Two CD1 genes map to the chicken MHC, indicating that CD1 genes are ancient and likely to have been present in the primordial MHC

Jan Salomonsen^{†‡}, Maria Rathmann Sørensen^{†‡}, Denise A. Marston[§], Sally L. Rogers[§], Trevor Collen[§], Andrew van Hateren[§], Adrian L. Smith[§], Richard K. Beal[§], Karsten Skjødtt[¶], and Jim Kaufman^{§||}

[†]Department of Pathobiology, Laboratory of Immunology, Royal Veterinary and Agricultural University, Stigbøjlen 7, DK-1870 Frederiksberg C, Denmark; [§]Division of Immunology and Pathology, Institute for Animal Health, Compton, Berks RG20 7NN, United Kingdom; and [¶]Department of Immunology and Medical Microbiology, University of Southern Denmark, Winsløvparken 21, DK-5000 Odense C, Denmark

Edited by Max D. Cooper, University of Alabama, Birmingham, AL, and approved April 6, 2005 (received for review December 10, 2004)

CD1 molecules play an important role in the immune system, presenting lipid-containing antigens to T and NKT cells. CD1 genes have long been thought to be as ancient as MHC class I and II genes, based on various arguments, but thus far they have been described only in mammals. Here we describe two CD1 genes in chickens, demonstrating that the CD1 system was present in the last common ancestor of mammals and birds at least 300 million years ago. In phylogenetic analysis, these sequences cluster with CD1 sequences from other species but are not obviously like any particular CD1 isotype. Sequence analysis suggests that the expressed proteins bind hydrophobic molecules and are recycled through intracellular vesicles. RNA expression is strong in lymphoid tissues but weaker to undetectable in some nonlymphoid tissues. Flow cytometry confirms expression from one gene on B cells. Based on Southern blotting and cloning, only two such CD1 genes are detected, located ≈ 800 nucleotides apart and in the same transcriptional orientation. The sequence of one gene is nearly identical in six chicken lines. By mapping with a backcross family, this gene could not be separated from the chicken MHC on chromosome 16. Mining the draft chicken genome sequence shows that chicken has only these two CD1 genes located ≈ 50 kb from the classical class I genes. The unexpected location of these genes in the chicken MHC suggests the CD1 system was present in the primordial MHC and is thus ≈ 600 million years old.

evolution | paralogy | antigen presentation | nonclassical | avian

MH C class I and II molecules play crucial roles in the immune system by presenting peptide antigen to T lymphocytes, as well as natural killer cells recognizing class I molecules. Both MHC class I and II genes have been described in all jawed vertebrates (including bony and cartilaginous fish) and are thought to have arisen at the same time as recombination activating, T cell receptor, and antibody genes ≈ 600 million years ago (1–3).

CD1 molecules play an equally important role in the immune system, presenting lipid, glycolipid, and lipopeptide antigens to T and NKT cells (4–8). CD1 genes have long been thought to be as ancient as class I and II genes, based on several pieces of evidence, but thus far they have been described only in mammals (4, 9–11).

That the CD1 system is related to MHC class I and II systems is clear, but the details of the evolutionary relationship are not (4–8). By comparison of gene and protein sequences as well as 3D folds, CD1 is equally related to class I and II sequences and proteins. However, by intron/exon structure and domain organization, CD1 is similar to class I genes and molecules, with a single transmembrane glycoprotein heavy chain of three extracellular domains that binds β_2 -microglobulin (compared to class II genes and molecules, with two transmembrane glycoproteins, each of two extracellular domains). Conversely, CD1 molecules resemble class II molecules in terms of exten-

sive intracellular trafficking, antigen loading in intracellular vesicles, and tissue distribution restricted to antigen-presenting cells and a few other cell types (compared with class I molecules, which have most antigen loading in the endoplasmic reticulum and a relatively ubiquitous tissue distribution). In terms of polymorphism, CD1 genes are like the monomorphic and oligomorphic nonclassical class I and II genes rather than the highly polymorphic classical class I and II genes. These observations have been used to suggest there are three separate and ancient antigen-presentation systems: class I, class II, and CD1.

Despite the similarities to MHC class I and II genes, CD1 genes are found in a cluster outside of the MHC, on chromosome 1 in humans and chromosome 3 in mice (4, 7, 11). In fact, many genes around the CD1 genes are related to genes in the MHC, as well as genes in several other locations in the genome. This has led to the hypothesis that there were two rounds of genome-wide duplication at the time of the appearance of the jawed vertebrates. In this view, a single primordial MHC region was multiplied into four paralogous regions, followed by differential silencing and evolutionary divergence of individual genes (12–14). Thus, CD1 genes could have come about in at least two ways. One possibility is that class I and II genes were present in the primordial MHC and were distributed into the paralogous regions, with class I genes evolving into CD1 genes in the CD1 region (13). Alternatively, the CD1, class I, and class II genes arose in the primordial MHC and were distributed into the paralogous regions, with the CD1 genes being deleted in the mammalian MHC and the class I and II genes deleted in the mammalian CD1 region.

Multiple CD1 genes have been found in every mammalian species examined (4, 7, 9). Five CD1 genes are found in humans, with the *CD1A*, *-B*, and *-C* genes forming the type 1 sequence family, the *CD1D* gene being the archetype of the type 2 family and the *CD1E* gene being considered as intermediate. In contrast, mice and rats have only CD1D genes, and sheep may have only type 1 genes. All of these genes encode molecules with hydrophobic amino acids in key positions in the antigen-binding domains. However, the exact hydrophobic amino acids in key positions vary, apparently so that different CD1 molecules bind different lipid structures, as seen in the 3D structures of human CD1a and CD1b and mouse CD1d molecules (15–18). The recycling motifs found in the cytoplas-

This paper was submitted directly (Track II) to the PNAS office.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. AY849318–AY849320).

See Commentary on page 8399.

[†]J.S., M.R.S., and D.A.M. contributed equally to this work.

^{||}To whom correspondence should be addressed. E-mail: jim.kaufman@bbsrc.ac.uk.

© 2005 by The National Academy of Sciences of the USA

mic tail of the molecules also vary, apparently so that different CD1 molecules can be loaded with lipid antigen in different compartments (7, 11, 19). Finally, the precise tissue distributions of the CD1 molecules vary (4, 7, 20, 21). Human type 1 molecules are expressed primarily on dendritic cells and other professional antigen-presenting cells as well as some thymocytes, with CD1c expressed on subsets of B cells. Human CD1d expression is broader, with expression on most dendritic cells, B cells, monocytes, and macrophages as well as some nonlymphoid cells, including some in gut and liver. Mouse CD1 is found in addition on subsets of thymocytes and T cells. The "traffic hypothesis" (11) attempts to explain at least some of these complexities by proposing that the important point is to have CD1 isoforms that penetrate enough different endosomal compartments to be loaded with the appropriate lipid antigens, a selection that then determines the optimal binding specificities and tissue distributions.

Despite the speculation, it has not been clear that the CD1 system is old, because CD1 genes have been described only in mammals. Moreover, there has been no evidence whether the CD1 system arose in the primordial MHC or evolved from class I genes after genome-wide duplication. Finally, there has been no evidence of what the ancestral gene might be like. In this paper, we describe two CD1 genes in chickens, going some ways towards answering these questions.

Materials and Methods

Animals, RNA, and Genomic DNA. Chicken lines were bred and maintained under specific pathogen-free conditions at the Institute for Animal Health (IAH), as detailed (22), or were from the Institute of Molecular Genetics, Prague (courtesy of Karel Hala). Tissues were removed from a 5-week-old C-B12 cockerel at the IAH and stored in RNAlater (Ambion, Austin, TX) at -20°C . Intestinal enterocytes from embryonic day 14–15 C-B12 embryos were isolated from the PBS-45% Percoll interface (23). The HD11 cell line, derived from chicken macrophages, was a gift of Fred Davison (Institute of Animal Health). Cell sorts from adult C-B12 spleen cells were performed by using an Automacs cell sorter (Miltenyi Biotec, Auburn, CA), by using the primary monoclonal antibodies (all from Southern Biotechnology Associates) AV20–phycoerythrin (PE) conjugate (Bu1⁺ cells), CT-8–PE conjugate (CD8 α ⁺ cells), TCR2–PE conjugate (TCR α β 1⁺ cells), TCR1–biotin conjugate (TCR γ δ ⁺ cells), TCR3–biotin conjugate (TCR α β 2⁺ cells), or CD4–FITC conjugate (CD4⁺ cells). Cells were sorted by using either anti-PE magnetic beads or anti-mouse IgG1 magnetic beads. Total RNA was extracted by using RNeasy kits (Qiagen, Chatsworth, CA), according to the manufacturer's instructions, eluted from the column using RNase-free water, and stored at -70°C . Genomic DNA was isolated from erythrocytes by salting-out.^{††} Standard molecular biology techniques were used (24), except where specified.

cDNA Coding Sequences. Total cDNA was produced by using the Advantage RT-for-PCR kit (Clontech), following the manufacturer's instructions, with oligo(dT) primer (GGCCACGCGTC-GACTAGTACTTTTTTTTTTTTTT), and incubated for 2 min at 70°C , then 1 h at 42°C , and then 5 min at 94°C . The coding region sequence of CD1.2 was from a clone with no 5'UTR and 51 bp of 3'UTR, derived using 0.5 μl (25 ng) of cDNA from various CB tissues and 20 pmol each primer [primer nos. 326119 (ATGTGGCCCCACTGCCTCTTC) and 331407 (GCTTAGAAATGGAGCAGGGAG)] in 50 μl of final volume including 0.8 mM total dNTPs, 1 $\times$ GC-rich enzyme buffer, 1 $\times$ GC-rich resolution solution, 2 units of enzyme (GC-Rich PCR

system, Roche Molecular Biochemicals), with amplification conditions of 3 min at 95°C , 10 cycles of 30 s at 95°C , 30 s with ramp from 70°C down to 60°C , and 1 min at 72°C , followed by 20 cycles of 30 s at 95°C , 30 s at 62°C , and 1 min at 72°C . The coding region sequence of CD1.1 was from a clone including 6 bp of 5'UTR and 318 bp of 3'UTR, derived using cDNA from chicken B cells from the C-B12 subline and 20 pmol each primer [primer nos. 382453 (GACACCATGCAGCCCCATGCCATC) and 380673 (CCCAAACCCCACTCACTCCCACACCTC)] with the same mixtures but amplification conditions of 1 min at 94°C , then 7 cycles of 30 s at 94°C and 3 min at 70°C , followed by 37 cycles of 25 s at 94°C , and 4 min with ramp from 60°C down to 50°C . Sequences for 5'UTR were determined from products of RACE-PCR by using the Marathon kit (Clontech).

Gene Sequences. Fragments of genes were initially amplified from CB genomic DNA by using primers from the cDNA sequence, and the remainder was cloned by PCR by using the Universal Genome Walker kit (Clontech). Four libraries were constructed after digestion with *DracI*, *EcoRV*, *PvuII*, and *StuI*, with an extra ligation with Ready-to-Go T4 DNA ligase (Amersham Pharmacia Biotech), followed by two amplifications using the GC-Rich PCR system (Roche Molecular Biochemicals). The first reaction used 60 ng of library DNA in a 50- μl final volume including 20 pmol each primer, 0.8 mM dNTP, 1 $\times$ GC-rich buffer, 1 $\times$ GC-rich resolution solution, 2 units of enzyme, and with amplification conditions of 1 min at 94°C , then 7 cycles of 25 s at 94°C and 3 min at 70°C , then 37 cycles of 25 s at 94°C and a 4-min ramp of 60°C down to 50°C end, then 7 min at 50°C . The second (nested) reaction used the product of the first reaction as template, and the PCR products were TOPO-cloned. Fragments were amplified from genomic DNA of other chicken lines by using primers based on the CB sequence. Genomic and cDNA sequences were analyzed by Prism 310 (Applied Biosystems) or CEQ8000 (Beckman-Coulter).

Expression Clones. The inserts of the expression constructs consisted of the signal sequence of chicken classical class I (BF2*1201) cDNA followed by a myc tag, followed by the cDNA sequence for the mature protein of CD1.1 or CD1.2. All amplifications were with 15 pmol of each primer, 0.2 mM total dNTPs, 1 $\times$ enzyme buffer, 1 mM MgSO_4 , 2 units of Pfx polymerase (Invitrogen). The signal sequence-myc fragment (with 5' BamHI site) was amplified from a cDNA clone by using primers c1622 (GCGGAAGCTTGCGATGGGGCCGT-GCGG) and c1784 (GGTATGCAGATCCTCTCTGAGAT-GAGTTTTTGTTCGGCCGCCGCCCGCCACAC) with amplification conditions of 2 min at 96°C , 15 cycles of 30 s at 96°C , 30 s at 52°C , and 30 s at 68°C . Myc-CD1 fragments (with 3' XbaI site) were amplified from cDNA clones by using CD1.1 primers c1796 (CCCTCTAGATTAGATGTGCCCTCCCTCATC) and c1799 (AAACTCATCTCAGAAGAGGATCTGCCA-GAGGGCTCCCACATGCTC) or CD1.2 primers c1797 (CCCTCTAGATTAGATGAGGAATGTCCGGAAATC) and c1800 (AAACTCATCTCAGAAGAGGATCTGTGCCCTC-CCCCTGAAGAATCAC) with amplification conditions of 2 min at 96°C , 25 cycles of 30 s at 96°C , 30 s at 60°C , and 30 s at 68°C . Joined fragments were amplified from 1 μl each of gel-purified PCR product by using CD1.1 primers c1798 (CTT-GGATCCGCGATGGGGCCGTGCGGGGCG) and c1796 or CD1.2 primers c1798 and c1797 with amplification conditions of 2 min at 96°C , 25 cycles of 30 s at 96°C , 30 s at 55°C , and 30 s at 90°C . Purified products were digested with BamHI and XbaI, cloned into pcDNA3.1⁺ (Invitrogen), and the sequence was confirmed. COS7 cells [10⁶ cells in 100 μl of Cell Line Nucleofector Kit V (Amaya, Gaithersburg, MD)] were transfected [program A-24 with a Nucleofector I (Amaya)] with 5 μg of DNA from ENDO-free maxipreps (Qiagen).

^{††}Miller, S., Dykes, D. & Polesky, H. (1988) *Nucleic Acids Res.* 16, 215 (abstr.).

Fig. 2. Phylogenetic analysis of the amino acid sequences of chicken CD1.1 and CD1.2. The tree was constructed by using full-length protein sequences, with the neighbor-joining method of the PHYLIP program from the GCG package (Genetics Computing Group, Madison, WI). Numbers indicate bootstrap values from 500 replicates. GenBank accession nos. are as follows: human CD1 isotypes as in Fig. 1 legend; classical class I sequences from human, chicken, and salmon [HLA-A2 (P01892), HLA-B27 (P10317), HLA-C (P10317), BF2*1201 (AL023516), Sasa-UBA (AF504019)]; nonclassical class I sequences from human and chicken [MICB (P79525), YF1w*7.1 (Q9BCW3)]; and classical class II sequences from human [HLA-DRB (P01913), HLA-DQB (P05537), HLA-DPB1 (Q30174)].

the PNAS web site). Comparison of amino acid and nucleotide sequences of CD1.2 from several chicken lines showed minimal polymorphism, like CD1 but unlike classical class I and II genes (Figs. 1 and 3). The total number of genes detected by Southern blots at low stringency (Fig. 3) and by cloning (data not shown) is low, unlike classical and closely related nonclassical class I sequences in many mammals. The tissue distribution shows the highest RNA expression in lymphoid tissues, with weaker expression in certain nonlymphoid tissues but undetectable expression in others, including embryonic intestinal epithelial cells (Fig. 4). Analysis of RNA from blood cells showed that both genes are expressed in Bu-1⁺ cells (B cells and some macrophages) and CD8 α ⁺ cells (some T and NK cells), and CD1.2 is expressed in TCR1⁺ cells ($\gamma\delta$ T cells) but is nearly undetectable in TCR2⁺ and TCR3⁺ cells ($\alpha\beta$ T cells). If anything, such tissue and cell distributions are more similar to CD1d than to other isotypes. Analysis by flow cytometry shows that the monoclonal antibody CB3 reported to identify cells of the B cell lineage (30) recognizes CD1.1 but apparently not CD1.2 expressed on the surface of transfected mammalian cells (Fig. 4), in agreement with the expression analysis at the RNA level. Interestingly, CD1.1 was well expressed compared with CD1.2, which might be because of the availability of a ligand in COS7 cells.

The Chicken CD1 Sequences Are Not Obviously More Like One CD1 Isotype than Another. The dendrogram shows that the chicken cDNA sequences cluster together outside the human CD1 sequence cluster (Fig. 2). This tree is similar to the topology of trees that include many more known mammalian sequences. The two chicken sequences are nearly identical in the α 3 (Ig-like) domain but are <30% identical in other domains (Fig. 1).

Comparison using individual domains gives dendrograms with similar topologies (Fig. 5). The presence of certain key residues in the antigen-binding domains suggests that the chicken CD1 proteins might bind lipid-containing ligands similar to one or another mammalian CD1 isotype (analysis not shown), whereas the presence of CD1.1 with a potential tyrosine-based motif and CD1.2 with a modified dileucine motif supports the “traffic hypothesis” of CD1 evolution (11).

Only Two CD1 Genes Were Found in the Chicken Genome, Next to Each Other and Located in the MHC. Very simple patterns were found by Southern blot at low stringency with the CD1.2 cDNA (Fig. 3B and data not shown), indicating only two genes with sufficient identity to be detected by nucleotide hybridization. Genome walking and sequencing identified the two genes, $\approx$ 800 bp apart and in the same transcriptional orientation (Fig. 3A).

Portions of the *CD1.2* gene were amplified from genomic DNA of six inbred white leghorn chicken lines and compared with the whole gene sequence from the CB line. Comparison of the gene sequences showed very limited polymorphism, which explained some of the Southern blot data (Figs. 1 and 3 and data not shown). In particular, the XhoI site in exon 2 of this gene was polymorphic, so it was used to map the gene in a backcross family. Unexpectedly, the gene was not separated by recombination in this family from the MHC (Fig. 3C and D).

During the preparation of this manuscript, the draft chicken genome sequence of a Red Jungle Fowl became available (www.ensembl.org/chicken), and BLAST analysis showed that the whole genome contains only these two sequences with sufficient sequence identity to be detected. Although the assembly of chromosome 16 was very poor (32), we were able to detect the 3' end of the *CD1.2* gene assembled at the far right of the major ultracontig from chromosome 16 (see Fig. 3 legend for details) and found fragments representing the rest of both genes in the unassembled and unassigned sequences. Together with data from the previously sequenced chicken MHC (33), these data locate our CD1 contig $\approx$ 50 kb from the outermost classical class I gene (*BF2*, previously called B-F major) of the B-F/B-L region. Furthermore, no CD1-like genes were found near the single spectrin gene thus far identified in the chicken genome, whereas other genes expected to be syntenic (like *C-reactive protein*, *KIRREL*, and *SLAMF7*) were found but not assembled in the chicken genome.

CD1 Genes Must Be Ancient and Were Likely in the Primordial MHC. By describing CD1 genes in a nonmammalian vertebrate, this paper makes three important points.

First, the CD1 system has long been thought to be ancient (4, 9–11), based on the equal phylogenetic distance from class I and II sequences and on the mix of class I and II features (intron/exon structure and intracellular recycling). However, these characteristics of CD1 could be accounted for by rapid sequence divergence under strong selective pressure. The presence of CD1 genes in the chicken shows that the CD1 system was present in the last common ancestor of mammals and birds at least 300 million years ago and is indeed ancient.

Second, the CD1 system has long been known to be located outside of the MHC in mammals, in a region later identified as an MHC-paralogous region (4, 7, 11–13). The general view is that a single primordial MHC region was duplicated twice into four paralogous regions in ancestors of the jawed vertebrates, followed by differential silencing, deletion, and divergence of individual genes. One possibility was that only class I (and II) genes were present in the primordial MHC and were distributed into the paralogous regions, with class I genes evolving into CD1 genes in the CD1 region and acquiring different sequence features under strong selective pressure (13). However, the simplest interpretation of the data in this paper is that the CD1

Fig. 3. Location of the chicken CD1 genes. (A) Intron/exon structure (from 9.1 kb of sequence from the CB chicken line, MHC haplotype B12, deposited with GenBank accession no. AY849318), including locations of SNPs for CD1.2 by comparisons of sequences from lines C-B4, 61, 72, WL, 151s, 0, and N. Exons indicated by boxes (closed for coding, open for 3'UTR) for CD1.1: exon 1, 1374 (start protein 1422)–1482; exon 2, 1672–1920; exon 3, 2791–3072; exon 4, 3160–3444; exon 5, 3531–3621; exon 6, 3697 (end protein 3829)–4208; polyadenylation site 4186–4191; CD1.2: exon 1, 5084 (start protein 5090)–5150; exon 2, 5478–5756; exon 3, 6259–6540; exon 4, 6629–6913; exon 5, 6994–7087; exon 6, 7160 (end protein 7202)–7772; and cryptic polyadenylation site 7746–7750 predicted by GENEVIEWER (<http://125.itba.mi.cnr.it/~webgene/geneviewer.html>). Analysis of the draft chicken genome sequence (www.ensembl.org/chicken) showed that the 3' end of the sequence (starting at position 6986 and including the last two exons of CD1.2) overlaps the end of contig 638.25, placing CD1.2 to the right of a seven-transmembrane receptor gene (GENSCAN0000040413) but in the opposite orientation (and likely to be in same transcriptional orientation to the major classical class I gene, BF2). SNPs (indicated by vertical lines) from partial sequences starting at position 4608 are T5077C, G5206del, T5851A, A5920G, C5942T, C6287A, C6433G, G6507A, C6531G (used for mapping), C6783T, C6985T, G7083A, T7211A, C8089T, and A8103G. Numbers below the line indicate the nucleotide position in kb. (B) Southern blot of genomic DNA from seven inbred lines of chickens with defined MHC type [lines 61 (B2), C-B4 (B4), C-B12 (B12), WL (B14), 151s (B15), P2a (B19), and N (B21)] digested with EcoRV followed by either NheI, StuI, or XhoI, and probed with a 1.3-kb exon 3 through exon 4 of the CD1.2 gene. The probe was produced from 5 ng of genomic B21 genomic DNA in 30 μ l of final volume including 20 pmol of each primer [c1124 (CGGTTTCATGATGAGATGAC), c1125 (CAACAGTGAAGATGCACTG)], 1 $\times$ enzyme buffer, 0.2 mM total dNTPs, 1.5 mM MgCl₂, 2.5 units of proof-reading polymerase Pwo (Hybaid, Middlesex, U.K.) with amplification conditions of 1 min at 96°C, 30 cycles of 30 s at 96°C, 30 s at 55°C, and 1 min at 72°C, and twice purified by agarose gel electrophoresis. Southern blots were performed as described (26, 27), with washing at low stringency (2 $\times$ SSC at 65°C for 30 min). Lines WL, 151s, and P2a lack the XhoI site and show only a single major band of $\approx$ 1,320 bp, with the other lines showing the CD1.2 gene cut into two bands of $\approx$ 815 and 505 bp, leaving behind the band representing CD1.1. (C) Mapping of the CD1.2 gene. Genomic DNA from the two parents [(151s $\times$ N) F₁ hen and 151s cockerel] and 55 progeny of the Compton backcross reference family (28) were used for mapping as described (27). For PCR, 5–50 ng of genomic DNA were amplified by using primers c1124 and c1125 and conditions as described in the Fig. 3B legend. The products were separated by an agarose gel, purified from excised bands by using Minelute columns according to the manufacturer's protocol (Qiagen), digested with 10 units of XhoI (Invitrogen) with the supplied buffer at 37°C for 5 h, and visualized with ethidium bromide on an agarose gel. M

Fig. 4. Tissue distribution of CD1. RT-PCR from RNA from (Top) six immunologically isolated spleen cell populations and (Middle) 10 tissues, one macrophage line (HD11 cells), and isolated embryonic enterocytes. First-strand cDNA was prepared by incubation of 2 h at 42°C of 4 μ l of RNA preparation and 80 units of RNasin (Promega) in 40 μ l of final volume including 1 $\times$ first-strand buffer, 10 mM DTT, 20 mM total dNTPs, 1 μ g oligo(dT), and 400 units of reverse transcriptase Superscript II (Invitrogen). Using primers for CD1.2 (c1124 and c1125, described in the legend to Fig. 3, amplifying 683 bp of cDNA) and β -actin [c765 (CTGGTCGTACCACTGGTATTG), c766 (CAGGTGGCGCAATGATCTTG)], the PCR mixtures were 0.5 μ l of cDNA in 50 μ l of final volume including 15 pmol each primer, 0.2 mM total dNTPs, 1 $\times$ enzyme buffer, 1.5 mM MgCl₂, and 2.5 units of recombinant Taq polymerase (Invitrogen), with amplification conditions of 2 min at 96°C, 35 cycles of 30 s at 96°C, 30 s at 55°C, and 1 min at 68°C. Using primers for CD1.1 [c1728 (GTGGAGCTGCATCACAGACACCATG), c1729 (CCCAAACCCCACTCACTCCACACCTC), amplifying 1,553 bp of cDNA], the PCR mixtures were the same but included 5% DMSO with amplification conditions of 2 min at 94°C, then seven cycles of 30 s at 94°C, 2 min at 68°C, followed by 35 cycles of 30 s at 94°C, 30 s at 60°C, and 1 min at 68°C. Flow cytometric analysis (Bottom) of peripheral blood leukocytes from a line C-B12 chicken and COS cells transiently transfected with pcDNA3.1 with no insert, with CD1.1 cDNA or with CD1.2 cDNA stained with TRT1 (IgG1 isotype-matched control, filled gray peak), monoclonal antibody to myc peptide (Sigma, gray line) or monoclonal antibody CB3 (ref. 30, solid line). Flow cytometric analysis was as described (31), with second antibody polyclonal FITC-conjugated goat anti-mouse Ig (DAKO) using FACSCalibur (Beckman).

and class I and II genes arose in the primordial MHC and were originally all present in the paralogous regions, with subsequent deletion of the CD1 genes from the mammalian MHC and

indicates the DNA markers (sizes of bands on the side); 15 and 15/N indicate DNA from male and female parents, respectively; and numbers indicate DNA from individual chickens in the mapping family. Progeny 31 and 32 have only the 1,320-bp band corresponding to two copies of the B15 allele, whereas progeny 29, 30, and 33 have 815- and 505-bp bands corresponding to one copy of the B21 allele and a 1,320-bp band corresponding to one copy of the B15 CD1 gene allele. Linkage analysis was performed by using MAP MANAGER (29). A logarithm of odds score (Z) value above three ($P = 0.001$) in two-point analysis was used as the significant threshold for linkage. Map distances were calculated with the Haldane map function (cM). (D) Current genetic map of chicken chromosome 16, with result of mapping CD1 to the B locus as defined by BG, BF, BLB, and TAP probes, among others. Genetic distances are in cM. As discussed (22), the distances and locus order in this map supersede those found in various web sites and publications.

deletion of the class I and II genes from the mammalian CD1 region. Thus, the data suggest that CD1 genes will be found in bony and cartilaginous fish, support the concept of evolution of the MHC and MHC-like genes by paralogous duplication, and argue strongly for a very early origin of CD1 genes, ≈ 600 million years ago.

Third, the CD1 system has multiple isotypes with different functions, whose representation varies among mammals (4–7). Humans have three type 1 genes, one type 2 gene, and an intermediate gene (*CD1E*), whereas mice have only type 2 genes, and sheep may have only type 1 genes. Thus, there are CD1-dependent responses in one mammal that are lacking or modified in other mammals, and it is unclear which functions are ancestral. The “traffic hypothesis” of CD1 evolution proposes that selection in each species is for several CD1 molecules to sample different intracellular vesicles, rather than to have par-

ticular isoforms (11). That the one chicken sequence has a modified dileucine motif that may limit the protein to endosomal recycling, whereas the other chicken sequence has a potential tyrosine-based motif that may allow recycling to lysosomes, provides support for this hypothesis. The possibility that there is selection to bind particular antigens awaits the identification of ligands for the two chicken CD1 proteins.

We thank Malene Dreyer Nielsen and Alma Lillian Orantes Poza Thorsen for excellent technical assistance, Fred Davison (Institute for Animal Health) for the gift of HD11 cells, Chen-Lo Chen (University of Alabama, Birmingham) for the kind gift of CB3 monoclonal antibody, Vincenzo Cerundolo for useful discussion, Gillian Griffiths for critical reading of the manuscript, two reviewers for valuable comments, and the Novo-Nordisk Foundation of Denmark and the Biotechnology and Biological Sciences Research Council of the U.K. for support.

1. Flajnik, M. & Kasahara, M. (2001) *Immunity* **15**, 351–362.
2. Kasahara, M., Suzuki, T. & Du Pasquier, L. (2004) *Trends Immunol.* **25**, 105–111.
3. Eason, D., Cannon, J., Haire, R., Rast, J., Ostrov, D. & Litman, G. (2004) *Semin. Immunol.* **16**, 215–226.
4. Porcelli, S. & Modlin, R. (1999) *Annu. Rev. Immunol.* **17**, 297–329.
5. Jayawardena-Wolf, J. & Bendelac, A. (2001) *Curr. Opin. Immunol.* **13**, 109–113.
6. Matsuda, J. & Kronenberg, M. (2001) *Curr. Opin. Immunol.* **13**, 19–25.
7. Brigl, M. & Brenner, M. (2004) *Annu. Rev. Immunol.* **22**, 817–890.
8. Moody, D., Young, D., Cheng, T., Rosat, J., Roura-Mir, C., O'Connor, P., Zajonc, D., Walz, A., Miller, M., Levery, S., *et al.* (2004) *Science* **303**, 527–531.
9. Martin, L., Calabi, F. & Milstein, C. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 9154–9158.
10. Blumberg, R., Gerdes, D., Chott, A., Porcelli, S. & Balk, S. (1995) *Immunol. Rev.* **147**, 5–29.
11. Dascher, C. & Brenner, M. (2003) *Trends Immunol.* **24**, 412–418.
12. Kasahara, M., Nakaya, J., Satta, Y. & Takahata, N. (1997) *Trends Genet.* **13**, 90–92.
13. Kasahara, M. (1999) *Immunol. Rev.* **167**, 17–32.
14. Abi Rached, L., McDermott, M. & Pontarotti, P. (1999) *Immunol. Rev.* **167**, 33–44.
15. Zeng, Z., Castano, A., Segelke, B., Stura, E., Peterson, P. & Wilson, I. (1997) *Science* **277**, 339–345.
16. Gadola, S., Zaccai, N., Harlos, K., Shepherd, D., Castro-Palomino, J., Ritter, G., Schmidt, R., Jones, E. & Cerundolo, V. (2002) *Nat. Immunol.* **3**, 721–726.
17. Zajonc, D., Elsliger, M., Teyton, L. & Wilson, I. (2003) *Nat. Immunol.* **4**, 808–815.
18. Batuwangala, T., Shepherd, D., Gadola, S., Gibson, K., Zaccai, N., Fersht, A., Besra, G., Cerundolo, V. & Jones, E. Y. (2004) *J. Immunol.* **172**, 2382–2388.
19. Moody, D. & Porcelli, S. (2003) *Nat. Rev. Immunol.* **3**, 11–22.
20. Blumberg, R., Terhorst, C., Bleicher, P., McDermott, F., Allan, C., Landau, S., Trier, J. & Balk, S. (1991) *J. Immunol.* **147**, 2518–2524.
21. Chen, Q. & Jackson, N. (2004) *J. Immunol.* **172**, 5512–5521.
22. Jacob, J., Milne, S., Beck, S. & Kaufman, J. (2000) *Immunogenetics* **51**, 138–147.
23. Goodman, T. & Lefrancois, L. (1988) *Nature* **333**, 855–858.
24. Sambrook, J., Fritsch, E. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
25. Sandoval, I. & Bakke, O. (1994) *Trends Cell Biol.* **4**, 292–297.
26. Kaufman, J., Salomonsen, J. & Skjød, K. (1989) *Immunogenetics* **30**, 440–451.
27. Salomonsen, J., Marston, D., Avila, D., Bumstead, N., Johansson, B., Juul-Madsen, H., Olesen, G., Riegert, P., Skjød, K., Vainio, O., *et al.* (2003) *Immunogenetics* **55**, 605–615.
28. Bumstead, N. & Palyga, J. (1992) *Genomics* **13**, 690–697.
29. Manly, K. (1993) *Mamm. Genome* **4**, 303–313.
30. Pickle, J., Chen, C. & Cooper, M. (1990) *Immunogenetics* **32**, 1–7.
31. Salomonsen, J., Dunon, D., Skjød, K., Thorpe, D., Vainio, O. & Kaufman, J. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 1359–1363.
32. International Chicken Genome Sequencing Consortium (2004) *Nature* **432**, 695–716.
33. Kaufman, J., Milne, S., Goebel, T., Walker, B., Jacob, J., Auffray, C., Zoorob, R. & Beck, S. (1999) *Nature* **401**, 923–925.