Molecular Cloning and Characterization of a Novel CD1 Gene from the Pig

Taehoon Chun, Kai Wang, Federico A. Zuckermann and H. Rex Gaskins

*J Immunol* 1999; 162:6562-6571; ;
http://www.jimmunol.org/content/162/11/6562

**References**

This article *cites 78 articles*, 29 of which you can access for free at:
http://www.jimmunol.org/content/162/11/6562.full#ref-list-1

---

**Why The JI?** Submit online.

- **Rapid Reviews!** 30 days* from submission to initial decision
- **No Triage!** Every submission reviewed by practicing scientists
- **Fast Publication!** 4 weeks from acceptance to publication

---

**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
Molecular Cloning and Characterization of a Novel CD1 Gene from the Pig

Taehoon Chun, Kai Wang, Federico A. Zuckermann, and H. Rex Gaskins

Much effort is underway to define the immunological functions of the CD1 multigene family, which encodes a separate lineage of Ag presentation molecules capable of presenting lipid and glycolipid Ags. To identify porcine CD1 homologues, a cosmide library was constructed and screened with a degenerate CD1 α3 domain probe. One porcine CD1 gene (pCD1.1) was isolated and fully characterized. The pCD1.1 gene is organized similarly to MHC class I and other CD1 genes and contains an open reading frame of 1020 bp encoding 339 amino acids. Expression of pCD1.1 mRNA was observed in CD3+ thymocytes, B lymphocytes, and tissue macrophages and dendritic cells. The pCD1.1 cDNA was transfected into Chinese hamster ovary cells, and subsequent FACS analysis demonstrated that mAb 76-7-4, previously suggested to be a pig CD1 mAb, recognizes cell surface pCD1.1. Structurally, the pCD1.1 α1 and α2 domains are relatively dissimilar to those of other CD1 molecules, whereas the α3 domain is conserved. Overall, pCD1.1 bears the highest similarity with human CD1a, and the ectodomain sequences characteristically encode a hydrophobic Ag-binding pocket. Distinct from other CD1 molecules, pCD1.1 contains a putative serine phosphorylation motif similar to that found in human, pig, and mouse MHC class Ia molecules and to that found in rodent, but not human, MHC class I-related (MR1) cytoplasmic tail sequences. Thus, pCD1.1 encodes a molecule with a conventional CD1 ectodomain and an MHC class I-like cytoplasmic tail. The unique features of pCD1.1 provoke intriguing questions about the immunologic functions of CD1 and the evolution of Ag presentation gene families.


Departments of *Animal Sciences and 1Veterinary Pathobiology, University of Illinois at Urbana-Champaign, Urbana, IL 61801; and 2Department of Molecular Biotechnology, University of Washington, Seattle, WA 98195

Received for publication November 5, 1998. Accepted for publication March 8, 1999.

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

1 This work was supported by U.S. Department of Agriculture Section 1433 Animal Health and Disease Grant ILLU-35-0911.

2 Genomic and cDNA sequence data reported in this paper were submitted to GenBank nucleotide sequence database and assigned the numbers AF056045 (genomic DNA) and AF059492 (cDNA), respectively.

3 Current address: Gwen Knapp Center for Lupus and Immunology Research, Department of Pathology, University of Chicago, Chicago, IL 60637.

4 Current address: Chiroscience R&D, Bothell, WA 98021.

5 Address correspondence and reprint requests to Dr. H. R. Gaskins, University of Illinois at Urbana-Champaign, 1207 W. Gregory Dr., Urbana, IL 61801. E-mail address: hrgaskins@uiuc.edu

6 Abbreviations used in this paper: β2m, β2-microglobulin; DC, dendritic cells; mCD1, mouse CD1; pCD1.1, porcine CD1.1; 5′UT, 5′ untranslated region; 3′UT, 3′ untranslated region; SLA, swine histocompatibility Ag; PEC-A, porcine endothelial cell-A; CHB, Chinese hamster ovary; CDPSS, cysteine-serine-serine-proline-serine-serine; DN, double negative.

Copyright © 1999 by The American Association of Immunologists 0022-1767/99/$02.00
genes are found in mice (25), rats (26), rabbits (23), and humans (2, 12). Mice and rats lack group I CD1 genes (25, 26), and the sheep genome encodes four copies of a CD1b gene homologue (24). Thus, mammalian species may harbor unique collections of CD1 genes, perhaps reflecting selective pressure after speciation.

Pigs constitute a major food-animal group, are used as a large animal model in biomedical research due to anatomical and physiological similarities with humans, and are potentially an important source of tissues or organs for xenotransplantation (27, 28). In addition, unique aspects of T cell biology in the pig make this species particularly suitable for studying the generation of T cell subset diversity and distribution (29–33). However, little is known about porcine non-MHC Ag presentation molecules. This knowledge is needed to better understand regulatory aspects of pig immunity, xenotransplantation issues, and comparative aspects of Ag presentation. We describe the molecular cloning and characterization of a porcine CD1 gene (pCD1.1) that has a conventional CD1 ectodomain with a class I-like cytoplasmic tail that appears to be a new member of the CD1 family.

Materials and Methods

CD1 probe

For porcine cosmid library screening, a CD1 probe was generated from PCR using degenerate primers based on human (2, 12) and mouse CD1 (25) gene sequences. PCR was performed with a PTC-100 thermocycler (MJ Research, Watertown, MA) for 35 cycles of 1 min at 94°C, 2 min at 55°C, and 2 min at 72°C, followed by a 10-min final extension step at 72°C using 200 ng of porcine genomic DNA. A 194-bp PCR product covering exon 4 (α3 domain) was generated by amplification with the following primers: 5′-GCTGGTATGTCATGTCTCTG (sense) and 5′-GACTGCTTACCCCACTTCT (antisense). The CD1 α3 domain PCR product (data not shown) was subcloned into the pGEM-T cloning vector (Promega, Madison, WI) and sequenced by the dyeoxy-mediated chain termination method (34) using a 373 automatic sequencer (ABI, Foster City, CA).

Porcine cosmid library construction and screening

Genomic DNA was isolated from PBMCs of an adult female Duroc pig using standard techniques (35). A five-genome-equivalent porcine cosmid library was constructed from partially HI-digested genomic DNA from an adult female Duroc pig using Ficoll-Hypaque (density, 1.077 g/ml; Sigma, St. Louis, MO) centrifugation at 900 × g for 30 min. Mononuclear cells at the gradient interface were collected by aspiration and washed twice in RPMI 1640 medium. Macrophages from peripheral blood and spleen were separated from lymphocytes according to differential plastic adherence (38).

Splenic DC were prepared as described previously (39, 40). Briefly, spleen cells were digested with collagenase and DNase in RPMI 1640 medium containing 2 mg/ml collagenase type IV (Sigma) for 15 min at 37°C. Digested fragments were filtered and resuspended in a dense BSA (1.08 g/ml; Sigma) solution, overlaid with RPMI 1640 medium and centrifuged at 9500 × g for 15 min. The low density cell population was resuspended in complete RPMI 1640 medium containing 10% FCS in a 60-mm tissue culture dish and was maintained at 37°C in a humidified 5% CO2 incubator. Nonadherent cells were removed by gentle washing with warm RPMI medium after a 2-h incubation, and the adherent cell population, containing DC and macrophages, was cultured overnight in complete RPMI 1640 medium containing 10% FCS. After overnight culture, detached DC were collected from the medium, whereas most macrophages remained adherent on tissue culture dishes (39, 40).

Epithelial cells were prepared from large intestine according to a previously described procedure (41). The large intestine was cut into small fragments (1 cm) and washed several times in calcium- and magnesium-free HBSS containing 1% penicillin-streptomycin, 1% fungizone, 50 μg/ml gentamicin, and 10 mM HEPES (Sigma). The tissue was then treated with 1 mM DTT (Sigma) in RPMI 1640 containing 5% FCS for 2 min. After washing, tissue segments were placed in calcium- and magnesium-free HBSS containing 0.75 mM EDTA (Sigma) and were maintained on an orbital shaker (100 rpm) at 37°C for 45 min. Pig macrophages, DC, lymphocytes, and epithelial cells were distinguished phenotypically by FACS analysis with purified mAbs against pig leukocyte Ags as described below (data not shown). Splenic DC, macrophages, lymphocytes, and peripheral blood leukocytes were positive for CD45, the common hemopoietic cell marker. Lymphocytes were negative for CD16 (FcRIII) and exhibited various levels of MHC class II (swine histocompatibility leukocyte Ag (SLA)-DR) expression. Splenic DC were negative for CD16 (FcRIII) and exhibited a high level of MHC class II (SLA-DR) expression, whereas macrophages were positive for CD16 (FcRIII) and displayed various levels of MHC class II (SLA-DR) expression. Epithelial cells from the large intestine (colon) showed a negative phenotype for all three cell surface markers.

The pig endothelial cell-A (PEC-A) line was established originally from aortic endothelial cells harboring a hybrid Rous sarcoma virus long terminal repeat promoter/SV40 large T Ag gene construct (provided by F. J. Candal, Centers for Disease Control, Atlanta, GA) (42). PEC-A cells express low levels of SLA class I constitutively, but do not express SLA class II (42). PEC-A cells were maintained in DMEM containing a final glucose concentration of 25 mM and were supplemented with Eagle’s nonessential amino acids, 44 mM sodium bicarbonate, 15 mM HEPES, 50 μg/ml gentamicin sulfate, and 10% FCS. Unless noted otherwise, cell culture reagents were purchased from Life Technologies.

isolation and culture of cells

Porcine mononuclear cells from peripheral blood, spleen, and thymus were isolated from an adult female Duroc pig using Ficoll-Hypaque (density, 1.077 g/ml; Sigma, St. Louis, MO) centrifugation at 900 × g for 30 min. Mononuclear cells at the gradient interface were collected by aspiration and washed twice in RPMI 1640 medium. Macrophages from peripheral blood and spleen were separated from lymphocytes according to differential plastic adherence (38).

The complete cDNA encoding pCD1.1 was generated by RT-PCR. Briefly, total RNA (1 μg) from adult female Duroc pig thymus was isolated according to a guanidium extraction method (37) and reverse transcribed into cDNA using 50 μl of murine Moloney leukemia virus-derived reverse transcriptase, 1 mM of each dNTP, 2.5 mM oligo(dT) primers, and 5 μM MgCl2. After cDNA synthesis, PCR was performed with the specific primers (25 pmol) listed below for 35 cycles of the temperature-time sequence: 10 s at 94°C, 30 s at 55°C, and 30 s at 72°C (Perkin-Elmer GeneAmp RNA PCR kit reagents (Perkin-Elmer/Cetus, Norwalk, CT) were used for RT-PCR. Amplified cDNA products were subcloned into the pGEM-T cloning vector (Promega) and were sequenced by the dyeoxy-mediated chain termination method using a 373 automatic sequencer (ABI). Primers (5′-AATGCTTGTTCGCAACTCCCAT, sense, 5′-translated region (5′UT); 5′-CCCTAGGACATCATGCCATAGACAG, antisense, 5′-translated region (5′UT) for cDNA amplification were designed from the genomic pCD1.1 DNA sequence and produced a 1236-bp PCR product (GenBank accession no. AF059492; data not shown).

Restriction mapping and genomic DNA sequencing

To characterize the putative CD1 gene, a partial restriction map of cosmid clone CWG-34 was generated from Southern blot analysis using the CD1 probe and BamHI digestion, separately and combined) with the CD1 α3 domain probe.

Sequence generation

The complete cDNA encoding pCD1.1 was generated by RT-PCR. Briefly, total RNA (1 μg) from adult female Duroc pig thymus was isolated according to a guanidium extraction method (37) and reverse transcribed into cDNA using 50 μl of murine Moloney leukemia virus-derived reverse transcriptase, 1 mM of each dNTP, 2.5 mM oligo(dT) primers, and 5 μM MgCl2. After cDNA synthesis, PCR was performed with the specific primers (25 pmol) listed below for 35 cycles of the temperature-time sequence: 10 s at 94°C, 30 s at 55°C, and 30 s at 72°C (Perkin-Elmer GeneAmp RNA PCR kit reagents (Perkin-Elmer/Cetus, Norwalk, CT) were used for RT-PCR. Amplified cDNA products were subcloned into the pGEM-T cloning vector (Promega) and were sequenced by the dyeoxy-mediated chain termination method using a 373 automatic sequencer (ABI). Primers (5′-AATGCTTGTTCGCAACTCCCAT, sense, 5′-translated region (5′UT); 5′-CCCTAGGACATCATGCCATAGACAG, antisense, 5′-translated region (5′UT) for cDNA amplification were designed from the genomic pCD1.1 DNA sequence and produced a 1236-bp PCR product (GenBank accession no. AF059492; data not shown).
mAbs and FACS analysis

Murine mAbs against pig leukocyte Ags have been characterized previously (43). These include BB23-8E6 (anti-CD3), K252.7e4 (anti-CD45), G7 (anti-CD16), MS3A-1 (anti-CD4), and anti-porcine Ig (H+L chains, B cell marker; South Biotechnology Associates, Birmingham, AL). Sources of these mAbs are as follows: BB23-8E6, M. D. Pescevitz (University of Indiana, Indianapolis, IN); K252.7e4, C. Stokes (Bristol, U.K.); G7, Y.-B. Kim (Chicago Medical School, Chicago, IL); and MS3A-1, J. K. Lunney (U.S. Department of Agriculture, Beltsville, MD). For FACS analysis, 10^5 cells of each population (macrophages, DC, lymphocytes, or epithelial cells) were incubated on ice in 50 μl of FACS buffer (PBS/1% BSA) with saturating concentrations of the indicated mAbs at 4°C for 1 h. Cells were washed twice and stained with 10 μg/ml of either PE- or FITC-conjugated sheep anti-mouse IgG in 50 μl of FACS buffer at 4°C for 30 min. After a final wash, cells were resuspended in 1 ml of FACS buffer. Fluorescence intensity was quantified using a Coulter XL flow cytometer (Hialeah, FL) equipped with an argon ion laser. Data were analyzed with Elite software from Coulter. Peripheral T lymphocytes and B lymphocytes, respectively, were isolated based on their specificities with the relevant mAbs by FACS. The CD3^+ and Ig^+ cells from PBMCs were classified as peripheral T lymphocytes and B lymphocytes, respectively. Mononuclear cells from peripheral blood or thymus were subsequently sorted after labeling with the appropriate mAbs at a rate of 3000 events/s on an EPICS 752 flow cytometer (Coulter).

RT-PCR analysis

Total RNA (1 μg) was isolated from each tissue and cell type described above using a guanidium extraction method (37) and was reverse transcribed into cDNA using 50 U of murine leukemia virus-derived reverse transcriptase, 1 mM of each dNTP, 2.5 mM oligo(dT) primers, and 5 mM MgCl_2. After cDNA synthesis, PCR was performed with the specific primers (25 pmol) listed below for 35 cycles of the following sequence: 1) 95°C for 30 s, 2) 60°C for 30 s, and 3) 72°C for 1 min. Specific primer sequences were as follows: pCD1.1, 5'-CGTCTGTTCCCTTTCTTTATGACG (sense) and 5'-CCCTAGGACATCATGCCATAGACAG (antisense), resulting in a 678-bp PCR product; and GAPDH, 5'-CCCTAGGACATCATGCCATAGACAG (antisense), resulting in a 225-bp product (45). RT-PCR products were confirmed by restriction enzyme digestion and DNA sequence analysis (data not shown).

Generation of pCD1.1-transfected cells

The complete pCD1.1 cDNA was ligated into the mammalian expression vector pTARGET (Promega) according to the manufacturer's protocol. The correct orientation of the cloned pCD1.1 cDNA was confirmed by restriction enzyme digestion and DNA sequence analysis (data not shown). The expression vector containing pCD1.1 cDNA was transfected into Chinese hamster ovary (CHO) cells using the Lipofectin reagent (Life Technologies). The pCD1.1-transfected CHO cells were selected in DMEM containing a final glucose concentration of 25 mM and supplemented with Eagle's nonessential amino acids, 44 mM sodium bicarbonate, 15 mM HEPES, 50 μg/ml gentamicin sulfate, 10% FCS, and 500 μg/ml genetin (G-418, Life Technologies). Twenty days after selection, pCD1.1-expressing CHO cells were screened by RT-PCR analysis and FACS analysis with mAbs (44). The correct orientation of the cloned transcript was confirmed by restriction enzyme digestion and DNA sequence analysis (data not shown). The carboxyl end comprising the cytoplasmic tail is a stretch of 14 hydrophilic amino acids.

Comparison of pCD1.1 nucleotide and amino acid sequences with other CD1 molecules

The percent similarities of pCD1.1 cDNA and deduced amino acid sequences compared with CD1 molecules of other species are shown in Table I. The pCD1.1 cDNA sequence shows the highest overall similarity with human CD1a, being 74% homologous at the nucleotide level and 63% homologous with the deduced amino acid sequence (Table I). However, the α1 and α2 domains of pCD1.1 exhibit marked amino acid dissimilarity and are most similar to either human CD1c (41% homology; α1 domain) or human CD1a (62% homology; α2 domain). The α3 domain of pCD1.1 shows a relatively high amino acid similarity (>75%) with other CD1 α3 domains. Alignment of the pCD1.1 amino acid sequence with other CD1 molecules revealed additional conserved and non-conserved residues (Fig. 2). Four putative N-linked glycosylation sites, predicted by the N-X-T/S (X: any amino acid) consensus sequence, were found at amino acid positions 41, 78, 149, and 217 of pCD1.1, with amino acid positions 41, 78, and 149 being well conserved among group 1 CD1 molecules (Fig. 2). The putative N-linked glycosylation site at amino acid position 217 in the α3 domain is unique to pCD1.1 (Fig. 2). Only human CD1b also has an N-linked glycosylation site in the α3 domain, found at amino acid position 261 (Fig. 2). Cysteine residues forming a potential intradomain disulfide bridge are well conserved among CD1 molecules and present at aa positions 123 and 187 in the pCD1.1 α2 domain and at aa positions 227 and 282 in the α3 domain (Fig. 2).
Additional cysteines found at aa position 49 in the pCD1.1 α1 domain and at aa position 166 in the α2 domain may also form interdomain disulfide bridges (Fig. 2).

**Structural comparison of pCD1.1 with other CD1 and MHC class I molecules**

The crystallographic structure of the mCD1.1 molecule has provided a structural basis for a CD1 Ag-binding pocket (52). Similar to MHC class Ia molecules, the overall structure of the mCD1.1 Ag binding motif is composed of an eight-stranded antiparallel β pleated sheet with two long antiparallel α-helical structures that sit atop and traverse the β sheet platform (52). However, the Ag binding motif of mCD1.1 creates two large pockets, A' and F', instead of six pockets (A–F) as found within MHC class Ia molecules (52). Further, hydrophobic amino acids within each wall of the mCD1.1 A' and F' pockets provide a microenvironment that favors binding of hydrophobic Ags (52). A comparison of mCD1.1 and the deduced amino acid sequence of pCD1.1 revealed conserved hydrophobic residues within each wall of the putative A' and F' pockets of pCD1.1 (not shown). Within the F' pocket of mCD1.1 and pCD1.1, all amino acid residues are hydrophobic, except one residue at the same position for both species (N94 in pCD1.1 and D94 in mCD1.1). Comparing A' and F' pocket amino acid residues of pCD1.1 with those of SLA class I, all eight consensus residues within SLA class I (PD1 and PD14) (53) are replaced with hydrophobic residues in pCD1.1, similar to the observations with Kβ and mCD1.1 (52).

An interaction between CD1 and β2m is required for stable expression of CD1 on the cell surface (54). However, the number of contact sites between β2m and mCD1 is less than that for β2m and MHC class Ia molecules (53). A comparison of amino acid residues involved in MHC class Ia and β2m interactions (55) with equivalent residues in pCD1.1 indicated that pCD1.1:β2m assembly may also differ from SLA class I:β2m assembly (not shown).

The most unique feature of pCD1.1 was found in the cytoplasmic tail (Fig. 3). The pCD1.1 cytoplasmic tail consists of 14 amino acids and differs in sequence from those of any other CD1 molecule characterized to date (Fig. 3A). Accordingly, the pCD1.1 cytoplasmic tail does not contain the tyrosine-based YXXZ (X, any amino acid; Z, bulky hydrophobic amino acid) motif found in most CD1 proteins that has been shown to target human CD1b to endocytic compartments (56, 57). Instead, the pCD1.1 cytoplasmic tail includes a 5-aa sequence, cysteine-aspartic acid-proline-serine-serine (CDPSS), with striking resemblance to a serine-containing motif conserved in the cytoplasmic tails of human and mouse MHC class Ia molecules (Fig. 3B; SD/EXSL; X, any amino acid) (58). The latter serine residue (position 335) in the SD/EXSL motif is a site of phosphorylation in HLA-B7 molecules (58) and is required for constitutive endocytosis of HLA class I molecules (59).

**Analysis of pCD1.1 mRNA expression**

To investigate the pattern of pCD1.1 mRNA expression, RT-PCR analysis was conducted with a variety of adult pig tissues using gene-specific primers based on the cDNA sequence of pCD1.1 (see Materials and Methods). As shown in Fig. 4A, pCD1.1 mRNA was detected by RT-PCR in a broad range of tissues, including lymphoid tissues, such as spleen, thymus, and small and large intestine, as well as nonlymphoid tissues, including kidney, liver, lung, and skin. The highest level of pCD1.1 mRNA expression was detected in the thymus (Fig. 4A). The lowest level of pCD1.1 mRNA expression was observed in the heart (Fig. 4). To examine cell type-specific patterns of pCD1.1 mRNA expression, thymocytes, macrophages, DC, lymphocytes, and epithelial cells were isolated from peripheral blood, thymus, spleen, or large intestine (colon). Results from RT-PCR analysis of these individual cell types indicate that pCD1.1 mRNA is expressed by CD3+ thymocytes and professional APCs such as DC, macrophages, and B lymphocytes in the periphery, but is not expressed by endothelial or epithelial cells (Fig. 4, B–F).

**FACS analysis of pCD1.1+ CHO cells**

The mAb 76-7-4 (mouse IgG2a) was derived from a panel of mAb isolated from mice immunized with porcine thymocytes (44) and
recognizes a 40-kDa molecule associated with β₂m on the surface of immature thymocytes, peripheral B lymphocytes, macrophages, and Langerhans cells, but not peripheral T lymphocytes (44, 60, 61). Therefore, this mAb shares characteristics with anti-human CD1 Abs. The pattern of pCD1.1 mRNA expression on CD3⁺ thymocytes and macrophages, and DC and B lymphocytes in the periphery is consistent with that of cell types stained by mAb 76-7-4. To determine whether mAb 76-7-4 recognizes pCD1.1 on the cell surface, pig thymocytes were isolated based on mAb 76-7-4 specificity (76-7-4₁ vs 76-7-4₂) and were analyzed by RT-PCR for pCD1.1 mRNA expression. Results from RT-PCR analysis correlated with thymocyte staining results using mAb 76-7-4 (Fig. 5A).

FIGURE 2. Alignment of deduced amino acid sequence of pCD1.1 ectodomains (aa 25–297) with those of other CD1 molecules. Amino acid substitutions, gaps (.), and identities (-) are shown. Three CD1 groups (a, group 1 CD1; b, CD1d; and c, CD1e) are compared. Putative N-linked glycosylation sites are underlined.
did not express pCD1.1 mRNA and were not recognized by mAb 76-7-4. To confirm that mAb 76-7-4 recognizes the pCD1.1 protein, the complete pCD1.1 cDNA was cloned into an expression vector and transfected into CHO cells. Twenty days after selection with geneticin, pCD1.1 CHO cells were screened by RT-PCR (Fig. 5B). Subsequent FACS analysis demonstrated that mAb 76-7-4 recognizes cell surface pCD1.1 on pCD1.1 CHO cells (Fig. 5B).

**Phylogenetic analysis**

Phylogenetic analysis was conducted to consider evolutionary relationships among CD1 molecules using the neighbor-joining method of Saitou and Nei (48). Three different regions, including the complete coding region, the ectodomains (α1–α3 domains), and the Ag-binding α1 and α2 domains, were compared among mammalian CD1 molecules. Results from each phylogenetic analysis were identical, and one phylogenetic tree, resulting from comparison of the complete coding region of each CD1 molecule, is shown in Fig. 6. Porcine CD1.1 clusters with human CD1a, while human CD1c clusters with CD1b molecules, and human CD1e clusters with CD1d (Fig. 6). Thus, three clades and six subclades were revealed among known mammalian CD1 molecules (Fig. 6). The branching order in each clade is consistent with the possibility that pCD1.1 diverged from human CD1a after CD1c diverged from CD1b and before CD1e diverged from CD1d (Fig. 6). The
duplicated mouse CD1d (mCD1.1 and mCD1.2) molecules appear to have emerged most recently (Fig. 6). These observations are supported by high bootstrap values (>70% bootstrap reliability; Fig. 6) and indicate that pCD1.1 represents a distinct branch on the CD1 phylogenetic tree.

Discussion

The novel pig CD1 gene described offers clues for a better understanding of the immunologic functions of CD1 and the evolution of Ag presentation gene families. The genomic organization of pCD1.1 is typical of that of other CD1 and MHC class I genes, with a 5′ leader sequence followed by three extracellular α domains encoded by separate exons and followed by transmembrane, cytoplasmic, and 3′-untranslated sequences. The pCD1.1-coding region bears the highest similarity with human CD1a, exhibiting 74 and 63% homology with the CD1a nucleotide and amino acid sequences, respectively. Among extant CD1 molecules, the α3 domain is more conserved than are the α1 and α2 domains. Further, α1 and α2 domains are much less conserved among CD1 family members than are the corresponding domains among class I MHC molecules (1, 4). Consistent with this pattern, the α1 and α2 domains of pCD1.1 are dissimilar to other CD1 molecules, whereas the α3 domain is relatively conserved. Nonetheless, the pCD1.1 gene, similar to other CD1 genes, encodes a hydrophobic Ag binding pocket and thereby probably contributes to host presentation of lipid or glycolipid Ags.

Notably different from other mammalian CD1 molecules, the pCD1.1 cytoplasmic tail consists of 14 aa and includes a 5-aa sequence (CDPSS) that is similar to a serine-containing motif (SD/EXSL; X, any amino acid) conserved in the cytoplasmic tails of human and mouse MHC class Ia molecules (58). Cytoplasmic tails of CD1b, CD1c, and CD1d consist of 7–9 aa and contain the well-conserved amino acid sequence YXXØ (Y, tyrosine; X, any amino acid; Ø, hydrophobic amino acid) (56, 57). This tyrosine-containing motif provides a signal for localization of CD1b to various acidic endosomal compartments, including lipid-rich late endosomes in which MHC class II molecules are loaded with peptides (20, 56, 57). The cytoplasmic tail of human CD1a consists of only 3 aa (CFC) and lacks a conserved motif or possible phosphorylation site as found in other CD1 and MHC class Ia molecules (1, 2, 4, 57). Thus, the pCD1.1 cytoplasmic tail provides the third example of notable dissimilarity in the cytoplasmic tail region of a CD1 gene.

The functional role of the pCD1.1 cytoplasmic tail is unknown. The related MHC class I SD/EXSL cytoplasmic motif is required for constitutive endocytosis of HLA class I molecules, and the latter serine residue (position 335) has been shown to be a phosphorylation site in HLA-B7 molecules (58, 59). Interestingly, cytoplasmic tail sequences of rat and mouse, but not human, MHC class I-related (MR1) genes also contain a similar serine-containing motif (NEGSS) (62–64). The pCD1.1 CDPSS, human and...
mouse MHC class Ia SD/EXSL, and rodent MR1 NEGSS sequences share the following features: an amino acid bearing a hydrophobic, uncharged polar functional side chain in position 1; an acidic amino acid in position 2; any amino acid in position 3; and a conserved serine residue in position 4. The pCD1.1 and MR1 motifs both have a serine in position 5, while leucine occupies that position in the MHC class I cytoplasmic motif. It has been suggested that diversity within the CD1 family may have evolved to efficiently survey distinct lipid Ags that are subsequently shuttled to different cellular compartments (57). This capability might provide for the host-variable Ag responsiveness, perhaps suited to specific pathogen niches. Distinct cytoplasmic tail sequences targeting different intracellular pathways would presumably be required to effect such a mechanism.

Perhaps also enabling varietal responsiveness to different pathogens, the number and diversity of CD1 genes appear to vary among animal species. For example, rodents lack group 1 CD1 genes (25, 26), the sheep genome encodes at least four copies of a CD1b homologue (24), and the rabbit genome is thought to harbor at least eight CD1 genes bearing similarity to CD1b or CD1d (23). Based on Southern blot analysis with a conserved CD1 α3 domain probe, we estimate that the pig genome encodes at least four distinct CD1 loci, assuming the existence of at least one restriction enzyme site within each gene and the absence of restriction sites within exon 4 (α3 domain) of distinct CD1 genes (23, 24) (T. Chun and H. R. Gaskins, unpublished observations). Preliminary sequence data indicate that another CD1 gene bearing similarity to CD1b is closely linked to pCD1.1 (T. Chun and H. R. Gaskins, unpublished observations). Our attempts to isolate pig CD1d-like genes from both genomic and cDNA libraries have been unsuccessful. Additional work is required to determine the extent to which the pig genome encodes homologues or variants of extant CD1 genes.

Variation in the collection of CD1 genes among mammalian species may reflect postspeciation selective pressures. The ancestral CD1 gene is thought to have diverged from MHC ancestral genes around the time of avian-mammal divergence (250–300 million yr ago) because CD1 genes are more closely related to avian class I MHC sequences than to mammalian class I or II MHC genes (65). Emerging evidence indicates that some MHC class Ib molecules share with CD1 a preference for relatively invariant hydrophobic Ags (66, 67). The MHC-encoded class Ib genes are thought to have evolved by independent duplications of class Ia genes following the divergence of the mammalian orders (68). Perhaps taxon-specific collections of CD1 genes may similarly reflect independent duplications from an ancestral CD1 gene after speciation. In that regard, taxa diversity in CD1 or class Ib MHC genes may relate to invariant hydrophobic Ag specificities associated with the major pathogens encountered or may relate to the diversity of Ag responsiveness otherwise enabled by the collection of MHC Ag presentation molecules within a particular taxon. These possibilities add intrigue to the objective of understanding where and how CD1 molecules are loaded with antigenic ligands and suggest that further characterization of CD1 genes in various taxa may clarify to some extent the evolutionary history of the CD1 family.

Similarly, detailed genomic maps of the chromosomal regions encoding pig CD1 and MR1 genes should provide insight into the emergence of MHC paralogous regions and perhaps refine our understanding of phylogenetic relationships among primates, rodents, and artiodactyls. A recent DNA sequence-based analysis estimated that pigs and primates last shared a common ancestor 80 million yr ago, which is 20 million yr after the estimated separation of rodent and primate lineages (69). The five CD1 genes and the single MR1 locus localize proximally on the long arm of human chromosome 1, while the syntenic region is split in the mouse genome, with mouse MR1 on chromosome 1 and the duplicated CD1 genes on chromosome 3 (70, 71). The human chromosome 1 region where MR1 and CD1 genes reside appears to have split into syntenic regions on chromosomes 1 and 3 in mouse, with the human chromosomal arrangement thought to reflect the ancestral CD1-MR1 genomic relationship (72). Recent results reported by Riegert et al. (73) demonstrate the conservation of what appears to be a single-copy MR1 gene in the porcine genome, and we have unpublished evidence indicating that the pig CD1 genes may lie in a cluster similar to those in other mammalian species (1, 65). It is now necessary to map the pig CD1 and MR1 genes and inspect closely their local genomic context.

Unique aspects of T cell biology in the pig make this species particularly suitable for studying the generation of T cell subset diversity and distribution. For example, pigs harbor sizable populations of both double-negative (DN; CD4−CD8−) and double positive (CD4+CD8+) lymphocytes (32, 33, 60). In addition, pigs share with other artiodactyls relatively prominent populations of γδ TCR+ T cells (30). Little is known about the functions of the pig CD4/CD8 DN or γδ TCR+ T cell populations or about possible artiodactyl-specific parameters underlying their existence. Further characterization of pig CD1 molecules may offer promise in that regard. Human CD1-reactive T cells that recognize target cells apparently in the absence of cognate Ag have been identified for three different CD1 molecules (CD1a, -c, and -d) (14, 74, 75). Those CD1-reactive T cells are often CD4/CD8 DN or express only the CD8 α-subunit at low levels and can express either αβ or γδ TCRs (1).

Additional knowledge of molecular and cellular aspects of Ag presentation in the pig is also needed in the xenotransplantation field. Several studies on human anti-pig responses have shown that human NK cells can kill porcine cells, bringing into question the nature of the targets recognized on porcine cells (76). Some MHC class Ia molecules are thought to deliver a negative signal to NK cells, such that cells bearing MHC class Ia molecules recognized by an inhibitory receptor on an NK cell are protected from cytolyis (77). Pig cells might lack such a signal due to incompatibility or, alternatively, other ligands may be recognized by NK cell receptors that transmit a positive signal for killing (77). Recently, CD1-restricted NK cells possessing αβ TCRs were identified in humans and mice (1, 78). Further, studies with MHC class II knockout mice have identified T cells that are reactive with mCD1.1 (78), and recent mCD1.1 knockout mouse models demonstrate that CD1d is critical for NK1+ T cell development and generation of the Th2 response (79–81). It may prove valuable to examine the relationship between human NK cell activity and pCD1.1 expression on APCs during xenograft rejection.

The pace at which the molecular basis of CD1 Ag binding is being defined exceeds the elucidation of the extent and immunologic functions of CD1 Ag presentation. The present data are consistent with the working hypothesis that distinct CD1 isoforms may diversify host options for immune responsiveness to a spectrum of pathogens, perhaps in a niche-dependent manner. A number of the more recalcitrant pathogens exhibit unique compartmentalized lifestyles, creating currently a significant obstacle for clinical immunology. Continuing efforts to define comparative aspects of CD1 Ag presentation should contribute both clinical and theoretical advances.

Acknowledgments
We are grateful for the technical assistance of R. Aminov, R. J. Feltes, H. Oakley, and W. Schnitzlein, and for suggestions from D. B. Anderson and...
References


