>
</tr>
<tr>
<td>NKG2F-rev</td>
<td>Antisense</td>
<td>ACA GTG GCC ATC AGG ACA ATG C</td>
</tr>
</tbody>
</table>
72°C for 1 min; followed by 72°C for 10 min to promote complete extension.

Primers were designed from available human cDNA sequences in the GenBank database. Initially, certain primer sets were designed to amplify more than one NKG2 gene; gene-specific primer sets were used subsequently. The following oligonucleotide primers were used for our cDNA amplifications: CD94-PCR and CD94-PCRrr (for CD94 and Prt-CD94), NK2G2ABC-PCRFR and NK2G2ABC-PCRRI (for NK2G2A/B and Pt-NKG2A/B), 0.25 mM of each dNTP, 0.15 U of each primer, 2.5 U of AmpliTaq DNA polymerase (Applied Biosystems), and 200 ng of genomic DNA isolated from a panel of higher primates using gene-specific oligonucleotide primers. The presence of PCR-amplified cDNA was cloned into the pCR-TOPO plasmid vector (Invitrogen). Plasmid DNA from multiple clones from each amplification were selected for analysis. Nucleotide sequences of plasmid DNA insert were determined on both strands using the Big-Dye Terminator Cycle Sequencing kit (Applied Biosystems) and a 377 Automated DNA sequencer (Applied Biosystems). In addition to the universal T3 and T7 oligonucleotide primers, the following primers were used to sequence completely the amplified cDNA; for NK2G2AB, NKG2ABC-R2 and NKG2ABC-R3, for Pt-NKG2A/B-R2, NKG2ABC-R2 and NKG2ABC-R3; for NK2G2C and NK2G2C-R1; for NK2G2D and NK2G2D-R1, F-DORa and DREV (Table I). DNA sequences were assembled and analyzed using the computer program AutoAssembler (version 2.1; Applied Biosystems) and the Wisconsin Package sequence analysis software (version 10.1; Genetics Computer Group, Madison, WI).

Nucleotide sequences

All common chimpanzee CD94 and NKG2 genes were assigned the prefix “Prt.” for P. troglodytes. New nucleotide sequences determined here have been deposited into the GenBank database under the accession numbers AF259061–63, AF260135–6, and AF350005–19 (Table II). Sequences previously reported by us have the accession numbers AF259054–60 and AF350016–18 (Table I). Sequences isolated from this study and reported by others are also shown in Table II; the “NM” prefix in the accession number denotes GenBank reference sequences. Numerous alternatively spliced variants were also identified in our analysis; these have not been deposited into the GenBank database but are available from B. P. Shum by request.

Analysis of the intron 4/exon 5 boundary of Prt-CD94

To examine whether alternative splicing could generate the chimpanzee CD94 and CD94B variants, which differ by insertion of three nucleotides (encoding a glutamine) between exon 4 and exon 5 in the cDNA, we determined the DNA sequence at the intron 4/exon 5 boundary of the Prt-CD94 gene. Genomic DNA from the common chimpanzee Cathy was amplified by PCR with oligonucleotide primers CD94-CD94E4 and CD94-CD94E5R2 (Table I). The 50-μl reaction contained 1× AmpliTaq buffer (Applied Biosystems), 0.25 mM of each dNTP, 0.15 U of each primer, 2.5 U of AmpliTaq DNA polymerase (Applied Biosystems), and 200 ng of genomic DNA. PCR was performed with an initial denaturation at 96°C for 2 min; 30 cycles of denaturation at 96°C for 30 s, annealing at 60°C for 30 s, extension at 72°C for 2 min; and a final extension at 72°C for 10 min. The PCR product (~1969 bp) was cloned into the pCR-TOPO vector (Invitrogen). Fifteen plasmid clones were isolated and their nucleotide sequence was determined using T3 and T7 universal primers.

Transmission of the common chimpanzee Prt-NKG2CII and CII genes

To examine the transcription of Prt-NKG2CII and Prt-NKG2CII genes in NK cells, we amplified cDNA by RT-PCR with gene-specific primers from individual NK cell clones; these NK clones established from the common chimpanzee Cathy have been described previously (35). Each cDNA typing reaction of 25-μl volume contained 1× AmpliTaq buffer (Applied Biosystems), 0.25 mM of each dNTP, 0.15 U of each primer, and 1 U of AmpliTaq DNA polymerase (Applied Biosystems). PCRs were performed with an initial denaturation at 96°C for 1 min; followed by 35 cycles of denaturation at 94°C for 15 s, annealing at 52°C for 15 s, and extension at 72°C for 30 s; then a final extension at 72°C for 10 min. Amplified product was visualized under UV light after separation by electrophoresis on a 1.2% agarose gel following manufacturer’s instructions (Invitrogen). Gene-specific oligonucleotide primers used for cDNA typing were as follows: for CD94 and Prt-CD94, 94-CyFl and 94-E6R1; for NK2G2A and

\[
\begin{array}{c|c|c}
\text{Table II. CD94 and NKG2 sequences} \\
\hline
\text{Sequence} & \text{Gene Symbol*} & \text{Accession No.} \\
\hline
\text{CD94} & \text{KLRC1} & \text{NM_002262} \\
\text{Prt-CD94} & \text{NM_002262} & \text{AF259054} \\
\text{NK2G2D} & \text{D1252489E} & \text{NM_007360} \\
\text{NK2G2D01} & \text{D1252489E} & \text{AF259055} \\
\text{NK2G2D02} & \text{D1252489E} & \text{AF260135} \\
\text{NK2G2D03} & \text{D1252489E} & \text{AF260136} \\
\text{NK2G2b} & \text{KLRCl} & \text{NM_002259} \\
\text{Pt-NKG2D} & \text{KLRCl} & \text{AF259055} \\
\text{Pt-NKG2a01} & \text{KLRCl} & \text{AF350005} \\
\text{Pt-NKG2a02} & \text{KLRCl} & \text{AF350056} \\
\text{Pt-NKG2a03} & \text{KLRCl} & \text{AF350005} \\
\text{NK2G2c01} & \text{KLRCl} & \text{Y13055} \\
\text{NK2G2c02} & \text{KLRCl} & \text{AF260134} \\
\text{Pt-NKG2c01} & \text{KLRCl} & \text{AF259057} \\
\text{Pt-NKG2c02} & \text{KLRCl} & \text{AF259058} \\
\text{Pt-NKG2c03} & \text{KLRCl} & \text{AF259059} \\
\text{Pt-NKG2c04} & \text{KLRCl} & \text{AF259060} \\
\text{Pt-NKG2Cl01} & \text{KLRCl} & \text{AF259061} \\
\text{Pt-NKG2Cl02} & \text{KLRCl} & \text{AF259062} \\
\text{Pt-NKG2Cl03} & \text{KLRCl} & \text{AF350016} \\
\text{Pt-NKG2Cl04} & \text{KLRCl} & \text{AF350017} \\
\text{Pt-NKG2Cl05} & \text{KLRCl} & \text{AF350006} \\
\text{Pt-NKG2Cl06} & \text{KLRCl} & \text{AF350007} \\
\text{Pt-NKG2Cl07} & \text{KLRCl} & \text{AF350008} \\
\text{Pt-NKG2Cl08} & \text{KLRCl} & \text{AF350009} \\
\text{Pt-NKG2Cl09} & \text{KLRCl} & \text{AF350010} \\
\text{Pt-NKG2Cl10} & \text{KLRCl} & \text{AF350011} \\
\text{Pt-NKG2Cl11} & \text{KLRCl} & \text{AF350012} \\
\text{Pt-NKG2Cl12} & \text{KLRCl} & \text{AF350013} \\
\text{Pt-NKG2Cl13} & \text{KLRCl} & \text{AF350014} \\
\text{Pt-NKG2Cl14} & \text{KLRCl} & \text{AF350015} \\
\text{Pt-NKG2F01} & \text{KLRCl} & \text{U96845} \\
\text{Pt-NKG2F02} & \text{KLRCl} & \text{NM_013431} \\
\text{Pt-NKG2F03} & \text{KLRCl} & \text{AF350018} \\
\text{Pt-NKG2F04} & \text{KLRCl} & \text{AF350019} \\
\text{Pt-NKG2F05} & \text{KLRCl} & \text{AF350020} \\
\text{Pt-NKG2F06} & \text{KLRCl} & \text{AF350001} \\
\text{Pt-NKG2F07} & \text{KLRCl} & \text{AF350012} \\
\text{Pt-NKG2F08} & \text{KLRCl} & \text{AF350013} \\
\text{Pt-NKG2F09} & \text{KLRCl} & \text{AF350014} \\
\text{Pt-NKG2F10} & \text{KLRCl} & \text{AF350015} \\
\end{array}
\]

* Human Gene Mapping Workshop approved symbols.
Phylogenetic analysis

Sequences were aligned with the Pileup program of the Wisconsin Package (version 10.1; Genetics Computer Group) and manually adjusted afterward. Phylogenetic analysis was performed using the PAUP* 4.0 software package (Phylogenetic Analysis Using Parsimony (*and Other Methods), version 4.0b4a; Sinauer Associates, Sutherland, MA). The neighbor-joining method was used to construct dendrograms. Confidence of individual nodes was evaluated by 1000 bootstrap replications, and majority-rule consensus trees were generated.

Results

Allelic and splice variations of human and common chimpanzee NKC genes

To investigate the polymorphism and interspecies divergence of CD94 and NKG2 genes, we determined their coding region cDNA sequences from four common chimpanzees and from healthy human donors. The panel of 14 human donors represented major ethnic groups, but not all of the genes were characterized for all donors (Fig. 1). No allelic polymorphism was found for CD94 in either human or common chimpanzee, whereas polymorphism was seen for all NKG2 genes in either one or both species. Many alternatively spliced variants emerged from our analysis, including the previously described CD94B and NKG2A/B (GenBank accession no. AJ000001) (7). Alternatively spliced variants of several types were identified: ones lacking exons, ones including introns, ones generated from cryptic splice sites, and one having combinations of the above. Assignments for cryptic splice sites were made from comparison of the variant cDNA sequence with the sequence of full-length cDNA and/or corresponding gene sequences. The presence of highly conserved “GT” or “AG” dinucleotides at most of the alternatively spliced junctions revealed the use of cryptic splice sites (data not shown) (39).

CD94

Data from six humans and four chimpanzees reveals no CD94 polymorphism in either species (Fig. 1). Human and chimpanzee CD94 differ at only two nucleotide positions in the coding sequence, one of which is nonsynonymous and the other synonymous (Fig. 2).

CD94B is a previously reported cDNA variant of human CD94 (GenBank accession no. AJ000001) that differs from CD94 by the coding of an additional glutamine residue (position 183 in Fig. 3). We found variants of both types in both human and chimpanzee (Figs. 1, 3, and 4). Comparison of the two cDNA sequences with the CD94 gene sequence indicates that the difference between CD94 and CD94B is due to alternative patterns of mRNA splicing at the junction between intron 4 and exon 5 (Fig. 4B). To determine whether the common chimpanzee Pt-CD94 and Pt-CD94B variants can be generated in the same manner, we amplified, cloned, and sequenced a genomic DNA segment spanning the boundary of intron 4 and exon 5 from chimpanzee Cathy, who expressed both Pt-CD94 and Pt-CD94B (Fig. 1B). Fifteen clones were analyzed and they all yielded a sequence that was identical to that of the human gene for 40 nt flanking each side of the intron/exon boundary. Thus the mechanism of alternative splicing by which the two CD94 variants are produced appears the same in chimpanzee and human.

Clones for five additional alternatively spliced forms of Pt-CD94 were obtained (Fig. 4). Four of these variants were spliced like CD94B but had additional changes. One lacked exon 2 (variant 1), one lacked exon 4 (variant 2), one lacked both exons 2 and 4 (variant 3), and one lacked the 3’ half of exon 4 (variant 4). Variant 5 was similarly missing the 3’ half of exon 4, but it is spliced like CD94; this truncation of exon 4 was likely due to a

A. Human Homo sapiens

<table>
<thead>
<tr>
<th>Donor</th>
<th>CD94</th>
<th>NKGD2</th>
<th>NKG2A (B)</th>
<th>NKG2B</th>
<th>NKG2C</th>
<th>NKG2E</th>
<th>NKG2F</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>01</td>
<td>02</td>
<td>03</td>
<td></td>
<td>01</td>
<td>02</td>
<td>03</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BS</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>5</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>DM</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>EA</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JG</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JL</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KM</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LG</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>N V</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>NY</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>PP</td>
<td>6</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RR</td>
<td>9</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WC</td>
<td>6</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YW</td>
<td>11</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total number of clones</td>
<td>41</td>
<td>27</td>
<td>14</td>
<td>46</td>
<td>17</td>
<td>11</td>
<td>31</td>
</tr>
</tbody>
</table>

B. Common Chimpanzee Pan troglodytes (Pt-)

<table>
<thead>
<tr>
<th>Donor</th>
<th>CD94</th>
<th>NKGD2</th>
<th>NKG2A (B)</th>
<th>NKG2B</th>
<th>NKG2C</th>
<th>NKG2E</th>
<th>NKG2F</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>01</td>
<td>02</td>
<td>03</td>
<td></td>
<td>01</td>
<td>02</td>
<td>03</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cathy</td>
<td>3</td>
<td>3</td>
<td>8</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Edwin</td>
<td>6</td>
<td>7</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elsouad</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mason</td>
<td>3</td>
<td>1</td>
<td>6</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total number of clones</td>
<td>11</td>
<td>11</td>
<td>29</td>
<td>14</td>
<td>7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

FIGURE 1. Individuals studied and the CD94 and NKG2 cDNA defined. For each human donor (A) and each common chimpanzee (B) the numbers of cDNA clones characterized for each allele defined are listed. In addition, the number of clones characterized the CD94B splice variant of CD94 and the NKG2B splice variant of NKG2A are shown. All alleles were defined by the sequences of more than one cDNA clone.
alleles for each gene. Numbering starts with the initiating codon nucleotide sequence difference within the set of human and chimpanzee.

**FIGURE 2.** Nucleotide and amino acid differences in human and common chimpanzee CD94 and NKG2 sequences. Listed are the positions of nucleotide sequence difference within the set of human and chimpanzee alleles for each gene. Numbering starts with the initiating codon “ATG” of the cDNA. Each sequence is compared with the left-most human sequence; identities are indicated by a dash and deletions by “del.” In the column headed “Amino Acid Substitutions,” both the residue in the reference sequence and other human and chimpanzee allotypes are indicated where there are differences. Substitution to a termination codon is indicated by an asterisk. Species-specific nucleotides are shaded gray. The exon numbering derives from the following sequences in the GenBank database (AJ000673–AJ001689).

Cryptic 5’ donor splice site within the exon. An alternatively spliced form of human CD94 lacking exon 2 has also been described previously (40). Only variant 1 maintains the proper frame of translation throughout, while variants 2–5 all have a frameshift and premature termination within exon 5.

In summary, CD94 is encoded by a nonpolymorphic gene that is highly conserved in human and chimpanzee, having a nucleotide sequence similarity of 99.6%. Alternatively spliced CD94 variants were found in both species.

**NKG2D**

Three NKG2D alleles were characterized from the analysis of 14 human donors (Fig. 1). Each allele was identified from more than one donor and at least five donors were NKG2D heterozygotes (Fig. 1A); NKG2D01 corresponds to the sequence originally described by Houchins et al. (7). The three NKG2D alleles are very similar, differing by substitutions at only two nucleotide positions, of which one is nonsynonymous and the other synonymous (Figs. 2 and 3). A single Pt-NKG2D sequence was encountered in the analysis of the four chimpanzees, and it differs from the human NKG2D alleles by 7- to 8-nt substitutions which produce 1- to 2-aa differences (Figs. 1B, 2, and 3).

Several alternatively spliced NKG2D variants were characterized, two of which (variants 1 and 2) were found in both human and chimpanzee (Fig. 4). Variants were found that included 106 bp of the 3’ end of intron 3, intron 6, 21 bp at the 3’ end of intron 8, or 23 bp within intron 9, or were missing exon 8. The additional intron sequence in variant 1 is in the 5’ untranslated region (3’ end of intron 3) and should not affect the protein product. The intron 6 inclusion (in variants 2, 3, and 4) results in a stop codon almost immediately after the start of the intron 6 sequence and the encoded proteins therefore lack an extracellular domain. Variant 5 has seven additional amino acid residues in the carbohydrate recognition domain (CRD) due to inclusion in the mRNA 21 bp from the 3’ end of intron 8. The additional 23 bp of intron 9 in variant 6 introduces a stop codon that prevents the translation of exon 10.

In summary, we have defined three very similar human NKG2D alleles and a single chimpanzee Pt-NKG2D allele. The sequences have a nucleotide identity of ~98.9%, demonstrating that NKG2D, like CD94, is a highly conserved gene. A number of alternatively spliced NKG2D variants were identified in both human and chimpanzee.

**NKG2A**

Whereas a single NKG2A sequence was identified from analysis of six human donors, three NKG2A alleles were found in the four chimpanzees (Fig. 1). The human sequence was identical to that previously reported (7, 41). The three Pt-NKG2A alleles differ at four nucleotide substitutions of which two are nonsynonymous and two are synonymous (Figs. 1–3). Human and chimpanzee NKG2A differ by 7–10 nt substitutions in the coding region and the encoded proteins differ by 4–6 amino acid substitutions (Figs. 2 and 3).

Previous analysis indicated that NKG2A and NKG2B are alternatively spliced forms of the same gene, with NKG2A including all exons in the message, whereas NKG2B message lacks exon 4 (Fig. 2).
4 and Refs. 7 and 41). Clones corresponding to NKG2B were obtained from all six human donors analyzed and from one of the four chimpanzees (Fig. 1). An additional splice variant found in chimpanzee Pt-NKG2A03 (variant 1) lacking exons 4 and 6 causes a frameshift and premature termination in exon 7.

In summary, NKG2A is a highly conserved gene with ~98.8% nucleotide sequence identity in human and chimpanzee and share a pattern of alternative mRNA splicing to give the NKG2B form. Human NKG2A appears to have no polymorphism, whereas chimpanzee Pt-NKG2A has modest polymorphism.

NKG2C

Analysis of nine human donors identified two NKG2C alleles, both of which were carried by several individuals, two of whom were shown to be NKG2C heterozygous (Fig. 1A). The NKG2C01 allele has a sequence identical to that deposited by Cantoni et al. (23) into GenBank (accession no. Y13055). NKG2C02 differs from it by two nonsynonymous substitutions at nucleotide positions 5 and 305 (Figs. 2 and 3). The NKG2C sequence described by Houchins et al. (7) (GenBank accession no. NM_002260) was not encountered in our study and differs from NKG2C01 by one nonsynonymous nucleotide substitution.

Six different Pt-NKG2C sequences were seen in the cDNA clones isolated from the four common chimpanzees (Fig. 1B). Four of the sequences encode proteins that are similar in length and sequence to human NKG2C (Pt-NKG2C01–04; Figs. 2 and 3). Of these the Pt-NKG2C01, 02, and 04 proteins have two extra amino acid residues in the carboxyl terminus when compared with the human NKG2C; Pt-NKG2C02 is five residues shorter due to an additional 21-bp deletion in exon 1 that maintains the proper reading frame (Figs. 2 and 3). The other two chimpanzee sequences (Pt-NKG2C0101 and 02) share a nucleotide deletion at position 358 that alters the reading frame near the start of the region encoding the CRD. Starting with position 140 (peptide alignment in Fig. 3) the proteins encoded by Pt-NKG2C0101 and 02 become totally different from those encoded by Pt-NKG2C and NKG2C, and they terminate 35 amino acid residues following the frameshift. Thus the predicted Pt-NKG2CII proteins have cytoplasmic, transmembrane, and extracellular stalk regions like NKG2C, but the CRD is replaced by a shorter region of unknown structure. Pt-NKG2CII terminates at the same position as human and chimpanzee NKG2F, but the two genes have different frameshift mechanisms (Fig. 3).

In comparison to the CD94, NKG2D, and NKG2A genes, NKG2C was found to be most diverged (Fig. 2). Human NKG2C and the chimpanzee alleles that encode similar protein products (Pt-NKG2CII) differ by 22–24 nucleotide substitutions (96.6–96.9% sequence identities) that produce 14–17 amino acid differences (Figs. 3 and 5A). Furthermore, that three different Pt-NKG2C sequences were isolated from two chimpanzees (Edwina and Elwood; Fig. 1B) indicates that there can be two Pt-NKG2C genes in common chimpanzee genomes. In addition, clones corresponding to both Pt-NKG2C and CII sequences were isolated from all four chimpanzees examined. Pairwise comparison of the coding region (Fig. 5A) and the 3’ untranslated region (Fig. 5B) of six Pt-NKG2C sequences clearly shows that they divide into two groups: one comprising the Pt-NKG2CII sequences, the other comprising the Pt-NKG2CII sequences. Comparison between the coding regions of the two common chimpanzee genes yields differences of 11–16 nucleotide substitutions, not including gaps (Fig. 5A). The range of nucleotide comparisons within each group (NKG2C, Pt-NK2C1, and CII) is lower and does not overlap with the range of differences for intergroup comparisons (Fig. 5). That both chimpanzee Pt-NKG2C genes are equally divergent in nucleotide sequence from human NKG2C is consistent with a history in which the Pt-NKG2CII and CII genes arose by gene duplication within the chimpanzee lineage after its split from the human lineage.

Alternatively spliced variants of both Pt-NKG2CII and CII were encountered (Fig. 4). Both Pt-NKG2CII variant 1 and CII variant 1 were missing the same 48- to 49-bp region at the 5′ end of exon 4 likely due to use of a cryptic splice site. Variant 2 from the Pt-NKG2CII02 allele might be generated by cryptic splice sites in exons 2 and 4. Pt-NKG2CII variant 2 was missing exon 5. All of these variants contain frameshifts that cause termination at the same place as Pt-NKG2CII and (Pt-)NKG2F.

Transcription of the Pt-NKG2CII and CII genes in a panel of NK cell clones was examined. Seventeen NK cell clones derived from the PBMC of chimpanzee Cathy (35) were typed by RT-PCR for the presence of Pt-NKG2CII and Pt-NKG2CII mRNA (Fig. 6A). Six of the clones (35%) typed positively for Pt-NKG2CII and three (18%) typed positively for Pt-NKG2CII. One of the NK cell clones typed positively for both Pt-NKG2CII and Pt-NKG2CII. The negative typing reactions were not due to poor quality mRNA, as all the NK cell clones typed positively for expression of Pt-NKG2D. In conclusion, there is differential expression of the two Pt-NKG2C genes in chimpanzee NK cell populations.

In comparison to CD94, NKG2D, and NKG2A, we find NKG2C to be both polymorphic within chimpanzees and humans and more divergent between the species (~3.2–3.3% nucleotide differences). Within-species variation is more striking for chimpanzee than for human, the former species having two distinctive, transcribed NKG2C genes.

NKG2E

Human NKG2E is closely related to NKG2C, but its mRNA splices differentially into an additional exon 7 not used by any other NKG2 gene (Fig. 4). An alternatively spliced NKG2HI form (more similar in its splicing pattern to NKG2C) has also been described (Figs. 3 and 4 and Ref. 42). The oligonucleotide primers we used for amplification were designed to detect the NKG2E form, but not the NKG2H form. Three NKG2E alleles were identified from nine human donors analyzed, with at least two donors being heterozygotes (Fig. 1A). The human NKG2E alleles vary at only two nucleotide positions, of which one is nonsynonymous and the other synonymous (Figs. 2 and 3). NKG2E01 is identical to the GenBank reference sequence NM_002261 (43).

The oligonucleotide primers used to amplify human NKG2E failed to amplify from chimpanzee cDNA. Successful amplification of chimpanzee Pt-NKG2E was eventually achieved with two other primer sets, one that amplified a segment of 552 nt corresponding to about three quarters of the coding region, and another that amplified the entire coding region. Clones corresponding to Pt-NKG2E were obtained from three of the four chimpanzees tested. Most of the clones corresponded to alternatively spliced variants or partial Pt-NKG2E sequences; only one full-length clone was obtained. The divergence between the human and chimpanzee NKG2E nucleotide sequences (~3.4%) is comparable to that seen for NKG2C. The nucleotide sequences of Pt-NKG2E clones analyzed provided evidence for five alleles, distinguished by the patterns of nucleotide substitution at six positions. Four of these positions involve nonsynonymous substitutions and two involve synonymous substitutions. All three chimpanzees from whom Pt-NKG2E sequences were obtained appeared to be heterozygotes; repeated attempts to amplify Pt-NKG2E from chimpanzee Elwood failed.

Two alternatively spliced variants were identified from the human NKG2E (Fig. 4, variants 1 and 2) which were distinct from the eight variants of Pt-NKG2E (variants 3–10), but all contain a
FIGURE 3. Predicted amino acid sequences of human and common chimpanzee CD94 and NKG2 polypeptides. Sequences are compared with the consensus and identities are indicated by dashes. Numbering refers only to the consensus sequence and not to any individual protein. Protein domains are designated according to annotations in GenBank reference sequences (NM_002259–62, 007360, 013431). Gaps were introduced to optimize alignment and are indicated with periods. NKG2H (NM_007333), an alternatively spliced form of the human NKG2E, is also included for comparison. The cytoplasmic, transmembrane, and stalk regions of NKG2D are probably not homologous with their counterparts in other NKG2 molecules, nor is the
change in the normal reading frame resulting in premature termination. Whereas all variants involved “conventional” splicing differences within the exons and introns of Pt-NKG2E, variant 10 (Fig. 4) consists of four exons from Pt-NKG2CI02 (though missing the 5’ third of exon 4), followed by 50 bp identical to a part of the NKG2E promoter (also identical to part of the NKG2C promoter), and then continues with exons 3 through 6 of Pt-NKG2E03 (but missing the 3’ part of exon 4). This suggests that the duplicated Pt-NKG2CII gene is directly upstream of Pt-NKG2E.

Of the NKC genes compared in chimpanzee and human, NKG2E and NKG2C show the highest interspecies divergence (~3.2–3.4% difference in nucleotide sequence) and they are both polymorphic within each species. Alternatively spliced variants of NKG2E were frequently encountered, all of them containing frameshifts that cause premature termination.

NKG2F

Four NKG2F alleles were obtained from the analysis of five human donors and five alleles from the four chimpanzees (Fig. 1), resulting in three NKG2F and five Pt-NKG2F allotypes (Figs. 2 and 3). The human NKG2F01 allele is identical to the original sequence from Plougastel and Trowsdale (44) (GenBank accession no. U96845) and NKG2F02 is identical to the GenBank reference sequence NM_013431 originally described by Glienke et al. (10). At least two humans and three chimpanzees appear to be NKG2F heterozygotes (Fig. 1). Clones corresponding to three NKG2F alleles were obtained from Elwood, the chimpanzee from whom NKG2E could not be amplified. The alleles Pt-NKG2FO4 and 05 differ from the others by a 3-bp deletion in the 5’ untranslated region and a 4-bp deletion in the 3’ untranslated region (data not shown).

Human and chimpanzee NKG2F have ~98.6% sequence similarity, differing by 5- to 9-bp differences between the two species. Striking is that only 2 of the 16 polymorphic positions are species specific (Fig. 2).

Three alternatively spliced variants of Pt-NKG2F01 were encountered, all involving exon 4 and the use of cryptic splice sites. Variant 1 lacked 183 bp within exon 4, but the deletion was to the 3’ side of the NKG2F stop codon. Variant 2 lacked 68 bp at the 5’ end of exon 4, which changes the reading frame to the one used by NKG2A, C, and E in this exon. Variant 3 lacked the same internal exon 4 sequence as variant 1 but also lacked 49 bp from the 5’ end of the exon and is prematurely terminated.

In summary, the NKG2F gene has a pattern of variation that distinguishes it from the other NKC genes examined. It combines relative conservation between the two species with greater polymorphism within each species.

Haplotypic variation in the chimpanzee NKC

Based upon the sequences of the human and chimpanzee CD94 and NKG2A, C, and D genes, a system of PCR sequence-specific primers typing of genomic DNA was developed and used to analyze samples obtained from 11 human donors and 69 apes representing six taxonomic species in total. Typing for the \( \beta_{\text{cm}} \) gene was included as a positive control (Fig. 6B). Typing for NKG2C involved three reactions: one aimed at a generic typing for (Pt-)NKG2C, one targeted at Pt-NKG2CI, and one targeted at Pt-NKG2CII.

All 80 samples typed positively for \( \beta_{\text{cm}} \) and NKG2A, and all but the three gibbons typed positively for CD94. The 11 humans gave identical typing reactions in which the only negatives were for the two Pt-NKG2C genes. In contrast, a heterogeneity of typing pattern was seen within the panel of 48 common chimpanzees. The most frequent pattern, seen in 30 individuals, comprised positive typing for all the genes tested. Six other patterns, all seen in at least two individuals, were distinguished by negative reactions for NKG2C and NKG2D genes. Most divergent from the dominant pattern 1 was pattern 2, in which typing for NKG2D and all three NKG2C reactions was negative (Fig. 6B). This pattern of reactivity was also given by the two gorillas tested.

The 11 pygmy chimpanzees tested gave an identical pattern which distinguished them from all individuals of the other five species. This pattern was comprised of positive reactions for all the genes tested, with the exception of Pt-NKG2CI. Species-specific patterns of typing were also seen for the five orangutans and three gibbons tested. Whereas the gibbons were negative for NKG2C and NKG2D (as well as CD94), the orangutans were negative for NKG2C but positive for NKG2D.

It must be emphasized that negative typing reactions do not show that a gene is absent, merely that at least one of the target sequences used for PCR is not present. Thus the results of the typing provide a measure of divergence from the human and chimpanzee sequences used to design the typing system. Most striking is the phylogenetic conservation of NKG2A and CD94 compared with NKG2C and NKG2D, a result that is consistent with and complementary to the comparison of cDNA sequences in human and common chimpanzee. Also impressive is the variability of the NKC genes within the common chimpanzee. The actual heterogeneity and polymorphism of the chimpanzee NKC is probably much greater than is apparent from this analysis because only individuals who are negative on both haplotypes will give the negative reactions needed to be distinguished from pattern 1.

Discussion

Upon phylogenetic analysis, the six related human NKC genes we studied subdivide into three lineages. CD94 and NKG2D comprise two of the lineages; the third lineage includes NKG2A, NKG2C, NKG2E, and NKG2F (Fig. 7A). Functions for the CD94, NKG2A, NKG2C, and NKG2D polypeptides have been identified, but none yet for NKG2E and NKG2F. Our assessment of polymorphism in the human genes indicates that all of the genes are relatively conserved with allotypes differing by no more than two amino acid substitutions (Figs. 2 and 3).

No polymorphism in human CD94 and NKG2A was found, indicating that the heterodimeric, inhibitory receptor formed by these two polypeptides with specificity for HLA-E (13–19, 21, 22) is largely conserved within the human population. CD94 also forms an activating receptor for HLA-E by association with NKG2C (22, 23). We identified two NKG2C allotypes differing at two amino acid residues. Both allotypes were well represented in the sample population, demonstrating modest polymorphism in this receptor, perhaps at a level comparable to that of the oligomorphic HLA-E

short cytoplasmic tail of CD94. Different translational frames are also shown for the carboxyl-terminal ends of Pt-NKG2CII, (Pt-)NKG2E, and NKG2H. Elements of secondary structure in the CRD of CD94 were based on the structure determined by Boyington et al. (65): loops are designated by L plus a number, \( \beta \) strands by \( \beta \) plus a number, and an \( \alpha \) helix as \( \alpha \). Shaded positions show the differences between the alternatively spliced forms of the CD94 gene product (CD94 and CD94B), interspecies variations, and allelic polymorphisms in NKG2 loci. A full-length cDNA clone was not obtained for Pt-NKG2E05 and the sequence shown is for a splice variant lacking exon 4. Boxed are the immunoreceptor tyrosine-based inhibitory motifs (V/IxYxxL) in the cytoplasmic region and positively charged residues in the transmembrane region.
FIGURE 4. Alternatively spliced variants of CD94 and NKG2 cDNAs. A. Alternatively spliced variants identified in this study. Genomic structures of the common chimpanzee genes are predicted from those of the human genes. Exons are depicted as numbered boxes and introns are not drawn to scale. Exon boxes are shaded to reflect homology between genes, with filled boxes indicating high homology to NKG2A and gray boxes indicating low or no homology. Open boxes represent untranslated regions of cDNA and dashed boxes represent pseudo exons in human genes. Below the gene structures are representations of the various cDNAs isolated from this study; also included is the human NKG2H coding cDNA (42). B, A schematic showing how the CD94 and CD94B variants are generated using two alternative “AG” acceptor splice sites at the 3’ end of intron 4. As a consequence, CD94B has an additional residue, a glutamine (Q), which is not present in CD94.
Further evidence for species-specific evolution within the NKC is seen in a tree made from just the primate NKG2A, C, E, F, and related sequences (Fig. 7B). This tree contains three distinguishable groups of sequences, corresponding to the NKG2A sequences, the NKG2F sequences, and the combination of NKG2C and NKG2E sequences. The two rhesus monkey (Macaca mulatta) sequences in the NKG2A group (48) are considerably diverged, both from human and chimpanzee NKG2A and from themselves (11 amino acid differences). Either there is a polymorphism greater than is seen in human and chimpanzee at a single Mm-NKG2A locus or there are two Mm-NKG2A genes. It is also intriguing that the rhesus monkey MHC-E is oligomorphic or polymorphic (59). Similarly, within the NKG2F group there are two divergent rhesus monkey NKG2F-like sequences (48) of which one is closer to human and chimpanzee NKG2F. Within the NKG2C group, the eight Mm-NKG2C sequences (48) form a species-specific subgroup that does not interleaf with the human and chimpanzee NKG2C sequences. All these data point to the NKC of the rhesus monkey having significant differences in gene content and haplotypic variation from the NKCs of human and chimpanzee.

Although the CD94 and NKG2 genes of the NKC are generally older, less variable, and slower evolving than the KIR gene family of the LRC (Fig. 8), phylogenetic comparisons suggest that they have nontrivially diverged during the course of mammalian evolution. That is clearly the case for the ligands recognized by these lectin-like receptors. Thus the Qa-1 class I ligand for mouse CD94: NKG2A and CD94:KIR receptors does not in its sequence have particular affinity with the HLA-E class I ligand for human CD94:KIR2DL1 and CD94:KIR2DS1 receptors (52, 53, 60, 61). Even more different are the divergent MHC class I-related ligands for NKG2D, which comprise MIC (24, 25) and ULBPs (47) in humans, and the retinoic acid early inducible RAE-1 and H60 in mice (62, 63).

Comparison of the human and chimpanzee NKC genes indicates that they have evolved under different selection pressures: CD94 is most conserved, followed by NKG2D, NKG2A, and NKG2F, with NKG2C and NKG2E being most divergent (Fig. 8) (35). The human and common chimpanzee CD94 have only 0.4% nucleotide difference, well below the average /H110111.24% variation as suggested for intergenic regions of human and chimpanzee genes (64). NKG2A and NKG2D also appear to be under purifying selection.

![FIGURE 5](https://www.jimmunol.org/)

Common chimpanzee Pr-NKG2CI and CII are equally diverged from human NKG2C. Numbers of nucleotide differences between sequence pairs are shown above the diagonal, while percentages of nucleotide differences are shown below. In parentheses are the maximum lengths of the regions compared; insertions and deletions were excluded from analysis. In the comparison in A, sequences downstream of the premature termination codon of Pr-NKG2CII alleles, corresponding to the coding regions of NKG2C and Pr-NKG2CI, were included. In the comparison of the 3’ untranslated region (B), Pr-NKG2CI02 was not included because only 190 bp of 3’ untranslated region was available due to the positioning of the specific oligonucleotide primer used to amplify this allele. The human sequences determined were similarly truncated in the 3’ untranslated region, so the NKG2C01 sequence used is from a previously reported genomic sequence (GenBank accession no. AJ001684).

![FIGURE 6](https://www.jimmunol.org/)

Typing for the expression and distribution of NKC genes. A. The results of RT-PCR typing for the transcription of Pr-NKG2CI and Pr-NKG2CII genes in NK cell clonal isolates from the common chimpanzee Cathy (35); typing for NKG2D transcription was included as control. B. Genomic DNA typing of different ape species for NKG2A, C, D, and CD94 genes; typing for /H9252 m was included as positive control. In parentheses are the numbers of individuals included for analysis. Only 7 of the 11 pygmy chimpanzee were typed for NKG2D. The percentage of positive typing results within each group is shown.
For CD94 and NKG2A, the identical inhibitory MHC class I specificity of the human and chimpanzee CD94:NKG2A receptors supports this mode of selection (35). The human and common chimpanzee NKG2C and NKG2E genes differ by 3.2–3.4% in nucleotide sequences, values higher than expected for the divergence of human and chimpanzee genes under neutral selection, and similar to those KIR genes which are orthologous in chimpanzee and human (Fig. 8). Most nucleotide substitutions that distinguish NKG2C and Pt-NKG2C1 change residues in the CRD loops (Figs. 2 and 3) predicted to interact with ligand (65). These differences might affect the ligand-binding specificity and be a consequence of selection upon activating receptors by pathogens. Such selection might also be the cause of the greater haplotypic diversity of NKG2C compared with other NKC genes, as seen both within species (as in the common chimpanzee) and between species.

The rodent NKC contains a diverse family of Ly49 genes that encode lectin-like receptors specific for polymorphic determinants of classical MHC class I molecules (66). In humans the Ly49 family is represented by a single nonfunctional gene, Ly49L (67), while KIR provide functions analogous to those of Ly49 in mice. In our investigation we did not examine the structure and diversity of the Ly49L gene. However, Mager et al. (68) have recently shown that a gene related to human Ly49L is present in common chimpanzee as well as in other primates (gorilla, orangutan, gibbon, baboon, and African green monkey) and other mammalian orders. In contrast, genes related to the expanded rodent Ly49 family were not detectable by Southern blot analysis in non-rodent species. Partial sequence analysis showed that common chimpanzee Ly49L has the same inactivating mutations as the human gene, whereas that was not so for the other primates. The extent to which these differences are fixed in humans, chimpanzees, and other species was not addressed, as analysis was confined to individuals of a species. For baboon, a complete cDNA sequence for Ly49L was determined (~95% nucleotide identity with the human Ly49L in the coding region) and it was shown to be expressed in baboon lymphocytes along with transcripts from several KIR genes. These observations raise the possibility that Ly49L encodes a functional receptor in some primate species and that its inactivation in humans and chimpanzee may be exceptional. Also worthy of consideration is that inactivation of the Ly49L gene may not be a property of all human and chimpanzee NKC haplotypes.
FIGURE 8. Polymorphism and variation among human and common chimpanzee CD94, NKG2, and KIR. The within-species and between-species variation in coding-region nucleotide sequences are shown as percentage differences. Within a species, the average pairwise differences between alleles are shown. A dash indicates that only one allele has been described. Between-species variation is the average of all possible pairwise comparisons between alleles of the two species. Orthologous genes are on the same line. The majority of human and chimpanzee KIR lineage III genes are not orthologous (35); they were considered together and compared only as species-specific groups (shaded). The exon sequences encoding the D0 domain in chimpanzee KIR3D of this group were not included in the interspecies comparisons because their human counterpart lack this domain (35).

Acknowledgments

We thank the staff at Yerkes Regional Primate Research Center of Emory University (especially Rickie Bass) and the Laboratory for Experimental Medicine and Surgery in Primates for assistance in obtaining chimpanzee tissue samples. We also thank Dr. Erin Adams for generating the chimpanzee B cell lines and Drs. Stewart Cooper and Ann Erickson for the RNA sample from the chimpanzee Todd.

References


44. Kurepa, Z., C. A. Hasemann, and J. Forman. 1999. Qa-1b molecules, bind to CMV glycoprotein UL16 and stimulate NK cytotoxicity through the Qa-1(b) molecule Qa-1b by putative activating receptors CD94/NKG2C and CD94/NKG2E on mouse natural killer cells. J. Exp. Med. 190:1381.


