University of Alberta

Genomic structure and cDNA analysis of two C-type lectin receptors, DCIR and DCAR, in duck (*Anas platyrhynchos*)

by

Xiaoxin Guo 

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science
in
Microbiology and Biotechnology Program

Department of Biological Sciences

Edmonton, Alberta
Fall, 2006

Your file Votre référence
ISBN: 978-0-494-22276-8
Our file Notre référence
ISBN: 978-0-494-22276-8

Abstract

Two C-type lectin immunoreceptor genes were identified in a previous expressed sequence tag project from duck spleen. These sequences were used to screen duck genomic and cDNA libraries. A genomic clone was fully sequenced, containing one dendritic cell immunoreceptor (DCIR) gene and two dendritic cell activating receptor (DCAR) genes arranged in tandem order. The duck DCIR gene featured an immunoreceptor tyrosine-based inhibitory motif (ITIM) in the cytoplasmic domain. The DCAR1 gene carried a stop codon in the carbohydrate recognition domain. Since no transcript was found for DCAR1, it was considered to be a pseudogene. cDNA clones matching the DCAR2 gene were identified and had features typical of activating receptors. Alternatively spliced transcripts were identified for both DCIR and DCAR2. These inhibitory and activating receptors may be involved in the regulation of duck antigen presenting cells.

Acknowledgements

Here I want to thank all the people who gave me endless support, help and love during the last three years.

I will never forget how fantastic Dr. Kathy Magor is, as a great teacher and mentor. I appreciate that she gave me the chance to study here and always encourage me to climb up. I will never forget the assistance and inspiration from our technician, Ms. Deb Moon. I will never forget the moments all lab members (including those who have left the lab and moved on) enjoyed together, celebrating one's birthday or successful defense.

Finally, the last gratitude is to my beloved family members: my wonderful husband who gave me confidence, encouragement and love forever, my parents who always provide me valuable advices and understanding, as well as my grandparents who believe that I can do better.

Table of Content

List of Tables

List of Figures

CHAPTER I. INTRODUCTION

1.1 C-type lectin family

The C-type lectin family is a group of lectins that recognize carbohydrate structures in a calcium-dependent manner. Whether soluble or membrane bound, all the C-type lectin family members share a highly conserved lectin prototype, namely the carbohydrate recognition domain (CRD). The length of all CRD domains is well-conserved in all C-type lectins, which is approximately 130 amino acids (Drickamer, 1988). Sequence alignment of these CRDs revealed 14 invariant and 18 highly conserved residues, including at least two pairs of cysteines that contribute to the formation of intramolecular disulfide bridges (Drickamer, 1993). The spatial structure of the C-type lectin CRD fold originally resolved from the crystal structure of rat is typically made up of two α-helices, five β-strands, and a calcium/sugar recognition surface with irregular loops (Drickamer and Taylor, 1993; Weis *et al.*, 1991). The CRD loop is formed by two closely arranged anti-parallel β-strands at the N- and C-termini, and the CRD region is flanked by two consecutive α-helices. The other three β- strands make up the core beneath the sugar-binding surface (Weis *et al.*,1991). The number of calcium-binding sites in the CRD domain varies among different C-type lectins (from 1 to 4). However, the second site, calcium-binding site 2 (Ca^{2+} 2), mediates

carbohydrate recognition and thus is the most conserved and crucial one (Weis *et al.*, 1992).

1.1.1 Structural basis of sugar recognition by C-type lectins

Most C-type lectins can recognize sugars, with different levels of binding affinity and distinct ligand specificity. Although the ligands of many C-type lectins have already been identified, little is known about the exact structure of the calcium-sugar-lectin binding complex. So far, the only confirmed structure of the binding complex was acquired through the crystal structure of natural mannose binding protein (MBP) binding a terminal mannose structure (Weis *et al.*, 1992). In the Ca^{2+} 2 complex, eight coordination bonds of calcium were formed, two with the 3- and 4- equatorial hydroxyl groups of mannose, and six with the side chains of five amino acid residues. At the same time, sugar-protein interactions are also present between the 3- and 4- hydroxyl groups of mannose and the NH_2-groups from four out of the same five amino acid residues through hydrogen bonds. Therefore, in the binding complex, the presence of five amino acid residues (including two asparagines, two glutamic acids and one aspartic acid) is crucial and becomes the criterion to predict the sugar binding specificity of C-type lectins. Alignment of several mannose binding C-type lectins has revealed that to achieve high mannose binding specificity, three of the five residues (EPN) must be conserved, among which Glu185 and Asn187 are invariable. Additionally,

Pro186, between E185 and N187, is a key connector to maintain the pocket-like conformation of the sugar-binding surface of the receptor (Weis *et al.*, 1996).

Mutation of a mannose-binding lectin in the binding sites, as well as neighboring residues were conducted. The E(185)PN(187) to QPD mutation resulted in the switch from strong mannose to weak galactose specificity (Iobst *et al.*, 1994. Moreover, a mutation at a neighboring residue from His 189 to Trp 189, led to a higher affinity for galactose. The effect of Trp 189 is to pack against the apolar face of galactose and consequently increase binding affinity (Iobst and Drickamer, 1994; Kolatkar and Weis, 1996).

Interestingly, the core structure of a tunicate C-type lectin TC14 displayed an EPS motif involved in galactose binding, instead of the QPD motif (Poget, *et al.*, 1999). The galactose used in this model is a 180° rotation of the one in the above mutation study. In this binding complex, Poget and coworkers only observed one functional calcium-binding site, at which two coordination bonds were formed between calcium and the 3- and 4-OH of D-galactose. Other sites of coordination were completed through ligation with Glu86, Asn89, Asp107 and Asp108. Sugar-protein hydrogen bonds were formed between 3-OH and Glu86 and Ser88, as well as between 4-OH and Asp107 and water. In addition, water molecules also mediated an indirect sugar-protein interaction between other -OH groups of D-galactose and neighboring amino acid residues. In a similar way as in the QPN mediated galactose binding complex, a Trp residue at the position of 100

endows the receptor an even higher affinity to less polar sugars such as D-fucose, than D-galactose.

1.1.2 C-type lectin classification

C-type lectins can be classified upon several different criteria. However, based on protein structure and domain organization or sequence homology, the same seven groups were defined (Drickamer *et al.*, 1993). Group I includes three proto-glycans, which are present in the extracellular matrix. Group II represents type II endocytic receptors, with a short N- terminal cytoplasmic tail containing a signal sequence, and C-terminal CRD domain. Mammalian asialoglycoprotein receptor, chicken hepatic lectin and lymphocyte receptors belong to this group. Group III molecules are called collectins, comprised of collagenous N-terminal region and a C-terminal CRD domain in one polypeptide. Collectins are soluble molecules, arranged in the form of homo-oligomers. Examples in this group include mannose binding proteins (MBP) and pulmonary surfactant apo-proteins, which function in humoral defense. Group IV was composed of selectin molecules. Three types of selectins were identified so far, all bearing a CRD domain at the N-terminal end, an epidermal growth factor-like domain, various numbers of complement-binding repeats at the stalk region, and a short cytoplasmic tail. Selectins recognize endogenous sugar structures and mediate the adhesion between leukocytes and endothelia. Natural killer cell receptors

constitute Group V C-type lectins. The molecules in this group are typically comprised of one CRD domain at the C-terminal end, a neck region, transmembrane domain, and a cytoplasmic domain with a signaling motif. Some receptors form hetero or homo-dimers through a disulfide bond at the neck region. The functions of these NK cell receptors involve distinguishing self or non-self ligands, and eliciting consequent activating or inhibitory signals. The endocytic receptors including macrophage mannose receptor, PLA2 receptor, DEC-205 and Endo 180 were identified as group VI, because they are type I transmembrane receptors. Members in this group bear multiple, usually 8-10, CRD domains at their N-terminus. Group VII molecules are distinct in that they only have a single CRD domain, with no other accessory regions. Molecules in this group are present more widely in invertebrates and lower vertebrates. In humans, only two pancreas-related group VII molecules were isolated so far, with unknown ligands and functions. We are interested in the C-type lectins expressed on antigen presenting cells, which are mainly group V and VI members (Figure 1-1).

1.1.3 C-type lectin like receptors

Some lectins possess part, or the entire classical CRD domain, but fail to recognize sugar and are functionally independent of calcium ions. Drickamer defined molecules of this kind as members of the C-type lectin-like superfamily and their special CRD domains as C-type lectin-like domains (CTLD)

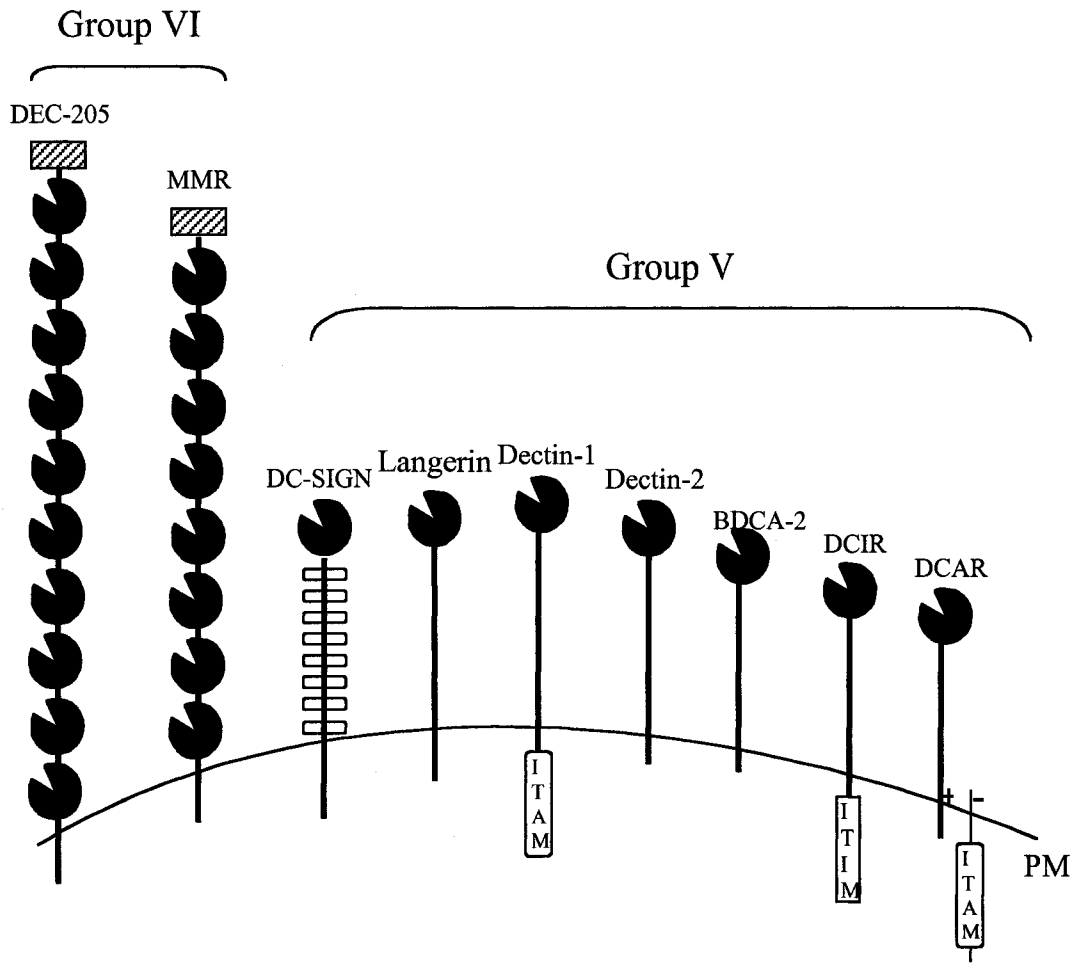


Figure 1-1. Schematic structure of APC expressed C-type lectin receptors. The CRD domains were indicated by black nicked balls. Hatched boxes show the fibronectin regions of type I receptor MMR and DEC-205. The tandem repeats found in the stalk region of DC-SIGN were shown by white boxes. The immunoreceptor tyrosine based inhibitory motif (ITIM) and immunoreceptor tyrosine-based activation motif (ITAM) are shown as boxed. This figure is adapted from Figdor. *et al* (2002).

(Drickamer, 1999). Due to divergence, this superfamily contains many branches, with only a small portion containing conserved CRD. For example, although some NK cell lectin receptors have a CRD domain, like the NKG2 family, they recognize proteins such as MHC molecules, instead of sugars. Such features were also seen in antifreeze proteins and lectin-like oxidized low-density-lipoprotein receptor-1 (LOX-1), which were the receptors for ice and lipids, respectively (Drickamer, 1999). Some molecules display neither similar sequences nor binding features to CRD members. However, the spatial structure of these molecules showed a folding format close to that found in classical C-type lectins. It is believed that convergence contributed to this topological similarity, yet evidence is still lacking (Drickamer, 1999).

1.2 Antigen presenting cells (APC) and C-type lectins expressed on their suface

Innate immunity plays a very important role in the host defense to pathogens. In vertebrates, the cells of innate immunity not only constitute the first line of defense, but also initiate and direct acquired immunity. At the interface of innate and acquired immunity, antigen presenting cells (APC) are key to presenting components of pathogens to lymphocytes, triggering specific immune responses and generating memory cells that can directly fight against pathogens when they encounter them the second time. Currently, three types of cells, including dendritic cells (DC), macrophages and B cells, are considered

professional APC. Among them, dendritic cells are the most potent at antigen presentation and T cell activation. In addition to antigen presentation, dendritic cells also function in mediating tolerance and autoimmunity.

1.2.1 Dendritic cells (DC) as professional APC

A great variety of DC populations have been identified so far. Based on different anatomical locations, DC are subcategorized into skin epidermal DC (termed Langerhans cells), dermal/interstitial cells (intDC), splenic marginal DCs (follicular DC), T zone interdigitating cells, germinal-center DCs, thymic DCs, liver DCs, and blood DCs (Banchereau *et al.*, 2000).

The precursors of DC migrate from bone marrow to peripheral organs, acting as sentinels capturing external antigen or danger signals. Once an antigen appears, DC will accumulate at the antigen deposition site. Chemokines guide the recruitment process through binding receptors expressed on DC.

DC takes up pathogens through three different ways: receptor-mediated endocytosis (Mahnke *et al.*, 2000), phagocytosis (Fanger *et al.*, 1996) and macropinocytosis (Sallusto *et al.*, 1995). Depending on the entry sites and properties, the antigens are presented by DC through two distinct pathways, mediated by MHC class I and II molecules. MHC class I or II molecules present the processed antigens to CD8 or CD4 T cells, respectively.

At the same time, pathogens or inflammatory factors trigger a phenotypic and functional transformation of DC, a process called maturation. During this stage, DC will undergo a series of changes. First, morphologically, many long dendrites extend out, which help with the searching and selection of T cells. Morphology changes facilitate antigen presentation and DC movement. Secondly, a number of changes occur regarding the presence of surface markers and receptors. The up-regulation of MHC molecules (Winzler *et al.*, 1997), co-stimulatory molecules (such as CD40, CD80 and CD86)(Inaba *et al.*, 1994), as well as adhesion molecules (Winzler *et al.*, 1997), are observed. This upregulation is critical for the recognition and activation of lymphocytes. In contrast, the maturation of the DC attenuates its capacity for antigen capture and internalization, with down-regulated expression of endocytic/phagocytic receptors. Thirdly, the components in cellular compartments also change.

The migration of activated DC is mediated by the binding of chemokines to their corresponding receptors. The activated DC up-regulates the expression of CCR7 (Yanagihara *et al.*, 1998), which is a receptor of CCL19 and CCL21, both present on lymphatic vessels (Baekkevold *et al.*, 2001; Gunn *et al.*, 1998). The binding of the chemokine receptors (e.g. CCR7) and their ligands (e.g. CCL19 and CCL21) lead to the migration of DC along the lymphatic stream (Cyster, 1999). Both CCL19 and CCL21 can also promote the adhesion and chemotaxis of naive T cells and thus play a key role in mediating the attachment of naive T cells to mature DC (Gunn *et al.*, 1998; Kaiser *et al.*, 2005).

The reaction of DC to pathogens is mainly dependent on the expression of two families of receptors: the Toll like receptor and C-type lectin families, both of which are pattern recognition receptors (PRRs) and trigger the downstream immune responses. This introduction will focus on reviewing C-type lectins on APC, which are the topic of this thesis.

1.1.2 Type I C-type lectin receptors on APC

Mannose receptors, together with another three C-type lectins, Endo180, PLA2R and DEC-205, have similar structures and thus constitute a mannose receptor family (EastIsacke, 2002a). As type I transmembrane proteins with the C-terminus within the cell, all family members have 1) a short cytoplasmic domain containing signaling motifs; 2) multiple (usually 8-10) CRD domains at the extracellular region; 3) a fibronectin type II (FNII) repeat motif outside of the CRD domains and 4) a cysteine rich region at the N-termini. Although sharing a similar structure, they have different expression spectra of ligand binding specificities and functions. For example, PLA2R was shown to bind and internalize various soluble forms of phospholipase A2, and participate in inflammation induced endotoxic shock (Hanasaki *et al.*, 1997). On the other hand, Endo 180 can bind to both mannose containing sugars (East *et al.*, 2002b) and collagen (Thomas *et al.*, 2005), functioning in the regulation of cell surface molecules and extracellular matrix remodeling.

1.1.2.1 Macrophage mannose receptor (MMR or MR)

Among the three sugar binding MR family members mentioned above, macrophage MR was identified first, and was the most extensively studied. MR is widely distributed on tissue macrophages, immature DC, lymphatic and hepatic endothelium, kidney mesangial cells, peri-vascular microglia, retinal pigment epithelium and Kaposi s sarcoma cells (Linehan *et al.*, 1999; Sallusto and Lanzavecchia, 1994; Uccini *et al.*, 1997).

At the genomic level, two transcription factors, the ubiquitous factor Sp1 and the lymphoid/myeloid factor PU.1, regulate the expression of MR synergistically (Eichbaum *et al.*, 1997). Pathogens and cytokines can also modulate MR expression. For example, Nef, an HIV-1 derived protein, is a post-transcription factor that reduced the surface expression of MR and therefore decreased the reactivity of host APC (Vigerust *et al.*, 2005). In terms of the regulatory effects of cytokines, various molecules collaborate and adjust the expression of MR according to the functional state of MR expressing cells. For instance, IL-4 and IL-10 enhanced the surface expression of MR (Martinez-Pomares *et al.*, 2003), whereas IFN-γ functions in balancing the phagocytic and antigen presentation capability of MR (Marodi *et al.*, 1993).

The ligands of the CRD of MR include terminal or branched mannose, fucose and N-acetylglucosamine, none of which are common in mammalian glycoproteins (Largent *et al.*, 1984). This feature is crucial for MR to recognize

exogenous non-self pathogens. Structural analysis found that only two of the eight CRD domains of MR, CRD4 and to a lesser extent CRD5, contained the required residues for sugar and calcium binding. Only having weak sugar binding ability, the rest of the CRD domains form a cluster and contribute to the high affinity of MR to multivalent glyco-conjugates (Taylor *et al.*, 1992). In addition to the CRD domain, the cysteine rich region and fibronectin II (FNII) repeat motif also possess ligand-binding ability. The cysteine rich region at the N-terminal end binds to sulfated carbohydrates, such as pituitary hormones, luthopin and thytropin, and regulate the level of circulating hormones (Leteux *et al.*, 2000). In contrast, the FNII domain recognizes collagen found in connective tissue, and might play a role in collagen clearance, or cell-matrix interaction (Napper *et al.*, 2006).

MR plays important roles in innate immunity, by recognizing a wide range of pathogens, including Gram positive and Gram negative bacteria, fungi and parasites. However, some pathogens take advantage of the binding and uptake specificity of MR, making the receptor a target to enter the body and escape from immunosurveillance. For instance, *Mycobacterium tuberculosis (M. tuberculosis)*, the pathogen of human tuberculosis, contains a lipoarabinomannan (LAM) structure in its cell wall, which displays a prominent terminal mannose cap in virulent strains, but not in attenuated strains. The recognition of MR to mannosylated LAM (ManLAM) mediates the binding and uptake of bacteria by macrophages. Research done by Kang *et al.* (2005) revealed that *M. tuberculosis*

limits the fusion of phagosomes and lysosomes after entering the phagosomes (Kang *et al.*, 2005). In addition to *M. tuberculosis*, an opportunistic parasite *Pneumocystis carinii (P. carinii)* can induce the shedding of MR from the macrophage surface, and therefore interfere with the normal function of the receptor (Fraser *et al.*, 2000).

In contrast to macrophage MR that takes up various pathogens through phagocytosis, dendritic cell MR mediates antigen uptake and directs the antigen to the MHC-II compartment through endocytosis. It was observed that immature DC presents mannosylated antigens far more efficiently than non-mannosylated antigens. In contrast, the expression level and endocytotic capability of MR were substantially reduced in mature DC (Sallusto *et al.*, 1995).

Other than recognizing non-self antigens, MR is also implicated in toxin clearance and lymphocyte trafficking through binding with self-ligands. The recognition and engulfing of lysosomal hydrolases and mannose-bearing serum glycoproteins mediated by MR facilitate the elimination of the toxic agents and helps to keep homeostasis of the body (Biessen *et al.*, 1997; Lee *et al.*, 2002).

1.1.1.2 DEC-205

First identified in mouse, DEC-205 is highly expressed on murine immature dendritic cells and thymic epithelial cells (Jiang *et al.*, 1995). *In situ* hybridization detected DEC-205 widely expressed in lymphoid tissues, especially

the T cell regions in spleen as well as lymph nodes, bone marrow stroma, mucosal tissues such as the epithelia of intestine and lung, and brain capillaries (Witmer-Pack *et al.*, 1995). Human DEC-205 showed 77% sequence similarity to the mouse homologue at the amino acid level and had a similar expression profile (Kato *et al.*, 1998). The extracellular region of DEC-205 is similar to MR, carrying multiple CRD domains and a cysteine-rich repeat motif. However, none of the 10 CRD bear the conserved carbohydrate binding residues and its cysteine-rich motif does not bind to sulfated sugars as in MR. These structural and functional differences suggests that the ligands of DEC-205 might not be sugar. Two functional motifs, the FSSVRY sequence and the acidic EDE motif, are localized in the cytoplasmic region of DEC-205. The former one resembles the coated pit localization sequence that is also found in MR, inferring its role in antigen endocytosis, whereas the latter is believed to mediate the intracellular trafficking of antigen (Mahnke *et al.*, 2000).

Although the specific ligands are still unknown, DEC-205 functions in a similar way as MR, demonstrating a strong capability for antigen uptake and presentation in APC. However, DEC-205 is unique in the way that it targets antigen to and then recycles from the late endosome, colocalized with MHC-II and LAMP-1(Mahnke *et al.*, 2000). The EDE motif in its cytoplasmic domain is the key structure that mediates the late endosomal targeting. This novel pathway mediated by DEC-205 greatly improved the efficiency of antigen presentation to CD4 T cells, even when the ligands are in low dose. Some subsets of DC may

also utilize DEC-205 to capture antigen and present it to CD8 T cells, through MHC-I molecules (Bonifaz *et al.*, 2004). However, this process is dependent on the status of DC, which means that only the mature DC is able to elicit CD8 T cell immunity, while the DC at its steady state leads to tolerance (Bonifaz *et al.*, 2004). This finding suggested that DEC-205 could be a useful tool in the process of anti-tumor or anti-virus T cell vaccination.

Other than antigen uptake and initiation of adaptive immunity, DEC-205 has another distinct function. When expressed on thymic cortical epithelial cells, DEC-205 was shown to bind tightly to the thymocytes that express apoptotic signals. This specific binding might contribute to the elimination of apoptotic thymocytes (Small and Kraal, 2003).

1.1.3 Macrophage specific C-type lectins

Macrophages are derived from the common myeloid progenitor and blood circulating monocytes. As a phagocyte, the macrophage is critical in host innate immunity, homeostasis maintenance and tissue remodeling. Depending on location, macrophage subpopulations are heterogeneous, displaying diverse phenotypes. A large number of macrophages are localized at pathogen entrance sites and well equipped with various receptors to recognize specific antigens. Generally, four families of receptors are usually present on macrophages, including scavenger receptors, integrins, Ig superfamily receptors and C-type

lectin superfamily receptors. Type I C-type lectins on macrophages have been discussed in the above section. Here, we describe two novel type II C-type lectin receptors, MCL and Mincle, both of which are predominantly expressed on macrophages.

1.1.3.1 MCL

Murine macrophage-restricted C-type lectin (mMCL), also called CLECSF8 in humans, is a novel macrophage specific type II C-type lectin (Arce *et al.*, 2004). The predicted polypeptide of murine MCL is 219 amino acids in length, containing a 20 aa cytoplasmic domain, 20 aa transmembrane domain, 36 aa neck region and 143 aa CRD domain (Balch *et al.*, 2002). Sharing 74% homology with its mouse homologue, the human MCL protein is four amino acids shorter, with similar amino acid distribution in each domain. The CRD domains of murine and human MCL are highly conserved, except for a pair of discrepancies at the carbohydrate recognition site. Unlike the usual EPN or QPD motifs, murine MCL has an ESN motif, with a serine, instead of a conserved proline in the middle (Balch *et al.*, 1998). An asparagine next to the serine in mMCL was substituted by an aspartate in hMCL (Arce *et al.*, 2004). Since the ligand of MCL is unclear, the functional significance of the different sugar-binding motif in mouse and human is still unknown. As in other APC-expressed C-type lectins, alternative splicing was found in hMCL. Interestingly, this spliced

form skips Exon IV, resulting in a molecule lacking part of the CRD domain (Arce *et al.*, 2004). In both humans and mice, the expression of MCL was restricted to macrophages and macrophage-rich organs, including bone marrow and peripheral blood leukocytes (Balch *et al.*, 2002; Arce *et al.*, 2004). The expression was upregulated when the cells were exposed to infection or various cytokines, including IL-6, TNF-α, IFN-γand IL-10. An internalization assay showed that cross-linking of membrane expressed MCL resulted in endocytosis, which suggested the potential role of MCL in antigen presentation (Arce *et al.*, 2004).

1.1.1.2 Mincle

Mincle is another macrophage-restricted infection-inducible C-type lectin, and is similar to MCL in terms of the size of polypeptide, structure and NKC-proximal localization. The expression of Mincle was detected in peritoneal macrophage two hours after LPS infection (Matsumoto *et al.*, 1999). Promoter analysis revealed a sequence motif at position -66 to -58 responsible for the activation by NF-IL6. These findings provided useful information on the mechanism of the regulation and function of Mincle. Variations at the 3 UTR poly-A site, rather than alternative splicing, resulted in two transcripts of Mincle in mouse (Matsumoto *et al.*, 1999).

1.1.4 Langerin

Langerin (CD207) is another type II C-type lectin, which is preferentially expressed on Langerhans cells (dermal DC) (Valladeau *et al.*, 2000). Langerin has a typical CRD domain, featuring a QPD motif. A proline-rich motif (WPREPPP) was found in the cytoplasmic domain of its mouse homologue, which might be implicated in protein interactions and cytoplasmic signaling (Valladeau *et al.*, 2002). Unique from most other APC expressed C-type lectins that lie in the NK gene complex or proximal to DC-SIGN, human Langerin was localized on chromosome 2 (Valladeau *et al.*, 2002). One alternative splicing isoform that lacked the stalk region encoding exon 3 was identified in mouse (Riedl *et al.*, 2003).

Subcellularly, Langerin accumulates in a special structure called the cytomembrane sandwiching structure, which was considered as the precursor of the Birbeck granule (BG). After binding to a specific mAb directed against the CRD domain, Langerin facilitated the rapid internalization of the ligand to the Birbeck granule, but not MHC-II-rich compartments (Valladeau *et al.*, 2002). Expression of Langerin on other cell types by transfection elicited similar reactions and induced BG formation. Furthermore, it was found that one amino acid mutation (F224-L) in the Langerin CRD domain could lead to a change of the structure of BG substantially (Valladeau *et al.*, 2002). This relationship between Langerin and BG structure was further confirmed by Verdijk *et*

al.(2005), who revealed that another point mutation (W264) in the CRD domain resulted in a naturally occurring BG absence in a healthy Caucasian man (Verdijk *et al.*, 2005). These results inferred that conformational changes in the extracellular domain of Langerin might interfere with the regular formation and structure of BG.

The expression of Langerin is not restricted to Langerhans cells. Other than skin where Langerhans cells accumulate, it was also detected in spleen, mesenteric lymph node and thymus, where no Langerhans cells reside (Valladeau *et al.*, 2002). Langerin was detected in lymph nodes and thymic medulla in cells displaying DC morphology, which is in accordance with its tissue distribution. Recent research conducted by Douillard and colleagues detected two different subsets of Langerin+ cells in lymph nodes, with one subset showing Langerhans cell features, whereas the second subset does not (Douillard *et al.*, 2005). In contrast, the Langerin+ cells located in spleen and thymus were predominantly subset 2. Taken together, Langerin is not an exclusive Langerhans cell marker, but its role in lymphoid DC is yet to be elucidated.

1.1.5 DC-SIGN

DC-SIGN is an APC specific C-type lectin that functions in antigen recognition and cell-cell interaction. Unlike other APC lectins, DC-SIGN has a short cytoplasmic domain and a relatively long stalk region, which bears seven

complete and one partial tandem repeats that mediate tetramerization (Mitchell *et al.*, 2001). DC-SIGN has a high expression level on DC derived from both monocytes and bone marrow, as well as dermal DC (Geijtenbeek *et al.*, 2000c). *In situ* hybridization showed that DC-SIGN was abundant in cells located in the T cell zone of lymph nodes, tonsil, spleen and dermal skin, as well as numerous mucosal tissues (Geijtenbeek *et al.*, 2000c). However, recent research using a new panel of monoclonal antibodies revealed that DC-SIGN was not DC-restricted, but was also expressed on macrophages present in lymph node medulla (Granelli-Piperno *et al.*, 2005). The expression of DC-SIGN on tissue macrophages is up-regulated by various stimuli, such as pathogen infections.

DC-SIGN was first identified as a ligand of adhesion molecule ICAM-3, which is highly expressed on resting T cells (Geijtenbeek *et al.*, 2000c). The calcium-dependent binding of dendritic cell DC-SIGN to T cell ICAM-3 provides close interaction for TCR to look for its ligands presented by MHC molecules (Geijtenbeek *et al.*, 2000c). In addition to mediating DC-T cell interaction, DC-SIGN also functions in DC trafficking through its interaction with another surface adhesion molecule, ICAM-2. ICAM-2 is richly expressed on the endothelium of blood and lymphatic vessels. The binding of DC-expressed DC-SIGN with ICAM-2 mediates the migration of DC from peripheral locations to blood and lymph vessels (Geijtenbeek *et al.*, 2000a). DC-SIGN also mediates the interaction between DC and neutrophils, through binding with a special Lewis (x) sugar displayed on Mac-1 (van Gisbergen *et al.*, 2005) and CEACAM-1 on neutrophils.

Apart from mediating cellular communication, DC-SIGN was shown to take an active part in pathogen recognition. It specifically binds to the high mannose N-linked oligosacharide structures that appear on the surface of viruses, bacteria, parasites, fungi, as well as endogenous antigens. For instance, DC-SIGN was highly expressed on DCs that were located at HIV-1 entry sites. Although DC-SIGN can t mediate the entry of HIV, it can capture the virus and facilitate HIV infection of permissive cell, including the cells that express CD4 or chemokines (CCR5) (Geijtenbeek *et al.*, 2000b). The specific binding of DC-SIGN to HIV is mediated by its high affinity to the surface protein gp120, in a calcium dependent manner. In addition, high expression of DC-SIGN was detected in some placenta cells, which carry CD4 and CCR receptors and thus are vulnerable to HIV infection (Geijtenbeek *et al.*, 2000b). This finding might shed light on the mechanism of vertical transmission of HIV infection. On one hand, the binding of DC-SIGN to gp120 facilitates the escape and infection of HIV. While on the other hand, it also provides possible routes to control the further transmission and infection of HIV by either blocking binding or changing the microenvironment of internalized viruses.

In addition to HIV, Cytomegalovirus (Halary *et al.*, 2002), hepatitis C virus (Barth *et al.*, 2005; Ludwig *et al.*, 2004), Dengue virus (Tassaneetrithep *et al.*, 2003), human herpesvirus 8 (Rappocciolo *et al.*, 2006) and Ebola virus (Alvarez *et al.*, 2002) can also take advantage of DC-SIGN to transmit to their susceptible cells.

1.1.6 CLEC-1 and CLEC-2

Both human CLEC-1 and CLEC-2 are type II C-type lectins, carrying non-classical CRD domains. Their CRD domains lack most of the crucial Calcium/sugar binding residues, indicating that the possible ligands of both receptors might be proteins or lipids, instead of sugar. Extra cysteine residues were found in the CRD domains of both receptors, which were predicted to mediate the formation of homo- or hetero- intrachain dimers. A tyrosine residue was found in the cytoplasmic domain of CLEC-1 within the sequence of Y-S-S-T for CLEC-1, which might be a potential tyrosine-based motif. In CLEC-2, a cytoplasmic tyrosine residue was also detected within the sequence of D-x-Y-x-x-L, which might function as an endocytic signal (Colonna *et al.*, 2000). Phylogenetic analysis of the CRD domain revealed that CLEC-1 is closer to CLEC-2, and NK C-type lectins such as NKG2s and CD94, than to other APC expressed lectins (Colonna *et al.*, 2000).

CLEC-1 is preferentially expressed in the APC abundant mucosal tissues, including lung, small intestine and placenta. RT-PCR also revealed that CLEC-1 was enriched on stimulated dendritic cells and epithelial cells (Colonna *et al.*, 2000), and was modestly expressed on monocytes (Sobanov *et al.*, 2001). Interestingly, it was found that both the entire molecule of CLEC-1 (Sobanov *et al.*, 2001) and its cytoplasmic domain truncated mutant, failed to be expressed on the surface of transfected cells (Colonna *et al.*, 2000). This observation suggests

that CLEC-1 may need another molecule to help with its membrane expression (Marco *et al.*, 2000). In contrast to CLEC-1, the expression of CLEC-2 was detected on the surface of transfected cells. Northern blot analysis showed that CLEC-2 was selectively expressed in liver and myeloid derived cells (Colonna Marco, 2000).

1.1.7 Dectin-1

The Dectin-1 transcript encoded a 244 amino acid polypeptide, including a 44 aa cytoplasmic domain, 45 aa transmembrane domain, and a CRD containing extracellular domain (Ariizumi *et al.*, 2000b). In contrast to other classical C-type lectins, the CRD domain of Dectin-1 lacks two important calcium/sugar binding positions (Ariizumi *et al.*, 2000b). Strikingly, an YTQL motif, which is a putative ITAM motif (YXXL), was found in its cytoplasmic domain. Alternative splicing leads to one major and six minor truncated isoforms of Dectin-1 in humans. Among them, the major isoform lacked exon 3, which encodes the stalk region (Willment *et al.*, 2001).

Displaying less typical APC C-type lectin features, together with its chromosomal localization at the NK gene complex, Dectin-1 was considered an unusual hybrid of C-type lectins and NK C-type lectins. Although it was originally found in dendritic cells and was termed DC-associated C-type lectin-1 (Dectin-1), Dectin-1 is expressed on a wide spectrum of cells. An extensive study

of Dectin-1 expression done by Reid *et al.* (2004) showed that Dectin-1 was highly expressed in macrophage/monocytes and dendritic cells, especially those within the tissues that are highly exposed to pathogens, such as lung, gut and spleen (Reid *et al.*, 2004). Moreover, the expression level of Dectin-1 on macrophages and DC are subject to the changing of cell activation/maturation status, as well as the regulation of cytokines (Willment *et al.*, 2003). The expression of Dectin-1 has also been found on neutrophils and a subpopulation of T cells (Taylor *et al.*, 2002).

Although lacking the conserved C-type lectin calcium/sugar binding motifs, Dectin-1 was known to bind to β-glucan, which is commonly found in the fungal cell wall (Brown, 2001). Mutation analysis has revealed that a unique WIH motif (aa221-223), instead of the typical E(/Q)PD(/N) motif in other C-type lectins, was crucial for β-glucan binding (Adachi *et al.*, 2004). To evade β-glucan recognition by Dectin-1, some pathogenic yeast developed an evasion mechanism by their switching morphology from a β-glucan-enriched cell wall to a filamentous form. In addition to pathogen recognition, both soluble and glycosylated Dectin-1 display binding specificity to activated T cells through unknown ligands. The binding is trypsin-sensitive, indicating that the potential ligand might be protein or glycoprotein (Ariizumi *et al.*, 2000b).

After binding to β-glucan containing antigens, APC undergo phagocytosis and produce cytokines, such as IL-2, -10, -12 (Rogers *et al.*, 2005). There are two pathways mediating downstream signal transduction, one is the MyD88-

dependent pathway that is activated by Toll like receptor-2, the other is independent of MyD88 and TLR and is brought about by Dectin-1 activation. An ITAM motif found in the cytoplasmic domain of Dectin-1 confers the signaling capability. Upon ligand binding, the phosphorylation of the membrane-proximal tyrosine within the ITAM motif is indispensable for the recruitment of Syk kinases, which will elicit downstream responses. The Dectin-1/ITAM pathway itself is sufficient for IL-10 production (Rogers *et al.*, 2005). However, studies have shown that the synthesis of other cytokines may require the collaboration between both pathways through unknown mechanisms (Gantner *et al.*, 2003).

1.1.8 Dectin-2

Dectin-2, originally reported as NKCL (Fernandes *et al.*, 1999), was isolated as an over-expressed gene from spleen cells of an experimental mouse model of aggressive acute phase leukemia. Carrying a conserved CRD domain, Dectin-2 shows high similarities to other type II APC lectins including DCIR and MGL (Ariizumi *et al.*, 2000a). So far, Dectin-2 has been identified in humans, mice and cows, with conservation of receptor structure, gene organization and the EPN motif (Ariizumi *et al.*, 2000a; Bonkobara *et al.*, 2006)]. The function of this motif in signal transduction is yet to be elucidated. In mouse, two truncated Dectin-2 isoforms resulting from alternative splicing were identified. One isoform lacked the entire exon 3, which encodes the neck region, whilst the other had

partial deletions at exon 5 and 6, skipping part of the CRD domain (Ariizumi *et al.*, 2000a). Similarly, the exon 3 deletion form was also detected in bovine Dectin-2 (Bonkobara *et al.*, 2006). In human Dectin-2, an exon 2 deletion version was detected, lacking the entire transmembrane domain and part of the intracellular domain (Gavino *et al.*, 2005).

Dectin-2 was most abundantly expressed in spleen and thymus (Ariizumi *et al.*, 2000a; Taylor *et al.*, 2005). Lung and bone marrow expression were also reported, which was not in agreement between different research groups (Fernandes *et al.*, 1999; Kanazawa *et al.*, 2004a). Cellular expression was restricted to myeloid derived cell populations including monocytes, tissue macrophages (Taylor *et al.*, 2005), as well as skin resident dendritic cells (i.e. Langerhans cells) (Ariizumi *et al.*, 2000a). The expression was even higher when monocytes were activated by various inflammatory stimuli, especially chronic inflammation (Taylor *et al.*, 2005). The analysis of the 5 flanking region of the Dectin-2 gene has revealed two possible enhancers, two repressors and a minimal promoter, in the region from -2741 to transcription initiation site (Bonkobara *et al.*, 2001). Interestingly, luciferase assay suggested that Langerhans cells contain the most Dectin-2 promoter specific nuclear proteins, and displays the highest luciferase expression among leukocytes. It was speculated that this Dectin-2 promoter could be used as a tool to develop Langerhans cell targeted gene therapy (Morita *et al.*, 2001).

The function of Dectin-2 was studied from two aspects: its ligand binding specificity and its role in ultraviolet -induced tolerance. It has been recently discovered that the CRD domain of Dectin-2 specifically recognized high mannose sugar, which is found on the surface of yeast and bacteria bearing complex polysaccharides (McGreal *et al.*, 2006). The binding specificity and efficiency of Dectin-2 suggests its roles in pathogen recognition and immuno-surveillance, which are also inferred by its tissue and cell distribution. Dectin-2 is also involved in the process of ultraviolet B -induced immune suppression (Aragane *et al.*, 2003). Generally, UVB is considered a hazard to animal life in that it induces skin aging and cancer. It also inhibits the cellular immune system via regulatory T cells resulting in immune supression. In the study of Aragane *et al.* (2003), a recombinant soluble Dectin-2 was used to block the function of normal Dectin-2. They found the soluble Dectin-2 could prevent induction of tolerance by UV light, indicating that Dectin-2 was involved in the UV induced immuno- suppression (Aragane *et al.*, 2003).

1.1.9 BDCA-2/DLEC

BDCA-2, also known as DLEC and CLECSF7, was first identified as a specific surface marker for human plasmacytoid dendritic cells (PDC) (Dzionek *et al.*, 2000). It was also recognized as a type II C-type lectin receptor that plays a role in antigen uptake and presentation. The CRD domain of BDCA-2 shows high

homology to Dectin-2 and DCIR. An EPN motif at the conserved position suggests that its binding specificity is for the mannose-type carbohydrates. Interestingly, BDCA-2 has a large number of alternatively spliced transcripts, which lack exon 2, exon 3, exon 4, exon 2 and 3, or exon 2 and 5 (Dzionek *et al.*, 2001). The expression of BDCA-2 is confined to plasmacytoid DC, but not all PDC-containing tissues. BDCA-2 was detected on PDCs located in the extrafollicular area of tonsil that is enriched for T cells, in reactive lymph nodes, in thymus and in the lymphoid aggregates that usually appear in testis seminoma (Dzionek *et al.*, 2001).

Although the ligand of BDCA-2 is still unknown, it was established that BDCA-2 plays an active roles in antigen uptake and presentation. The binding of BDCA-2 with its specific antibody elicited both protein tyrosine phosphorylation and intracellular calcium mobilization, which occurs in an Src-family tyrosine kinase dependent manner (Dzionek *et al.*, 2001). It was observed that ligand internalization was followed by the transition of the ligand to antigen processing compartments where class-II loading takes place. Containing no endocytic motif like that found in DC-SIGN and MMR, BDCA-2 makes use of an acidic triad (EEE) in its cytoplasmic domain to deliver its ligand to late endosomes (Dzionek *et al.*, 2001).

PDC are natural producers of IFN-α and β, especially when stimulated by various pathogens or some IFN-inducing factors. However, the binding of BDCA-2 to a specific monoclonal antibody can block the induction of IFNα/β (Dzionek

et al., 2001). Blocking these type I interferons results in skewing the T cell response away from a Th1 antiviral immune response. Viruses, like influenza, may evade immune responses by triggering BDCA-2 on PDC.

1.1.10 Paired APC type II receptors

1.1.10.1 ITIM and ITAM motifs

Originally found in NK cell and B cells, paired immunoreceptors utilize immunoreceptor tyrosine based inhibitory motif (ITIM) and immunoreceptor tyrosine-based activation motif (ITAM) to perform opposite functions and balance inhibitory and activation signals. The binding of antigen/ligands to their specific receptors on lymphocytes triggers conformation changes and clustering of receptors, which in turn recruits cytoplasmic enzymes and initiates downstream signal transduction in the cell. After recruitment, the Src family protein tyrosine kinases (PTKs) phosphorylate tyrosine residues situated either on the cytoplasmic domain of the receptor or on receptor-associated adaptor molecules. The phosphotyrosines then bind with Src homology 2 (SH2) domain containing PTKs and elicit subsequent activating signals.

The ITAM motif was identified in the cytoplasmic domain of B cell receptor (BCR) and T cell receptor (TCR) associated proteins (Isakov, 1997). The canonical ITAM motif contains two tyrosine residues separated by 9-12 amino

acids. Upon activation of the receptor, the two tyrosines in the ITAM motif are phosphorylated by Src family PTK, and bind with SH2 containing PTKs, which trigger the downstream signals (Isakov, 1997).

Within the murine FcγIIB, a receptor that can inhibit B cell activation, a 6 amino acid sequence in the intracellular domain was found to mediate inhibitory functions (Daeron *et al.*, 1995). Subsequent studies revealed that this 6 aa sequence (I/VxYxxL) was highly conserved in all of the inhibitory receptors whose inhibitory functions depended on the ITAM motif, and the six residues were named ITIM. Following co-aggregation with other ITAM-bearing receptors, the tyrosine residue on FcγIIB is phosphorylated, followed by the recruitment of an SH2 domain-containing PTK. The phosphatase then dephosphorylates downstream effectors and blocks cell signaling (Coggeshall, 1998).

1.1.1.2 DCIR

Dendritic cell immunoreceptor (DCIR), also reported as LLIR and CLECSF6, is the first reported APC expressed C-type lectin that contains an ITIM motif in its cytoplasmic domain. Located at the telomeric end of the NKC, the DCIR gene is close to Dectin-2, Mincle and MCL (Kanazawa *et al.*, 2003). So far, DCIR homologues have been identified in human, mouse and rat. The amino acid sequence of mouse DCIR shares 65% and 74% identity to that of human and rat, respectively (Bates *et al.*, 1999; Flornes *et al.*, 2004). Due to

alternative splicing, three alternate DCIR isoforms have been found in human myeloid cell lines and neutrophils (Huang *et al.*, 2001; Richard *et al.*, 2002). These three isoforms are missing exon 2 that encodes the transmembrane domain, exon 3 that encodes the stalk region, or both.

The expression profile of DCIR in human and rodents are consistent. The tissue distribution of DCIR includes: a) lymphoid organs, for instance, lymph nodes and tonsil, b) mucosal tissues including lung and intestine (Kanazawa *et al.*, 2002), and c) hematopoietic tissues such as peripheral blood, spleen and bone marrow (Bates *et al.*, 1999). In regard to the cellular expression, DCIR has been detected in antigen presenting cells, such as dendritic cells, macrophages, and B cells, indicating its potential role in antigen recognition and processing (Bates *et al.*, 1999; Huang *et al.*, 2001). Furthermore, DCIR has a much higher expression level in the interstitial/dermal DC subsets than in the epidermal DC (like LC) (Bates *et al.*, 1999). The expression of DCIR is downregulated during DC maturation as shown by the increased expression of co-stimulatory receptors and MHC and DC migration to secondary lymphoid organs (Bates *et al.*, 1999). Richard and colleagues also detected that the expression of DCIR in neutrophils was activated by GM-CSF (Richard *et al.*, 2003).

Analysis of the predicted amino acid sequence of human DCIR has revealed that the extracellular domain contains an EPS motif, which is conserved in the CRD domain in DCIR and DCAR and is the putative sugar-binding site. The six cysteines that contribute to CRD folding are conserved. In the

cytoplasmic domain, an ITYAEV motif was found, which perfectly matches the consensus sequence of the ITIM (Bates *et al.*, 1999).

In order to characterize the inhibitory function of DCIR, a chimeric receptor containing the extracellular domain of tagged FcγIIB and the intracellular domain of DCIR was constructed and expressed in a variant of mouse B cell, which lacks cell surface Fc receptors. After stimulation by $F(ab)_2$ or intact antimouse IgG antibodies, DCIR-FcR absolutely blocked the B cell receptor (BCR)-mediated calcium mobilization on co-ligation with BCR at the same level as the positive control (CD72-FcR, CD72 is a known ITIM-bearing receptor that can negatively regulate signaling though BCR). In contrast, the ITIM mutant DCIR did not show this inhibition (Kanazawa *et al.*, 2002). This construct was used to test whether DCIR had an inhibitory effect on BCR-mediated tyrosine phosphorylation (Kanazawa *et al.*, 2002). The result illustrated that after co-ligation with BCR, the DCIR-FcR reduced cellular tyrosine phosphorylation significantly. Meanwhile, the ITIM tyrosine mutant showed no effect. Therefore, DCIR had the ability to inhibit cellular signaling through its ITIM motif. So far, there is no definite conclusion on the mechanism of the inhibitory effect of the DCIR ITIM. The binding of DCIR ITIM to phosphatase SHP-1 and -2 has been reported (Huang *et al.*, 2001; Richard *et al.*, 2006).

1.1.1.3 DCAR

Dendritic cell immunoactivating receptor (DCAR) was discovered while searching for activating counterparts of DCIR. The amino acid sequence of DCAR is highly similar to DCIR in the CRD domain, but very different in the transmembrane and cytoplasmic domain. The cytoplasmic domain of DCAR is much shorter than DCIR and lacks the ITIM motif. Instead, it contains an arginine, a conserved charged amino acid residue that is also found in other two short-tail C-type lectins, BDCA-2 and Dectin-2 (Kanazawa *et al.*, 2003). This charged amino acid might contribute to its association with ITAM-containing adaptor proteins, which convey activation signals. It was reported that the FcR γ chain could not only promote the surface expression of DCAR, but also act as the adaptor protein. Through the association of the ITAM to FcR γ chain, activating effects are observed and are accompanied by both increased calcium mobilization and increased kinase cascade activation (Kanazawa *et al.*, 2003).

Two DCARs were found in the rat and mouse (Flornes *et al.*, 2004). Among them, rat DCAR2 is a pseudogene, which contains exon 2 and 3 only. No human DCAR homologue has been detected so far. One alternatively spliced form of DCAR was found in mouse, lacking the neck region (Kanazawa *et al.*, 2003).

1.3 C-type lectin gene clusters

Families of C-type lectins most likely originate from common ancestors and are evolutionarily related. Mapped to certain gene clusters, these genes shared

homologies to some extent. Currently, the natural killer gene complex and antigen presenting lectin-like receptor complex represent two large gene clusters that host a large number of C-type lectin receptors.

1.3.1 Natural killer gene complex

The natural killer (NK) cell is a key component in innate immunity, being capable of distinguishing missing self by recognizing the MHC-I molecules on the target cells. NK cells have both activating and inhibitory receptors that are used to control and regulate their cytotoxic function. Generally, MHC-I molecules are expressed on all healthy host cells. When NK cells encounter self-MHC-I bearing cells, the inhibitory receptors will suppress NK lysis. Therefore, this inhibition prevents the normal cell from being killed. In other cases, when NK cells encounter tumor cells or microbial infected cells, the activating receptor transmits an activating signal and triggers the cytotoxic reaction in the target cells. These cells often lose their identification- MHC-I molecules, and are recognized by the receptors on NK cells. A delicate balance between activating and inhibitory receptors determines the outcome of NK cell activation.

NK activating and inhibitory receptors include members from two superfamilies. One is the Ig superfamily for killer cell Ig-like receptors (KIRs) at the leukocyte receptor complex, and the other is the C-type lectin like receptors (CTLR) for CD94, NKG2 molecules, CD69, at the NK gene complex (NKC)

(Lanier, 1998). NK cell expressed C-type lectins belong to the group V lectins within the C-type lectin family, all of which are type II transmembrane receptors and have a single CRD at the extracellular domain. The NK gene complex (NKC) was first found in the distal region of mice chromosomal 6, showing the genetic linkage of NK receptors mNKRP family, mCD69, mCD94, mNKG2 family and mLy-49, from telomere to centromere (Yokoyama *et al.*, 1991). The short arm of both human chromosome 12p12-p13 and rat chromosome 4 are both syntenic with this mouse NKC (Renedo *et al.*, 1997; Dissen *et al.*, 1996). Most genes encoded in this region belong to the C-type lectin superfamily and are selectively expressed on NK or other cells in the hematopoietic lineage.

1.3.2 APLEC cluster

Some APC, especially dendritic cells, express C-type lectins from a gene cluster that shares the same genetic region with NK receptors. Flornes *et.al*,(2004) mapped the APC C-type lectin genes cluster to a region that is located telomeric to the NKC. This cluster was named the antigen presenting lectin-like receptor complex (APLEC) (Flornes *et al.*, 2004). The APLEC region was found to be evolutionarily conserved among rat, mouse and human. The interval between APLEC and NKC is around 5Mb in rodent and 1Mb in human, and contains no other C-type lectin superfamily genes (Flornes *et al.*, 2004). Genes found in the APLEC include DLEC, DCIR, DCAR, Dectin-2, MCL and Mincle. Sharing even

higher homology and containing signaling motifs, DCIR, DCAR, Dectin-2 and DLEC/BDCA-2 were further grouped as the DCIR family (Kanazawa *et al.*, 2004b). The order and orientation of genes, as well as distances between genes are consistent from rat and mouse to human (Figure 1-2). Also common among the three species, Mincle, MCL and Dectin-2 are positioned at the centromeric end of APLEC and arranged in the same orientation and order. Nonetheless, human APLEC display big differences from rodents with respect to the length of the whole region and gene arrangement. Human APLEC is much shorter than rodents with a compact gene organization. In addition, unlike rodents that have four DCIR and two DCAR loci, humans have only one DCIR and no identified DCAR gene. Compared to rodents with genes distributed evenly within the APLEC, humans have a 0.4Mb empty region between DCIR and Dectin-2 with no known gene locus (Flornes *et al.*, 2004).

1.4 Avian lectins

Several C-type lectins have been identified in chicken, some of which were assigned to the two chicken MHC regions, the B complex and Rfp-Y complex. The chicken B complex contains genes encoding MHC class Iα and II β chains, TAP genes and Tapasin, and hence is the main region determining antigen presentation, allograft rejection and disease resistance (Jacob *et al.*, 2000; Kaufman *et al.*, 1999). In contrast, the Rfp-Y region contains genes encoding

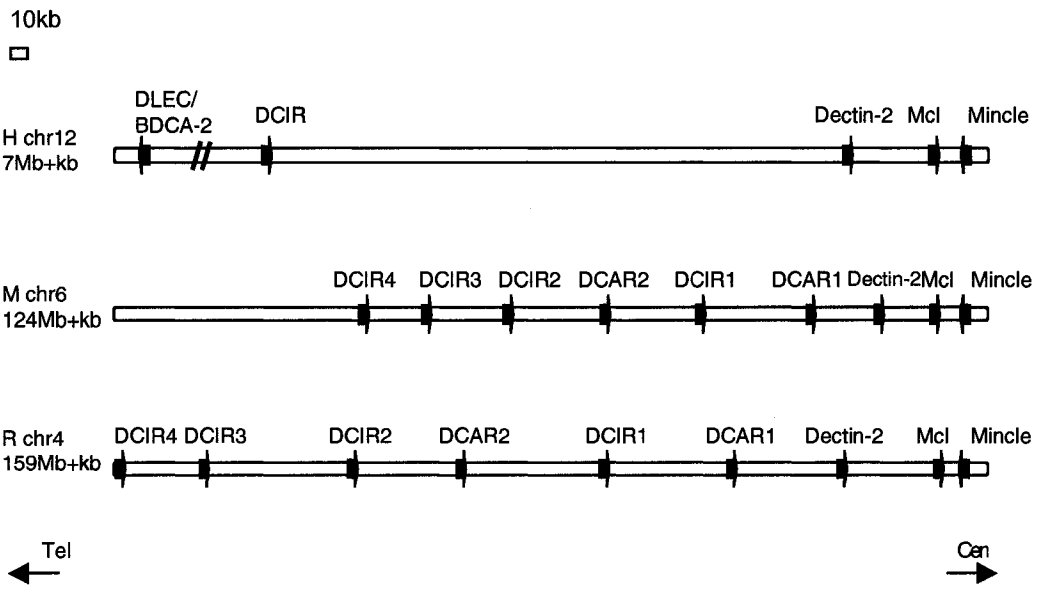


Figure 1-2. Gene organization and orientation of the APLEC region of human, mouse and rat. (Adapted from Flornes *et al.*, 2004)

non-classical MHC molecules. Due to the polymorphism of Rfp-Y MHC molecules, an intermediate effect was observed in skin allograft rejection (Thoraval *et al.*, 2003). Other functions of the Rfp-Y region in the chicken immune response were still unknown. While B and Rfp-Y regions were mapped to the same chicken microchromosome (chromosome 16), they are not genetically linked (Miller *et al.*, 1996).

So far, B-NK and B-lec were identified in chicken B complex, whilst Y-lec1 and Y-lec2 were situated in the Rfp-Y complex. All of the four lectin genes encode type II transmembrane receptors, and belong to Group V (NK cell receptor) C-type lectins. The CRD domain of all four lectins lack crucial calcium and sugar binding residues, although exhibiting the typical CRD folding. B-NK displays a conserved ITIM motif in its cytoplasmic domain, whereas B-lec has a canonical endocytosis motif (Rogers *et al.*, 2005). Both Y-lec1 and Y-lec2 show no obvious signaling motifs (Rogers *et al.*, 2003). Except B-NK that contains six exons, the other three genes have five exons. In the amino acid sequence, B-NK is most related to NKR-P1 family receptors, while B-lec is more related to CD69, AICL and LLT1 (Rogers *et al.*, 2005). Unlike B region lectins, Rfp-Y lectins are relatively farther from known mammalian lectins, but are either grouped with other lectins like ch17.5, or alone (Rogers *et al.*, 2003).

The co-localization of the chicken lectins with the MHC locus suggested that the evolution of the two gene complexes, NKC and MHC, are linked. Based on phylogenetic analysis and gene organization of several species, it has been

proposed that MHC and NKC shared a common ancestral region, which contained C-type lectins, MHC genes, and MHC associated genes (Rogers *et al.*, 2005; Trowsdale, 2001). This primordial region underwent genome-wide duplications first, and then various duplication, inversion or deletion events within each paralogous region.

1.5 Duck immune system

1.5.1 Relevance of ducks to human diseases

Avian species are the natural host of influenza A viruses. So far, all of the influenza viruses that caused global outbreaks, including the currently endemic H5N1, have been confirmed to be related to avian viruses, two of which originated from avian while the other two were human and avian reassortants. Acting as important carriers of influenza viruses, domestic and wild ducks both contribute to the propagation of viruses and may play a role in the transmission of viruses across species. In wild ducks, influenza viruses infect and replicate in the intestinal tract, then were secreted to feces and transmitted to other animals through direct contact or are waterborne. Previously influenza viruses were not able to cause lethal disease in ducks. In 2002, with the emergence of H5N1 virus, influenza acquired high pathogenesis and killed a large number of domestic and migratory birds, including ducks (Sturm-Ramirez *et al.*, 2004). In addition to the

change in pathogenesis, the new H5N1 viruses preferred the aerosol transmission route more than traditional oral-fecal route. A study carried out by Hulse-Post and colleagues (2005) showed that compared to old viruses, the H5N1 that appeared in recent three years has a relatively longer duration of virus shedding and lower pathogenesis to ducks yet is still lethal to other species. They also found that the less pathogenic virus variants were selected during the passage of viruses from the originally infected duck to others (Hulse-Post *et al.*, 2005). Further understanding of the evolution of influenza virus in duck and the adaptation of the duck immune system to viruses may be helpful to control bird flu spread and limit the chance of transmission from avian to human.

Duck is also a good model for research on Hepatitis B Virus (HBV) infection since duck Hepatitis B Virus (DHBV) belongs to the same *Hepadnaviridae* family as human HBV, and has a similar transmission pathway and infection outcome.

1.5.2 The significance of the research on duck lectins

A previous duck spleen EST project in our lab revealed several novel duck C-type lectins, including both APLEC and NKC lectins (Branton WG, thesis). In this study, we analyzed the genomic clones containing duck DCIR and DCAR genes, which is part of the potential duck APLEC region. Moreover, we characterized the genomic structure and mRNA transcripts of the two genes. The

study of these two duck lectins will enable us to 1) have greater understanding of the antigen presentation process in ducks; 2) develop a duck APC marker; 3) acquire novel approaches to manipulate duck immune responses; and 4) contribute to the big picture of lectin immunoreceptor evolution.

CHAPTER II. MATERIALS AND METHODS

2.1 cDNA library screening

From the preliminary EST screening of 3,000 colonies, two cDNA clones, 14B7 and 16F11 were isolated, both of which had significant identity to mammalian DCIR and DCAR genes, respectively. The two clones were named thereafter and sequenced from both strands. Based on the unique regions of the two clones, two pairs of overgo primers were designed: aDCIR-og F ^{5}AAGGAA ATAAGAAGACAGATGAAATAC^{3} and aDCIR-ogR ^{5}ATTGATTGCATTTTG TTGGTATTTC^{3}; DCAR-ogF ^{5}TGCTATTACTTCTCAGAGATCAG^{3} and aDCAR-ogR ^{5}TCTCATCCCAGGGCATCTGATCAT^{3}. Another cDNA clone, 9H3, shared high sequence identity to 16F11 with three gaps, and thus was thought to be an isoform of the duck DCAR gene.

A duck spleen cDNA library was arrayed on nylon membrane. cDNA clones were hybridized using a pool of probes, which included the overgo probes for DCIR and DCAR, along with TAP and AICL. Among the 120,000 cDNAs, 124 positive clones were isolated and sequenced from 5 end. Genetool analysis (Biotools Inc., Edmonton, Alberta) grouped the 108 valid ESTs into 10 contigs and 16 unmatched single sequences. The contigs and sequences then were annotated by NCBI BlastX (http://www.ncbi.nlm.nih.gov/ BLAST) and aligned

with two previous isolated duck C-type lectin sequences, 14B7 and 16F11. Contig 4 and contig 3, each contain four or five cDNA clones, and showed homology to 16F11 and 14B7 at the DNA sequence level, respectively. Therefore, the above nine duck DCIR and DCAR related clones were selected and subjected to further sequence analysis.

2.2 Fosmid genomic library screening

A genomic library from erythrocyte DNA of a male duck, *Anas platyrhynchos* (duck 26), was previously constructed in the fosmid vector pCC1FOS (Moon and Magor, 2004). Approximately 124,488 independent colonies were screened using a pool of probes including the DCIR, DCAR and AICR overgo probes. Positive clones were rearrayed and hybridized with each lectin overgo probe separately. Four genomic clones, 110O1, 140L23, 205H3 and 260J15, hybridized to both overgo probes. The sizes of the genomic inserts were estimated from BamHI restriction digests by field inversion gel electrophoresis (FIGE; BioRad Laboratories, Mississauga, Ontario). The four fosmid clones displayed similar restriction digestion and hybridization patterns. Since 110O1 showed more lectin hybridizing bands and contained the largest insert, it was selected for complete sequencing.

2.3 DNA sequencing

Generally, all the target sequences were subcloned into a plasmid vector. Sequencing reactions were performed on both strands using the BigDye Terminator 3.1 cycle sequencing kit (Applied Biosystems, Foster City, CA) on an ABI377 Automatic Sequencer (Applied Biosystems, Foster City CA). Chromatogram editing, contig assembly, nucleotide alignments and percent identity analysis were performed using Genetool 2.0 software and by visual inspection.

To sequence fosmid clones that contained genomic genes, restriction enzymes were used to make smaller subclones. In this case, BamHI digestion produced four unique subclones, which completely covered the original fosmid insert. Southern analysis showed one DCIR and DCAR double hybridizing clone with a 10kb insert, two DCAR hybridizing clones with 11kb and 12kb inserts respectively, and a 1.5kb subclone containing only the intergenic sequence. The 1.5 kb subclone was sequenced completely on both strands by primer walking. The large subclones were sequenced using the EZ: TN<KAN-2> transposon insertion strategy (Epicentre, Madison, WI), by which primer islands were randomly inserted into the subclones.

The order of the four BamHI subclones in the original fosmid was determined by two methods: restriction enzyme digestion and junction sequencing. Through aligning the end sequences of the original fosmid and subclone inserts, the 1.5kb subclone and 11kb subclone were localized to the two

ends of 11001, while the 10kb and 12kb fragment were in the middle. 10 kb subclone was localized next to 1.5kb subclone by amplifying overlapping junction regions. The orientation of the 12kb fragment was determined by KpnI digestion, which cuts the 12kb fragment unequally. The result was also confirmed by EcoRI and XhoI digestion.

2.4 Analysis of coding regions

The coding region of the edited 110O1 final sequence was predicted by Genescan (http://genes.mit.edu /GENCAN.html) and verified by aligning to the corresponding cDNA sequences. The exon-intron junctions were decided using BL2Seq (http://www.ncbi.nlm.nih.gov/BLAST/bl2seq/). Homology searches were done using NCBI BLAST. Putative amino acid sequences were predicted by Genetool® v2.0 and were aligned by CLUSTAL-W (http://www.ddbj.nig.ac.jp). Functional motifs in putative proteins were searched by Interpro (http://www.ebi. ac.uk/InterProScan/). Putative transmembrane domains and CRD domains were identified and labeled. Repeatmasker (http://repeatmasker.org) was used to identify the sequence repeats. GrailEXP v3.0 identified the location of repeats and significant CpG islands. Pipmaker (http://bio.cse.psu.edu/pipmaker) was used to produce a dot-matrix plot of the entire fosmid 110O1 contig against itself.

2.5 Promoter analysis

Upstream sequences of DCIR, DCAR1 and DCAR2 genes were analyzed for potential transcriptional factor binding sites. The 400 bp 5 upstream sequences were submitted to TESS string search (http://www.cbil.upenn.edu/cgi-bin/tess/). Conserved sites within the promoter regions were identified by eye.

2.6 Southern hybridization for clone mapping

Enzyme digested fosmid clones were separated on 0.8% agarose gel and then transferred to Performa nylon transfer membrane (Genetix USA Inc, Boston, MA) and immobilized by UV cross-linking. Blots were hybridized to DCIR or DCAR overgo probes. Hybridizations were carried out for 16 hours at 62°C in aqueous hybridization solution (5x Denhardt s, 6 x SSC, 5% dextran sulfate, 1% SDS and 100 mg/ ml salmon sperm DNA (1 x SSC: 0.15M NaCl, 15mM $Na_3C_6H_5O_7$ pH 7.6, 1mM EDTA)). Membranes were then washed at high stringency in 0.1 x SSC, 0.1% SDS at 68°C. Overgo probes were radiolabelled with ^{32}P α-dCTP. The blots were exposed to X-OMAT film (Kodak) at -80°C for 16-18 hours. Blots were stripped in 0.1%SDS at 100°C for 1 hour, stripped twice when needed.

To select DCIR or DCAR positive BamHI subclones, dot blots were prepared for hybridization. The plasmid DNA of the subclones was directly loaded to the membrane stabilized in a vacuum blotting manifold. Dot blot hybridizations were conducted with similar conditions as stated above.

CHAPTER III. RESULTS

3.1 Selection of duck genomic clone 110O1

To determine the genomic structure and organization of duck C-type lectins, a duck genomic library was screened with overgo probes designed on EST sequences of duck DCIR and DCAR. Four positive clones, namely 110O1, 140L23, 205H3 and 260J15, were obtained. The four fosmid clones were then digested by various enzymes, including PstI, PvuII and BamHI. Digestion patterns were shown in Figure 2-1. Cutting less often, BamHI was selected to make subclones. After BamHI digestion, 110O1 insert released four fragments, with the size of 1.6kb, 10kb, 11kb and 12kb, respectively. The aggregate size of 110O1 fragments was approximately 35kb, which was larger than the other three clones. In addition, 110O1 and 260J15 both contained one DCIR and multiple DCAR hybridizing fragments, as shown in the Southern blot (Figure 2-2A and 2-2B). Taking this information into consideration, clone 110O1 was chosen for further sequencing analysis.

To select DCIR or DCAR containing fragments, spot hybridization was done using the same sets of DCIR or DCAR overgo probes. Subclones of the four fosmid clones were loaded to the membrane. As seen in Figure 2-3, each fosmid clone contained one subclone that hybridized with both DCIR and DCAR probes.

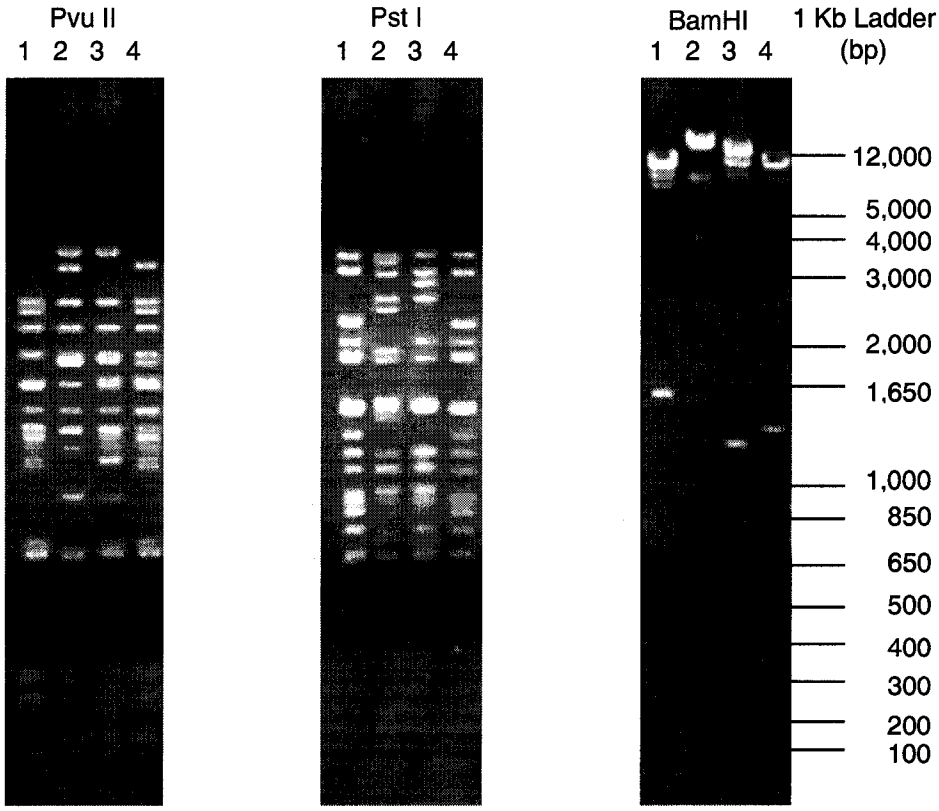


Figure 2-1. Restriction enzyme digestion of the four positive fosmid clones, 110O1 (1), 140L23 (2), 205H3 (3) and 260J15 (4). Fosmid clones were digested using PvuII (the left panel), PstI (middle panel) and BamHI (right panel). The 1 Kb ladder was used as molecular weight marker.

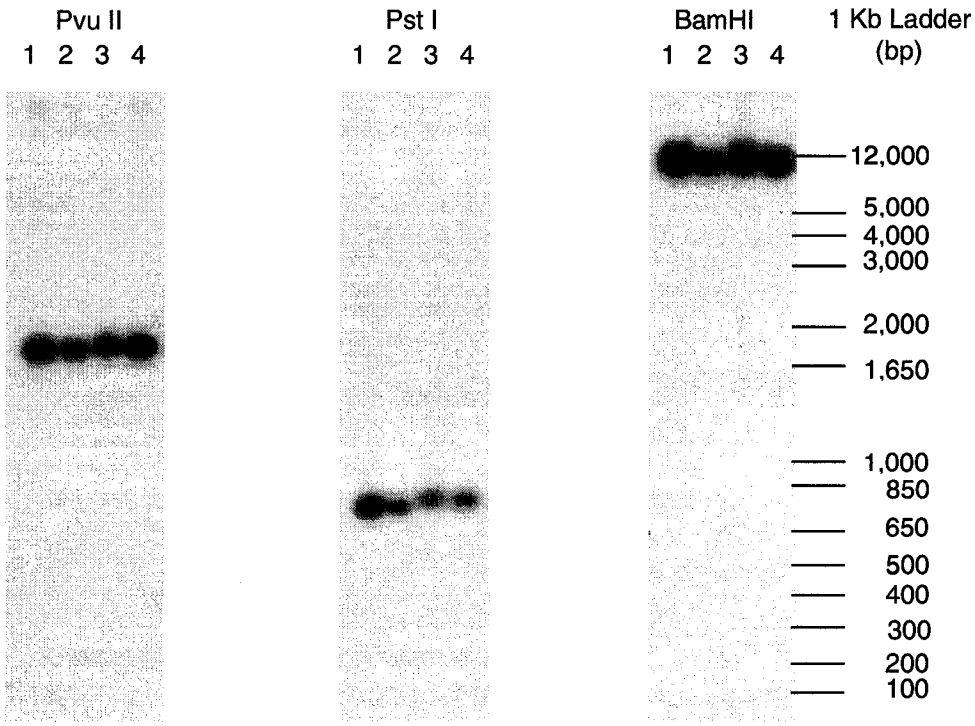


Figure 2-2A. Southern blot of the hybridization of enzyme digested fosmid fragments to duck DCIR overgo probes, which was designed based on the unique sequence of the CRD domains. The gel shown in Fig. 1 was transferred to nylon membrane. Lanes correspond to: 1: 110O1; 2: 140L23; 3: 205H3 and 4: 260J15. The 1 Kb ladder was used as molecular weight marker.

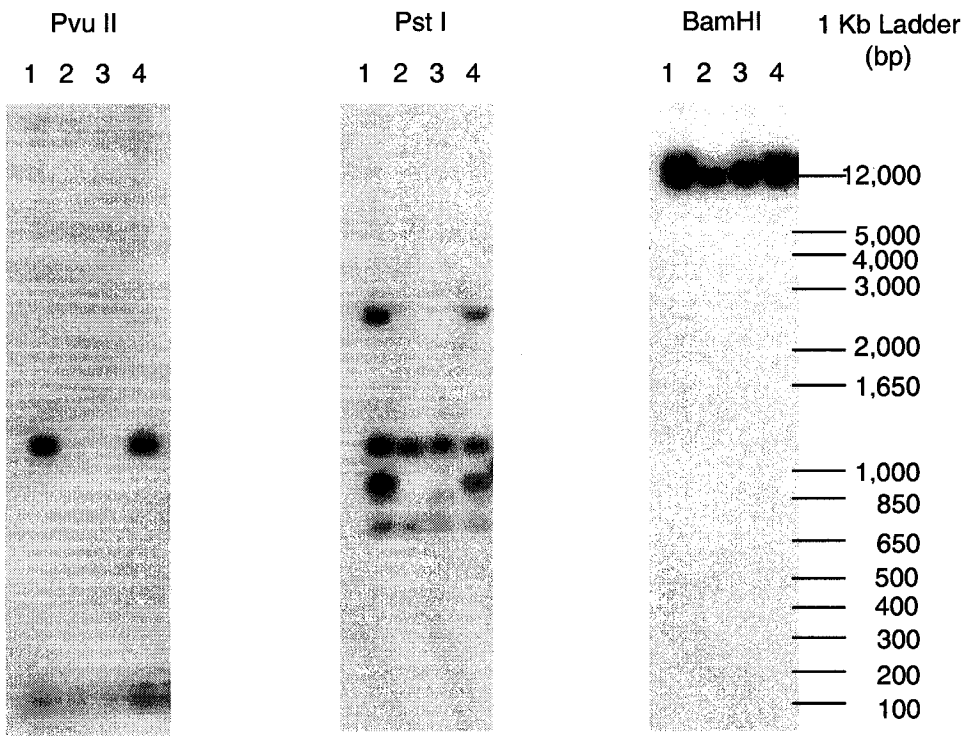


Figure 2-2B. Southern blot of the hybridization of enzyme digested fosmid fragments to duck DCAR overgo probes. Each fosmid clone contained one or two DCAR positive Bam HI digested fragments. Lanes correspond to: 1: 110O1; 2: 140L23; 3: 205H3 and 4: 260J15. The 1 Kb ladder was used as molecular weight marker.

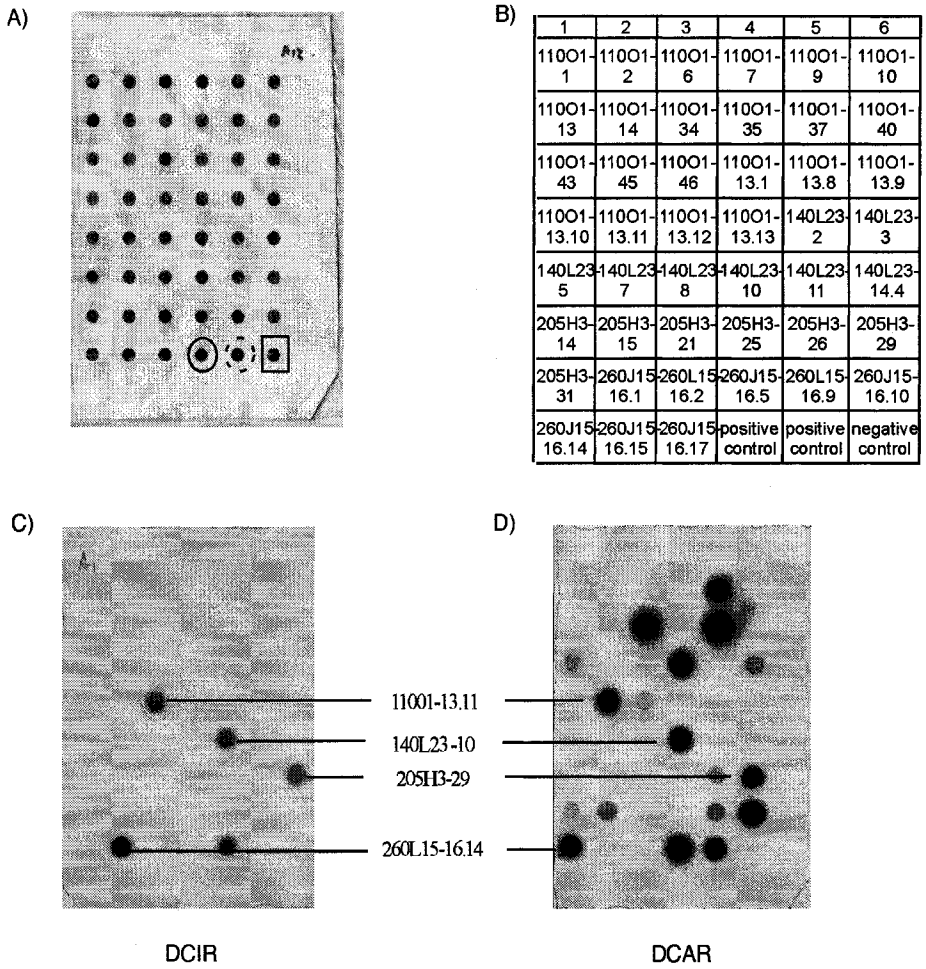


Figure 2-3. DCIR or DCAR positive subclones were selected by dot blot. Bam HI digested fragments were subcloned into pBS-KS+ vector. Subcloned DNAs were extracted and loaded on nylon membranes (as shown in (A), the top panel). Solid and dashed circles, as well as the black square indicate the position of a double positive, a DCAR positive and a negative control, respectively. The order of loading clones is shown in (B). Southern blots were hybridized with the same DCIR (C) and DCAR(D) probes as described in Fig. 2.

The double positive subclones were 110O1- 13.11, 140L23-10, 205H3-29 and 260J15-16.14. The other twelve subclones hybridized to the DCAR probe only.

Alignment of the end sequences of these DCAR positive subclones suggested that they might be the same, or alleles of two gene fragments, already represented by110O1 46 and 110O1 34.

3.2 Genomic organization of duck DCIR and DCAR genes

For each of the four subclones derived from 110O1, the sequence was obtained by transposon insertion. The four fragments were put together by restriction enzyme digestion and sequencing of overlapping fragments. A contig of 34.9kb containing one DCIR positive gene and two DCAR positive genes (DCAR1 and DCAR2) was assembled. The three genes were arranged tandemly and in the same transcriptional orientation (as seen in Figure 2-4A). The restriction enzymes used to subclone and to determine the orientation of the fragments were shown in Figure 2-4B. The overall G+C nucleotide content over the whole contig was 49.7%. CpG island analysis was done by Pipmaker. No significant CpG islands were detected in the entire contig.

For insight into the duplication events, we analyzed for the presence of repeat elements in the entire sequence. Repeat elements were predicted using GrailEXP and shown in Figure 2-4C. Four copies of long interspersed nuclear elements (LINE) CR1, which represents the most abundant repetitive element in

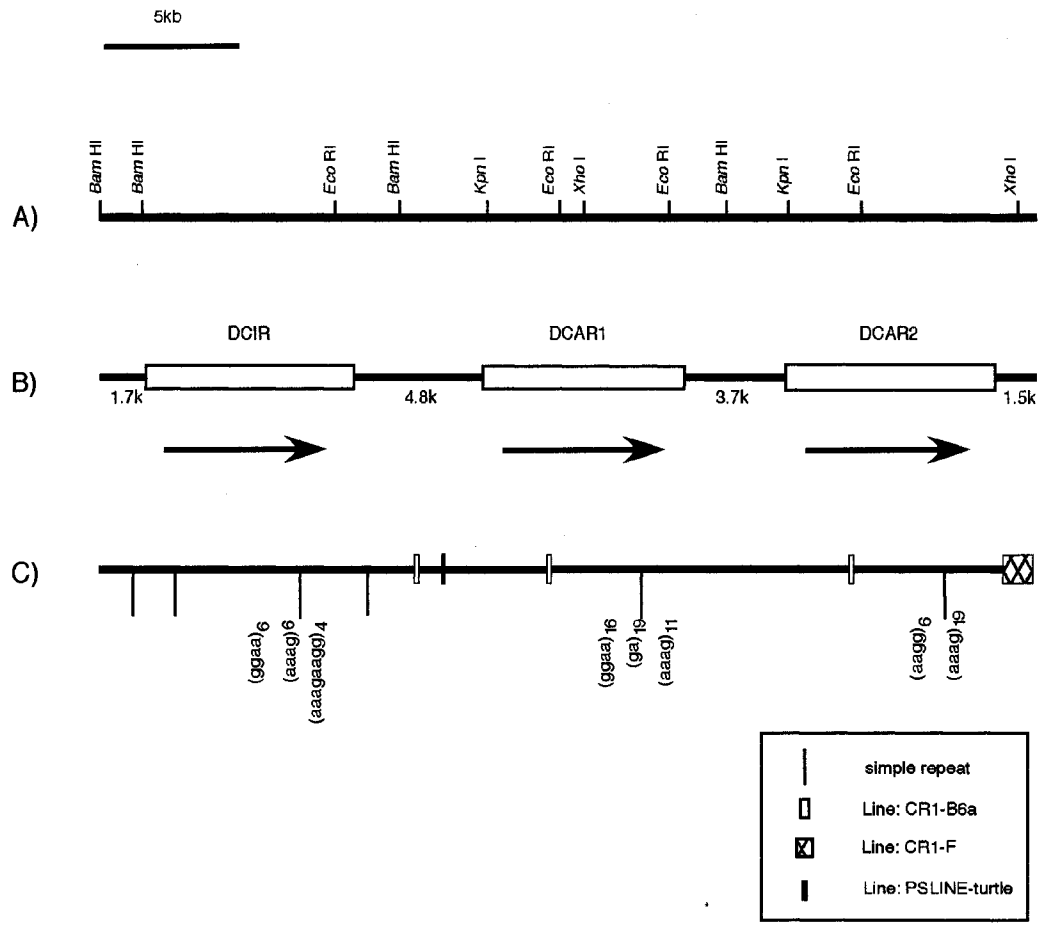


Figure 2-4. Schematic view of duck genomic clone 11001. A) Restriction digestion sites used for mapping are shown. Four Bam HI digested fragments were subcloned and sequenced. EcoRI, KpnI and XhoI were used to organize subclones and determine gene orientation. B) Duck DCIR, DCAR1 and DCAR2 genes are located in one genomic clone 110O1. They are arranged tandemly and in the same transcriptional orientation. C) Repetitive elements were identified using Grail EXP and Repeatmasker. Four LINE-CR1 repeats and one PSLINE-turtle repeat were detected. Three GA rich simple repeat regions were also identified. They were located at similar positions within each gene.

the chicken genome, as well as one PSLINE repetitive element, were identified. Among the four CR1/LINES repeats, three CR1-B6a LINE were the same length and shared high sequence similarity. Two of the CR1-B6a repeats found within DCAR1 and DCAR2 were 100% identical, and both were 75% identical to the other one between DCIR and DCAR1. The similarity of sequence and the location of repeats suggested that DCAR1 and DCAR2 were a more recently duplicated pair than DCIR and DCAR. In addition, three AG rich simple repeat regions were detected located at similar positions in all three genes, which also supports the proposed gene duplication.

A dot plot (Figure 2-5) analysis of the entire insert of 110O1 against itself illustrates the gene duplication events within this region. Gene fragments encoding the CRD domain of all three genes including exons and introns, were highly conserved, as indicated by the diagonal lines within the shaded regions. Moreover, whole gene conservation is also observed between the two DCAR genes, DCAR1 and DCAR2. The diagonal lines extend outside the exons, suggesting that duplication includes part of the intergenic region. Again, results from the dot plot suggested that the duplication of DCAR1 and DCAR2 is more recent, which is consistent with our previous interpretation of the repetitive elements analysis.

The 110O1 genomic contig contains a 1.7kb 5 flanking region of the duck DCIR gene and the whole 5 upstream sequence of the DCAR1 and DCAR2 genes. To identify possible promoter proximal regulatory elements, 400 bp of

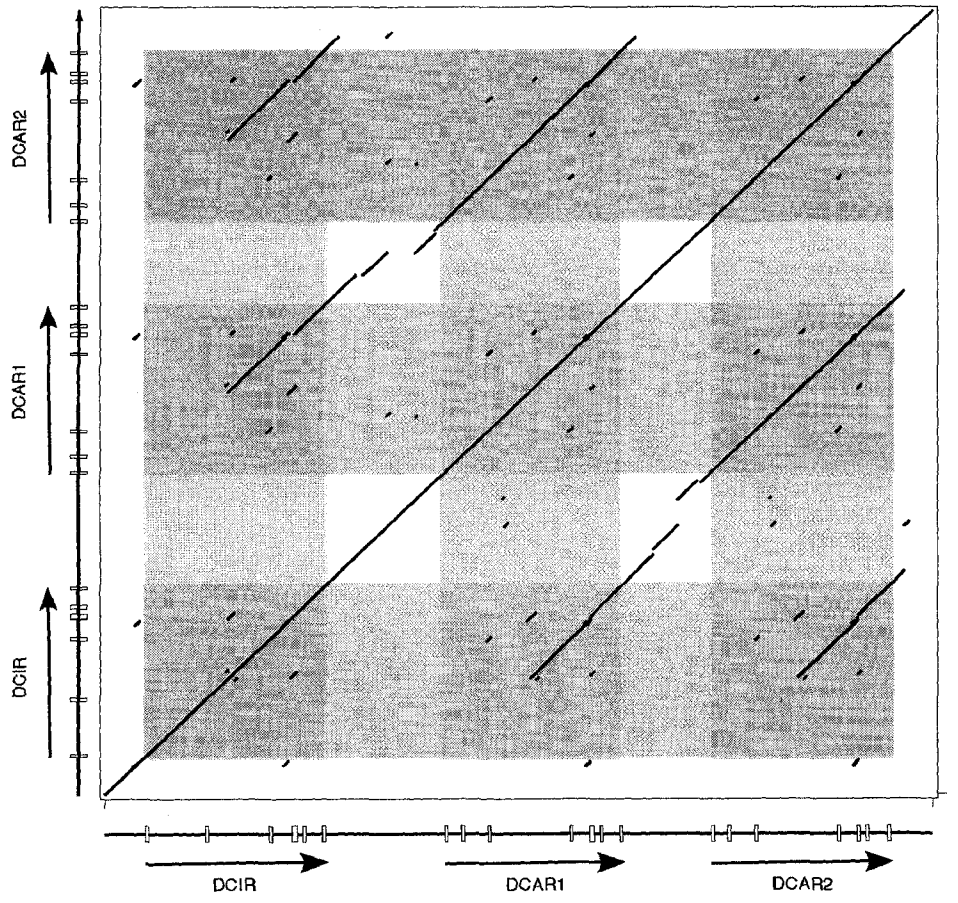
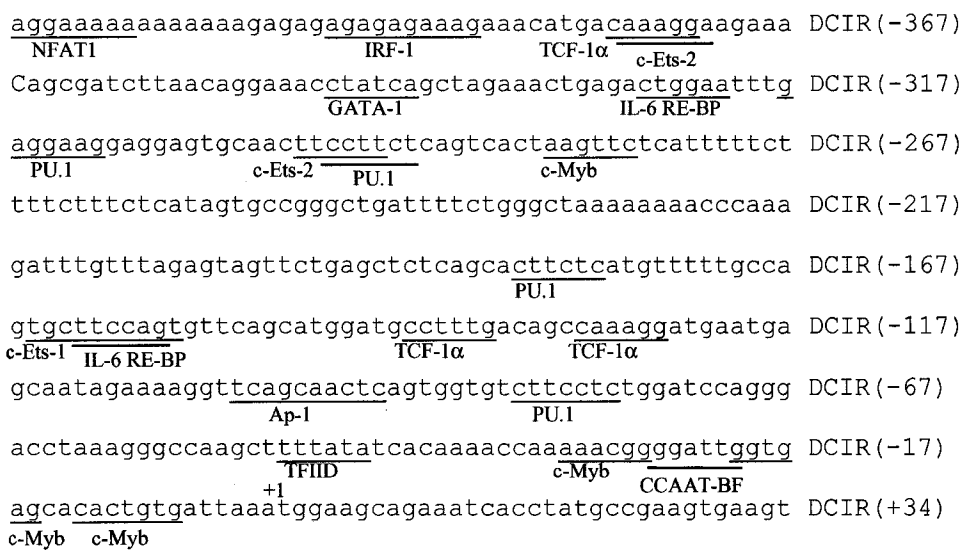
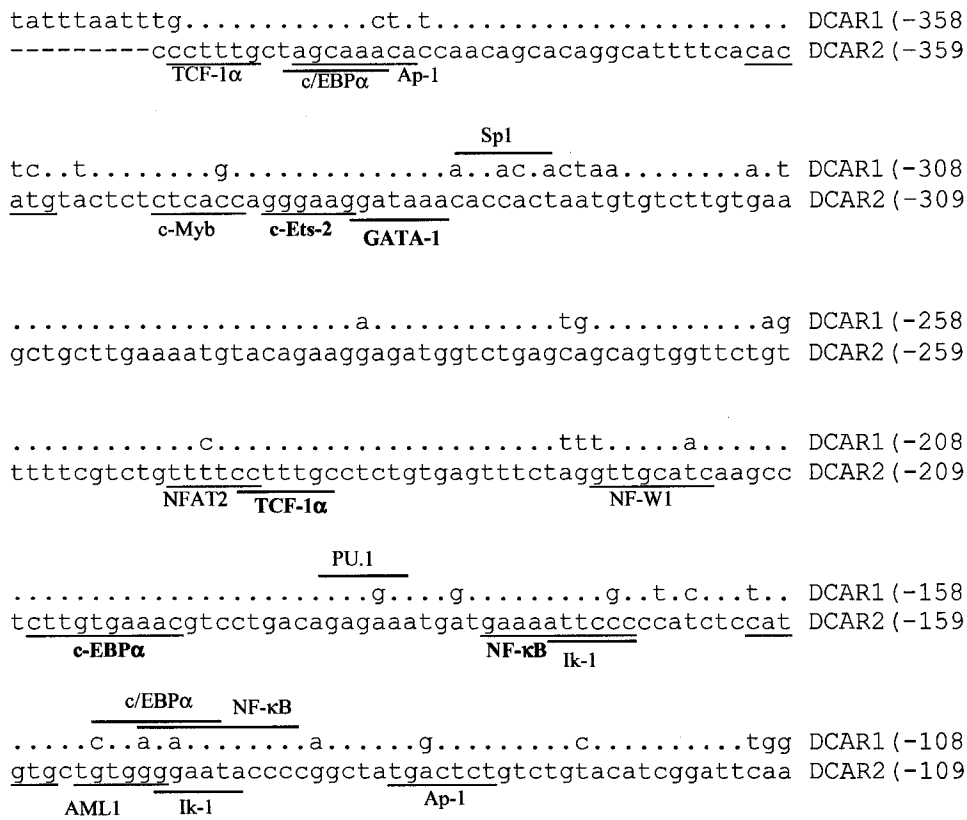


Figure 2-5. Dot plot analysis of duck APLEC region versus itself. Diagonals indicate the regions sharing significant sequence identity.

A)



B)



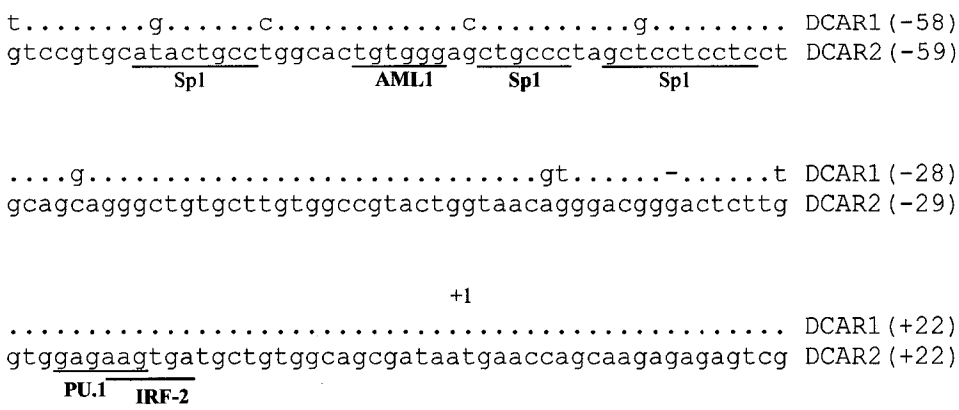


Figure 2-6. Sequence analysis of the 400bp promoter regions of duck DCIR (A) and DCAR 1 and 2 (B) genes. The binding motifs for transcription factors are indicated with lines under the DNA sequences. Since the transcriptional initiation sites are unknown, the start of the coding region was set as +1. All binding sites shown here have at least 85% sequence identity to the consensus binding motif. Bold indicates the binding sites of regulatory element that common in DCAR1 and DCAR2. DCAR2 specific regulatory elements binding sites were shown in normal font. Abbreviations: IRF, interferon γ regulatory factor; TCF, T cell factor; IL-RE BP, Interleukine 6 regulatory element binding protein; c/EBP, CCAAT-enhancer binding protein; AML, acute myeloid leukemia gene factor.

upstream sequence of DCIR, DCAR1 and DCAR2 genes were analyzed using the TESS program. As seen in Figure 2-6, several transcription factor binding sites that are essential for regulation in hematopoesis were found in the sequence upstream of both DCIR and DCAR2 genes, including the binding motif of PU.1, CCAAT/enhancer binding protein α (c-EBPα), c-myb, and acute myeloid leukemia gene 1(AML1) that functions in myeloid development and differentiation, and of T cell factor 1α (TCF-1α) Ikaros (Ik-1)and nuclear factor of activated T cells 1 and 2 (NFAT 1 and 2) that regulate lymphocyte differentiation. Some motifs commonly found in RNA polymerase II promoter module were also detected, including those specific for GATA-1, CCAAT-binding protein and Sp1. It is worth noting that interferon regulatory factor binding regions have been identified, an IRF-1 in DCIR and an IRF-2 in DCAR. Although there was no classical TATA box found in the promoters of both genes, a TFIID binding site was found in DCIR -50 site, indicating that this region might be the core promoter of the gene. The identification of above transcription factors indicates that duck DCIR and DCAR2 genes are subjected to the regulation by both ubiquitous elements and specific factors, whereas the pseudogene DCAR1 had fewer regulatory elements. Despite the high sequence similarity (87%) of DCAR1 and DCAR2 5 upstream regions, only 9 regulatory elements are common in both (as shown in Figure 2-6B).

3.3 DCIR or DCAR on other genomic clones

To screen for the DCIR locus on the other three fosmid clones, two strategies were adopted: sequencing of selected regions and hybridization using a new specific probe based on the unique ITIM motif. Fosmid clones were digested by selected restriction enzymes and blotted to Nytran. As indicated by the Southern blot (Figure 2-7A), 110O1, 140L23 and 205H3 contained only one DCIR hybridizing band, while 260J15 had no hybridization. 110O1 released a slightly larger hybridization band than 140L23 and 205H3, because of the appearance of the vector multi-cloning site at the 5 end. End sequencing of the four DCIR containing subclones revealed that the 260J15 insert starts from the first intron of the DCIR gene and thus missed the ITIM containing exon I. The results in Figure 2-7B and Table 2-1, demonstrate the possible gene organizations of the other three fosmid clones. Our data at hand are not sufficient to draw any conclusions about whether the four fosmid clones were made from one allele or two.

3.4 Gene structure of duck DCIR

To determine which genes were expressed, a cDNA library was screened using probes prepared from EST clone sequences. Five clones were recovered and completely sequenced. Duck C-type lectin cDNAs that were isolated from a cDNA library constructed from the spleen of the same duck, were aligned with

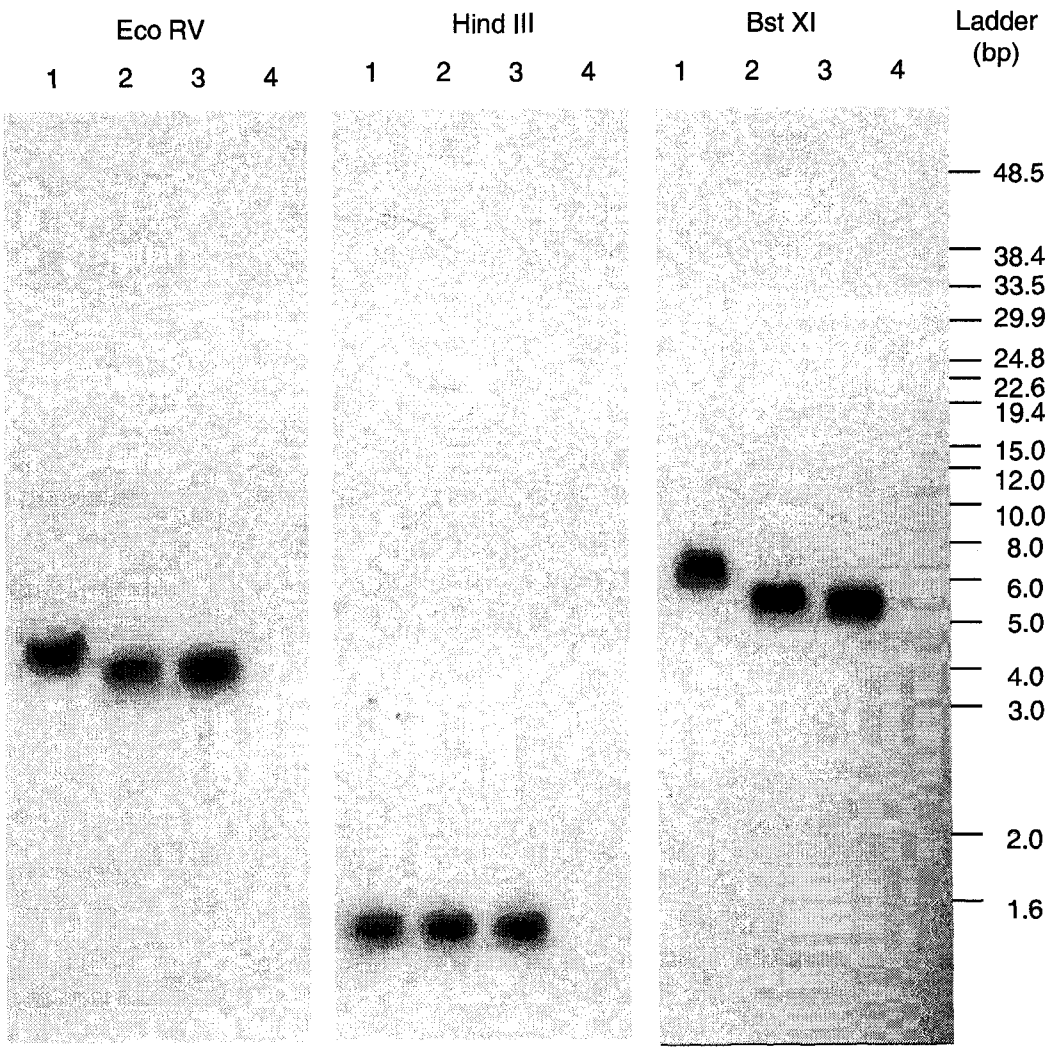


Figure 2-7A. Southern blot of the hybridization of enzyme digested fosmid fragments to new duck DCIR overgo probe. The new probes were designed based on the unique ITIM motif of duck DCIR. Lanes correspond to: 1: 110O1; 2: 140L23; 3: 205H3 and 4: 260J15. The 1 Kb ladder and 8-48 Kb high molecular weight ladder were used as molecular weight markers.

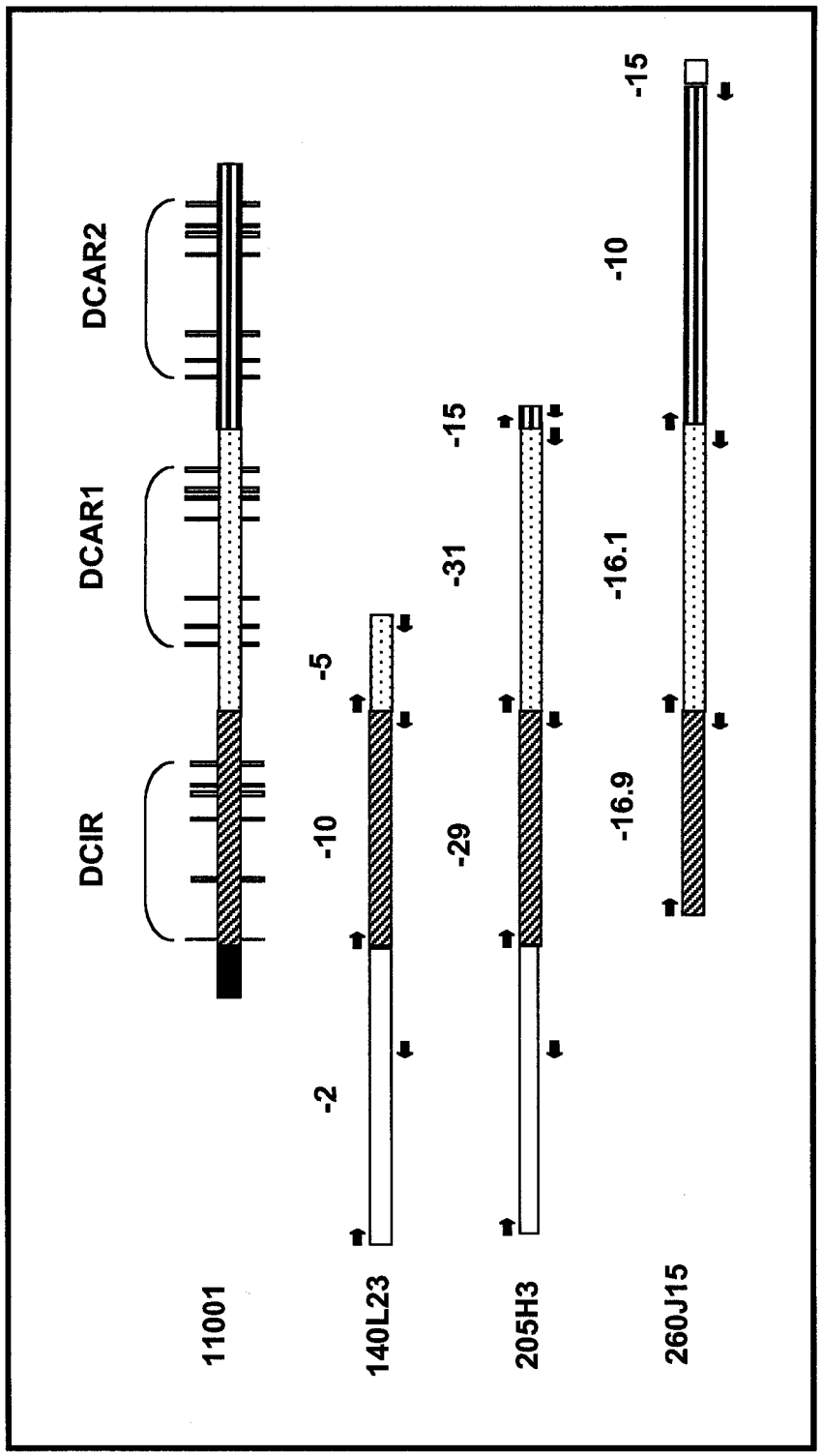


Figure 2-7B. Gene organization for 110O1, and the hypothetical organization of three fosmid clones.

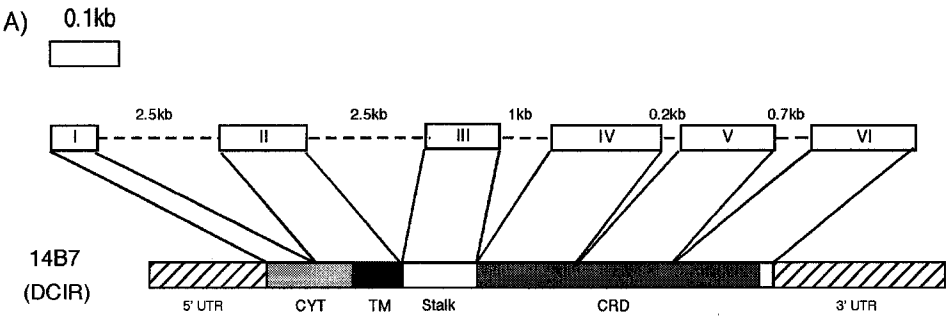
fosmid clone	number of fragments	size of insert (kb)	subclone	hybridization to DCIR/DCAR	Coding for
110O1	4	12	110O1-46	-/+	duck DCAR1
		11	110O1-34	-/+	duck DCAR2
		9.5	110O1-13.11	+/+	duck DCIR (from exon II)
		1.6	110O1-35	-/-	duck DCIR (UTR+exon I)
140L23	3	>12	uk (could be -2)	-/-	upstream of DCIR
		9.5	140L23-10	+/+	duck DCIR (from exon II)
		4.2	140L23-5,-3	-/-	part of DCAR1
205H3	4	>12	uk	-/-	upstream of DCIR
		12	205H3-31,-26	-/+	duck DCAR1
		9.5	205H3-29	+/+	duck DCIR (from exon II)
		1.3	205H3 -15	-/-	part of duck DCAR2
260J15	4	12	260J15-16.10	-/+	duck DCAR1
		11-12	260J15-16.1	-/+	duck DCAR2+some downstream
		9	260J15-16.9	+/+	part of duck DCIR
		1.4	260J15-15	-/-	duck DCAR2 downstream, should be back to back with 16.1

Table 2-1. A summary of BamHI digested fosmid fragments. UK stands for unidentified subclones.

three genomic sequences. One clone, 14B7 was identical to the duck DCIR gene, the structure of which is shown in Figure 2-8A. Clone 14B7 was 1117 base pairs in length, with an open reading frame of 711 bp. The putative translation start codon was located at 167 bp, which is not in a consensus Kozak context. There were several other upstream ATGs, however they were followed by in-frame stop codons. A protein of 237 amino acids was deduced from the cDNA sequence, as shown in Figure 2-8B. Starting from the translation initiation site, the gene was 6864 base pairs in length and was composed of six exons. Following a 166 bp untranslated region (UTR) at 5 - end, Exon I encodes 22 aa of the cytoplasmic domain. Exon II encodes the remaining cytoplasmic domain and the entire transmembrane domain. Exon III encodes the stalk region. Exon IV, V and VI construct the CRD domain followed by a 183 bp 3 UTR.

Analysis of the amino acid sequence identified duck DCIR as a type II C-type lectin with the following features: first, a conserved CRD domain ranging from aa99 to aa231; second, a 23 aa transmembrane domain; and third, six conserved cysteine residues in its CRD domain. Interestingly, an ITYAEV motif identical to the conserved sequence of the mouse and human ITIM (I/VXYXXL/V) was found in the cytoplasmic domain. The identification of an ITIM motif, that is also common in NK cell inhibitory receptors, infers that duck DCIR may also play an inhibitory role in duck immune responses. BLASTP results showed that the CRD domain of 14B7 had greatest homology to human, mouse and rat DCIR (45%, 41% and 40%, respectively), human and rat Mincle

A)



B)

```
   1 gtgcttccagtgttcagcatggatgcctttgacagccaaaggatgaatgagcaaagcaat
  61 agaaaaggttcctcagtggtgtcttcctctggatccagggacctaaagggccaagctttt
                                                          E1 ↱
 121 atatcacaaaaccaaaaacggggattggtgagcacactgtgattaaatggaagcagaaat
                                                      M  E  A  E
                                                                E2 ↱
 181 cacctatgccgaagtgaagttcaagaatgcatcaccaactgaagaggttgaagtacctca
   I  T  Y  A  E  V  K  F  K  N  A  S  P  T  E  E  V  E  V  P  Q
 241 gaagaagcagcagcagcatgagcaacatacgcagacatgccctccatggctcccgtggct
     K  K  Q  Q  Q  H  E  Q  H  T  Q  T  C  P  P  W  L  P  W  L
                                                                E3 ↱
 301 gatctcactgctcctgctcctggtgtgcgttgcccttgttgttgttctcctagtcactca
     I  S  L  L  L  L  L  V  C  V  A  L  V  V  V  L  L  V  T  H
 361 cgtcccccagagctgtgacaagcccgcagtcctgcagcggaaccacacagggtggcactg
     V  P  Q  S  C  D  K  P  A  V  L  Q  R  N  H  T  G  W  H  C
                                  E4 ↱
 421 catcttggcagtgcatcaaggcaaagaggacagctggaagtgctgtccagagggctggag
     I  L  A  V  H  Q  G  K  E  D  S  W  K  C  C  P  E  G  W  R
 481 acccctttcaggaaagctgctattacttctcagatgatcagatgccctggaatgagagcaa
     P  F  Q  E  S  C  Y  Y  F  S  D  D  Q  M  P  W  N  E  S  K
                                                          E5 ↱
 541 gaagaactgcagtgggatgggctcccagctggtggtgatcaatacagaagcagagcagga
     K  N  C  S  G  M  G  S  Q  L  V  V  I  N  T  E  A  E  Q  D
 601 tttcctctataaggaaataagaagacagatgaaataccaacaaaatgcaatcaatttatt
     F  L  Y  K  E  I  R  R  Q  M  K  Y  Q  Q  N  A  I  N  L  F
 661 catcggtctgagggcacaggaggtgggccagtggcgctgggcagaccagactccctataa
     I  G  L  R  A  Q  E  V  G  Q  W  R  W  A  D  Q  T  P  Y  N
                E6 ↱
 721 tgaatcagcagcgttctggaggtctggggagccaagtaataaatctgatgagctgtgtgt
     E  S  A  A  F  W  R  S  G  E  P  S  N  K  S  D  E  L  C  V
 781 tgtaatccatcacaaaacagaaaacctccggaactggaatgatgtcccgtgcagaatacg
     V  I  H  H  K  T  E  N  L  R  N  W  N  D  V  P  C  R  I  R
 861 ttcttatcggatttgtgagactgcagcagtaactctatgatggaggaatcctcatcctga
     S  Y  R  I  C  E  T  A  A  V  T  L
 921 gattagcagcgaactgggaacagcagagggctgtgttgggaggggtgggagagccttgga
 981 gcctttatcttgcctctgctggtgggatgatgagactggggaactatgacgcctcaagta
1041 gaaacaataaatgctagagaactctgaaaaaaaaaaaaaaaaaaaa
```

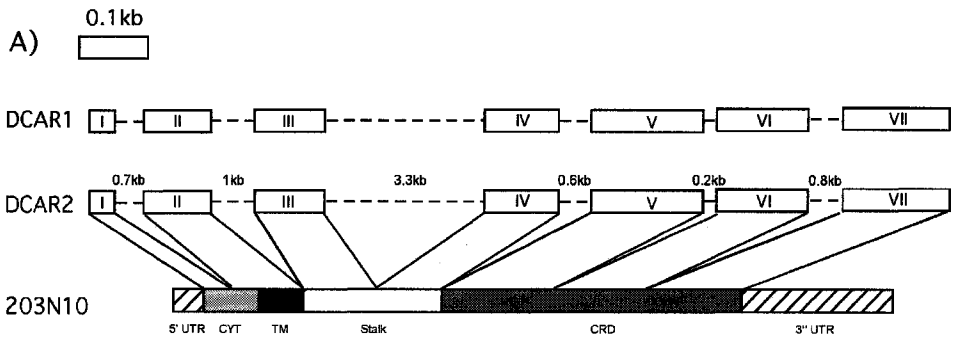
Figure 2-8. Schematic view of the genomic structure of duck DCIR (A) and its cDNA sequence (B). A) Boxes and lines represent exons and introns respectively. The DCIR gene contains six exons (I-VI). The region of the cDNA contributed by each exon is indicated. B) cDNA and putative amino acid sequences of DCIR are shown. The amino acid sequence encoding the transmembrane domain is indicated in bold. The CRD domain is underlined. Exon junctions are labeled above the DNA sequence.

(45% and 44%, respectively), human and rat MCL (42%), human BDCA-2 (40%), mouse DCAR (37%) and rat DC-SIGN (39%).

3.5 Gene structure of duck DCAR

To determine the genomic structure of the duck DCAR genes, the two DCAR genomic sequences were aligned with corresponding cDNA clones. One clone, 203N10 matched exactly the translation of DCAR2 mRNA from a consensus Kozak start codon at 76bp, followed by an ORF 750 base pairs long, encoding a 249 aa polypeptide. The length of DCAR1 and DCAR2 genes are 7568bps and 7798 bps, respectively, starting from the translation initiation site, The DCAR2 gene consists of seven exons, instead of six, like most C-type lectins found in the APLEC gene cluster. The gene structure of DCAR2 is depicted in Figure 2-9. After the 5 UTR, exon I encodes part of the cytoplasmic domain. Exon II encoded the remaining cytoplasmic domain and the 21aa transmembrane domain. It has an unexpectedly long 63 aa stalk region, which is encoded by exon III and IV. Similar to DCIR and other APLEC C-type lectin genes, the CRD domain of duck DCAR2 is also composed of three exons, exon V, VI and VII. A 210 bp 3 UTR follows exon VII. The cDNA was also characterized as a type II C-type lectin containing a conserved CRD domain with six cysteine residues and a 21 aa transmembrane domain. A striking feature found in the transmembrane

A)



B)

```
                                                                E1 ↱
   1 ggacgggactcttggtggagaagtgatgctgtggcagcgataatgaaccagcaagagaga
                                                       M  N  Q  Q  E  R
                                    E2 ↱
  61 gtcggtcctgggactgcagccccagcagaagggagcagctgttcccgcctgagcccctgg
      V  G  P  G  T  A  A  P  A  E  G  S  S  C  S  R  L  S  P  W
                                                                E3 ↱
 121 gtcttcctcgcttctgcccttgccgtcaaaactgccctcatgaccgttggcctcgttgtt
      V  F  L  A  S  A  L  A  V  K  T  A  L  M  T  V  G  L  V  V
 181 ctctttcacaggagctgtggccagtacaagactctgccccagaatgctccagggtggcac
      L  F  H  R  S  C  G  Q  Y  K  T  L  P  Q  N  A  P  G  W  H
                                          E4 ↱
 241 tgcatccccaatggatctgcaagccaaatcactcatgtctcccagagctgtgacaagccc
      C  I  P  N  G  S  A  S  Q  I  T  H  V  S  Q  S  C  D  K  P
 301 gcagccctgcagggaaccatacagagttgcactgcatcttggcggtgcatcaaggcaaa
      A  A  L  Q  G  N  H  T  E  L  H  C  I  L  A  V  H  Q  G  K
E5 ↱
 361 gggcgagacctgaagtgctgttcagagggctggagacccttcaggaaagctgctattac
      G  R  D  L  K  C  C  S  E  G  W  R  P  F  Q  E  S  C  Y  Y
 421 ttctcagatgatcagatgccctgggatgagagccagcagaactgcagtgggatgggctcc
      F  S  D  D  Q  M  P  W  D  E  S  Q  Q  N  C  S  G  M  G  S
                                          E6 ↱
 481 cagctggtggtgatcaatacaaaagcagagcaggctttcctctataaggaaatacagatg
      Q  L  V  V  I  N  T  K  A  E  Q  A  F  L  Y  K  E  I  Q  M
 541 aaataccgacaaaatggaatcaatttatacatcggtctgagggcacagaaggtgggccag
      K  Y  R  Q  N  G  I  N  L  Y  I  G  L  R  A  Q  K  V  G  Q
                                                E7 ↱
 601 tggcgctgggcagaccagactccctataatgaaagagcagcgttctggaggcgtggggag
      W  R  W  A  D  Q  T  P  Y  N  E  R  A  A  F  W  R  R  G  E
 661 ccaagtgatcaaccaagtgatgagctgtgtgttgtaatccattaccagaaagatattttc
      P  S  D  Q  P  S  D  E  L  C  V  V  I  H  Y  Q  K  D  I  F
 721 cggaactggaataatgtcccatgcacaatccactcttattggatttgtgagactgcagca
      R  N  W  N  N  V  P  C  T  I  H  S  Y  W  I  C  E  T  A  A
 781 gaaacaatatgatggaggaatcctcatcctgagatgagcagcgaactgggaacagcagag
      E  T  I  *
 841 ggctgtgttgggaggggtaggagagccttggagcatttgtcttgcctctgctcgtgggat
 901 gatgagactgggagtgatgttgctctgcacacagcatcccttgtgcatgtgtatttctca
 961 aagtacccctgatgtggaaacaataaatgctagagaactctgaaaaaaaaaaaaaaaaaa
1021 aaaaaaaaaaa
```

Figure 2-9. Schematic view of the genomic structure of duck DCAR1 and 2 (A) and the cDNA sequence of DCAR2 (B). A) Boxes and lines represent exons and introns respectively. The DCAR2 gene contains six exons (I-VII). The region of the cDNA contributed by each exon is indicated. B) cDNA and putative amino acid sequences of DCIR were shown. The amino acid sequence encoding the transmembrane domain was indicated in bold. The CRD domain was underlined. Exon junctions were labeled above the DNA sequence.

domain was a positively charged Lysine residue (K), suggesting the involvement of adaptor molecules in a manner similar to activating immunoreceptors.

A BLASTP search revealed that the predicted amino acid sequence of the CRD domain of DCAR2 had 44% homology to canine and rat Mincle, 42% to mouse MCL, 41% to human MCL or BDCA-2, 39% to rat and mouse DCIR, and human Dectin2, as well as 36% to mouse DCAR

The DCAR1 gene had a similar structure with DCAR2 in Exon I-V, except for a few nucleotide mismatches in Exon II, IV and V. A premature stop codon in the middle of exon VI makes it a truncated version, with the deletion of 31 aa of exon VI and the whole exon VII, leading to the absence of more than half of the CRD domain. DCAR1 displayed 89% overall sequence similarity to DCAR2. Most sequence mismatches were concentrated in the 5 and 3 non-coding regions. Since no transcripts identical to this gene were found, it was inferred that the DCAR1 gene might not be a functional gene.

3.6 Duck DCAR and DCIR shared high sequence identity

Three novel genes were aligned and compared at both genomic and amino acid sequence level. High sequence identity was observed. At the genomic level, the three genes share two conserved regions. One is a 2.3 kb region starting from the middle of DCIR intron 2 (intron 3 for DCARs due to an extra exon encoding the stalk region) and ending at the middle of DCIR intron 3. The other is a 1.9 kb

region starting from DCIR exon IV (exon V for DCARs) to around 500bp right after the stop codon, spanning the genomic region that encodes the whole CRD domain. Sequence comparison suggested that DCAR1 and 2 are more similar to each other than to DCIR.

At the amino acid sequence level, duck DCIR and DCAR2 were compared and found to share 60.8% sequence identity within the whole gene. As shown in Figure 2-10, most of the mismatches were located at the cytoplasmic domain, transmembrane domain and half of the stalk region. Their CRD domains were consistent in length and share 80% identity at amino acid level. A 15 aa shorter cytoplasmic domain of duck DCAR2 as compared to DCIR was also observed, which agrees with what was found in rodents. The highly dissimilar transmembrane domain suggests different functions for this region in the two proteins. Interestingly, it is apparent that the stalk region of DCAR is 28 aa longer than that of DCIR. Considering the sequence of other APLEC lectins, they all contain a stalk region with a length less than 40 aa. However, duck DCAR2 has an encoded stalk region comprised of two exons with a length of 65 aa. In Figure 2-10, putative calcium and sugar binding residues were also found and labeled. An EPS motif, which was thought to be a potential sugar-binding motif, was conserved between two genes. Calcium binding site 2 residues (as indicated by triangles in Figure 2-10) were found in both genes at the conserved position.

A phylogenic tree for the CRD domain only was constructed and shown in Figure 2-11. The whole molecule comparison showed that duck DCIR and DCAR

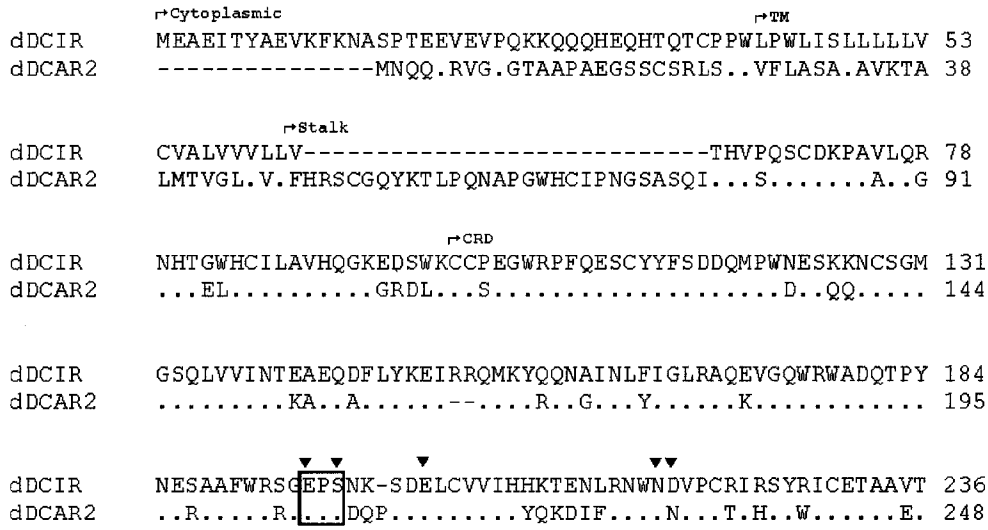


Figure 2-10. Alignment of putative amino acid sequences of duck DCIR and DCAR2. Amino acid identities were shown with dots. Conserved calcium/sugar binding sites were indicated with triangles above. The EPS motif, believed to determine the specificity of sugar binding, was boxed.

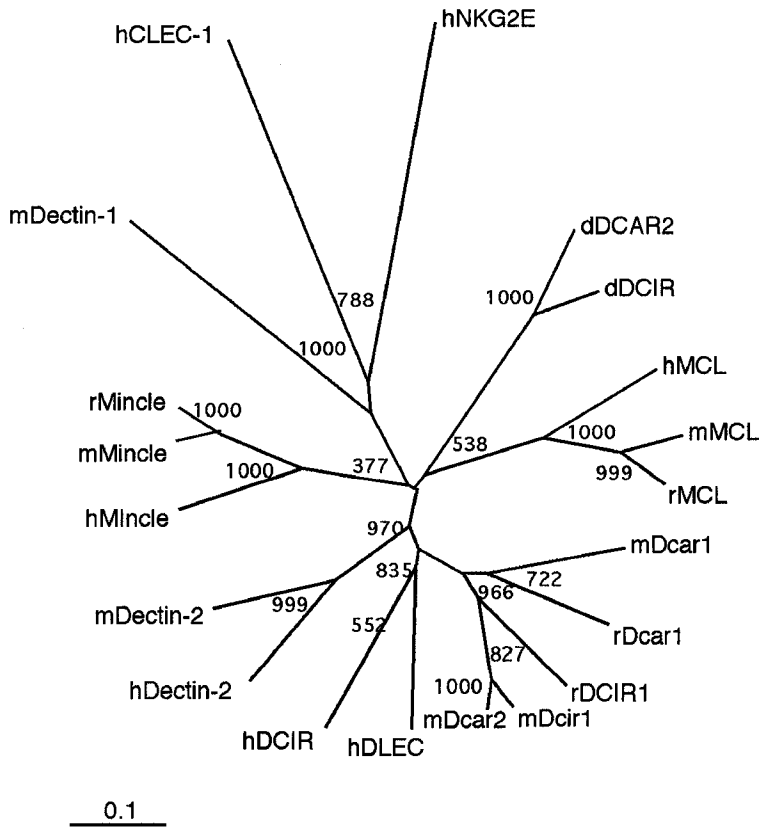


Figure 2-11. Phylogenic tree constructed based on CRD domain of APLEC/NKC lectins in various species. Comparison was made based on the amino acid sequences of the CRD domain. The bootstrap values out of 1,000 replicates are shown.

are closer to NKC lectins (NKG2E, Dectin-1 and CLEC-1)(data not shown). However, as indicated by the phylogenic tree, this association does not hold for the CRD domain comparison. The CRD domains of duck DCIR and DCAR are grouped with the mammalian MCL. This result suggested that the avian and mammalian APLEC complex genes evolved differently from a common ancestral gene.

3.7 Analysis of probe specificity

The high sequence similarity, especially in the CRD domain, between duck DCIR and DCAR genes lead us to question the specificity of our previously used overgo probes. To address this question, an alignment of probes and gene sequences was done. As shown in Figure 2-12A, the DCIR overgo probe is specific, by displaying a gap and several nucleotide mismatches compared to both DCAR genes. However, the DCAR specific overgo probe has only one nucleotide mismatch compared to DCIR gene (Figure 2-12B). In addition, it also displayed high sequence identity to an intron region between DCIR exon I and II. Considering the cross-hybridization of DCAR overgo probe, this explains why all the DCIR containing fragments can hybridize to both DCIR and DCAR probes. New DCAR probes have been designed, based on the region outside the CRD and will be used in an expression profile experiment.

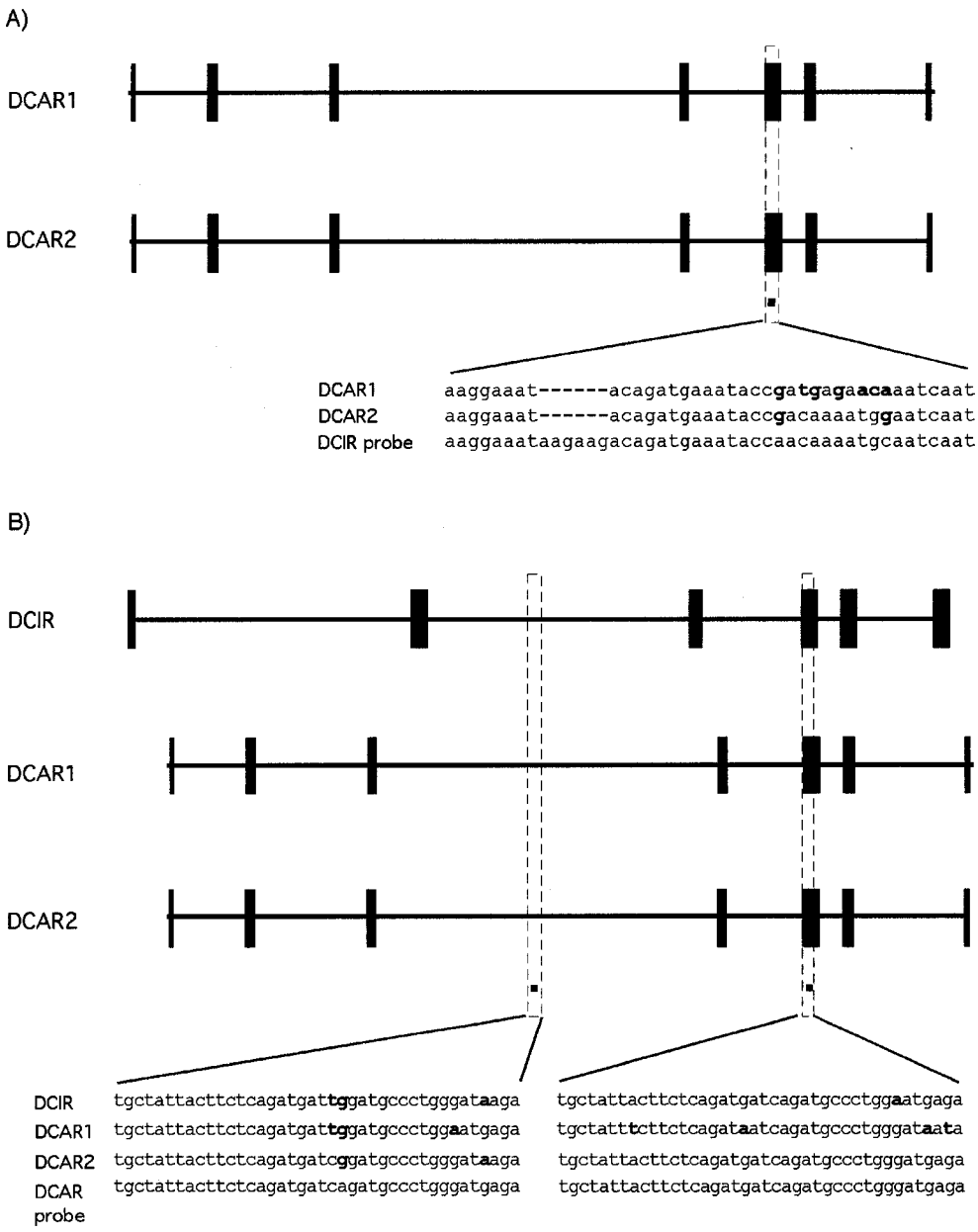


Figure 2-12. A diagram showing the cross-hybridization of DCIR (A) and DCAR (B) overgo probes. Solid black bars represent the location of exons, and lines indicate intron regions. Probes are identical with a square and regions hybridizing indicated with a box with dotted lines. The nucleotide mismatches are labeled in bold.

3.8 Alternative spliced DCIR transcripts

To search for other expressed isoforms of DCIR and DCAR as found in human and mouse, a duck spleen cell cDNA library was screened. Five DCIR positive clones were isolated, representing two isoforms of DCIR. A schematic comparison between the two isoforms is shown in Figure 2-13A. Four of the five transcripts represented the same new short isoform of DCIR, which was 120bp shorter than the full length transcript 14B7. Alignment with 14B7 showed that the DCIR variant (DCIRv) lacked exon II, which encodes the transmembrane domain of the receptor, as indicated in Figure 2-13. DCIRv was identical to 14B7 in all other parts. At the genomic sequence level, the alternative splicing of exon II may result from a poor donor site at the beginning of exon II. The deletion of the transmembrane domain is predicted to result in the production of a cytoplasmic protein instead of a membrane bound receptor.

3.9 DCAR-like transcripts

In addition to 203N10, the full length transcript of DCAR2, six other DCAR2-like transcripts were identified through screening of the duck cDNA library. DCAR2 cDNA and the other six cDNA sequences were aligned, as shown in Figure 2-14. According to the sequence similarity and gene structure, we divided the six transcripts into three groups. 224D17 and 10O10 represent the first group, encoding the same polypeptide with 98% nucleotide identity to DCAR2. It

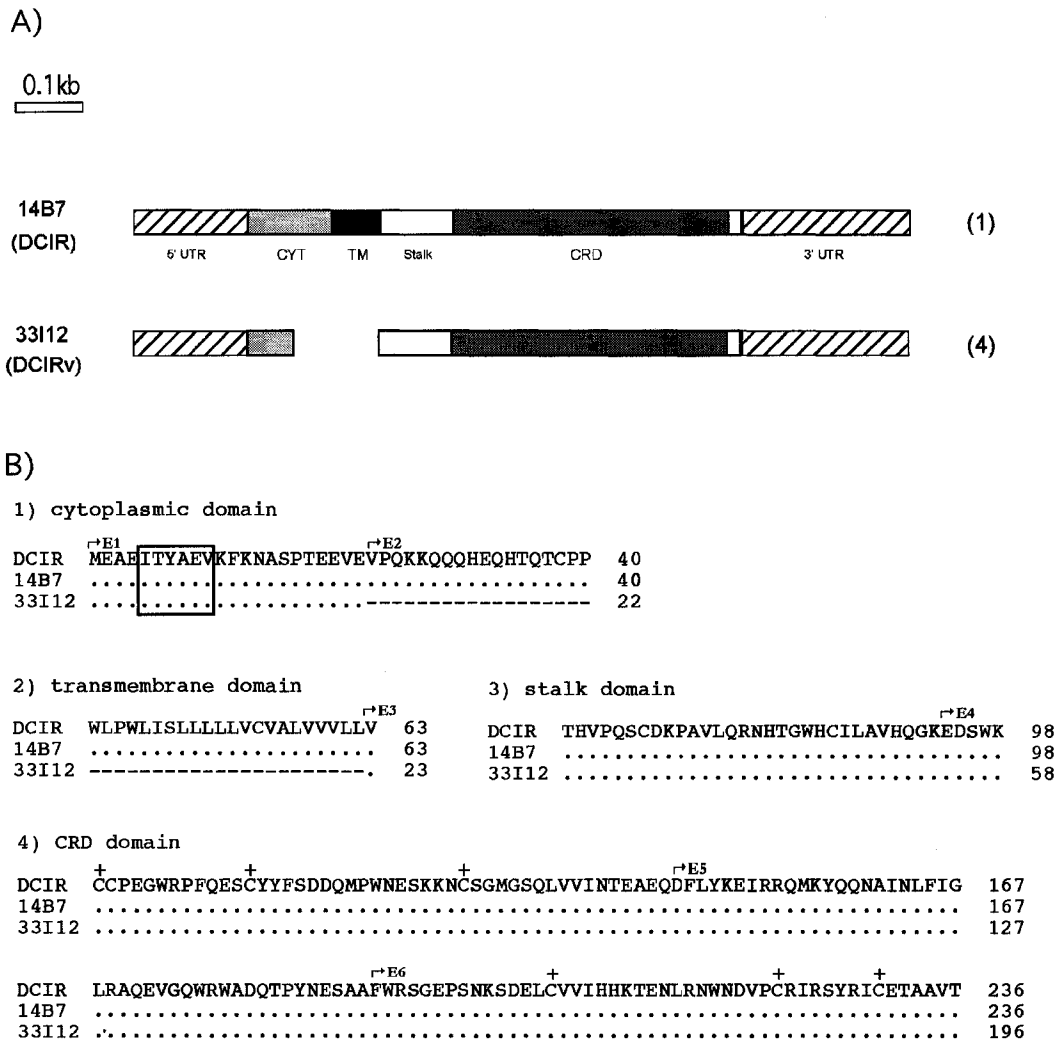
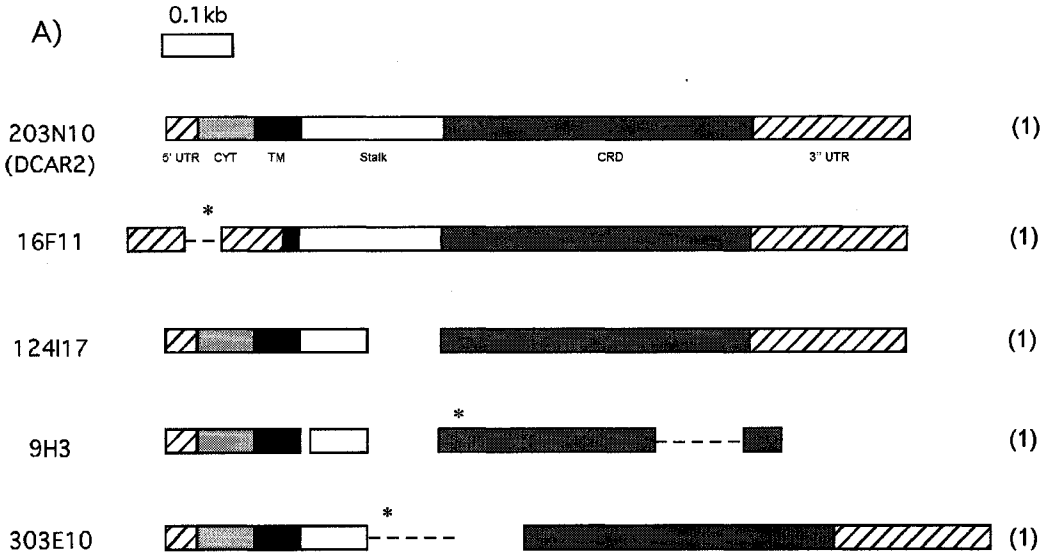


Figure 2-13. Schematic view (A) and sequence alignment (B) of the duck DCIR transcript and an alternatively spliced variant. 33I12 is a DCIR variant lacking Exon II, which encodes the transmembrane domain. The six conserved cysteine residues are labeled (+). The featured ITIM motif at the cytoplasmic domain of DCIR was boxed. The number of clones with each structure was shown in parentheses.

A)



B)

1) cytoplasmic domain

```
                ↱E1          ↱E2
DCAR2     MNQQERVGPGTAAPAEGSSCSRLSP  25
203N10    .........................  25
16F11     -------------------------  25
243K13    .........................  25
10O10     .........................  25
124I17    .......S.R................  25
9H3       .......S.R................  25
DCAR1     .........................  25
303E10    ...E...S...................  25
```

2) transmembrane domain

```
                                  ↱E3
DCAR2     WVFLASALAVKTALMTVGLVV  46
203N10    ......|..|.............  46
16F11     ------|--|-------------  7
243K13    ....V.|..|.............  46
10O10     ....V.|..|.............  46
124I17    ..L.I.|..|.............  46
9H3       ..L.I.|..|.........--  44
DCAR1     ..L.V.|..|.............  46
303E10    ......|..|.............  46
```

3) stalk region

```
                                                  ↱E4                                                  ↱E5
DCAR2     LFHRSCGQYKTLPQNAPGWHCIPNGSASQITHVSQSCDKPAALQGNHTELHCILAVHQGKGRDLK  111
203N10    ..................................................................  111
16F11     ..................................................................  72
243K13    ................SE..............................................W.  111
10O10     ................SE..............................................W.  111
124I17    ....I..........SE....TS.P.....------------------------------KDSW.  76
9H3       ---GSVASTRLC.RMLQSGTASPVDLQATR.AG.AVQRAGDPFRKAAITSQM.RCPGIRAS.TAV  106
DCAR1     ..........................................W.......V...P.........  111
303E10    ............E..Q......G.....TPQLQS.SQI*-------------------------  85
```

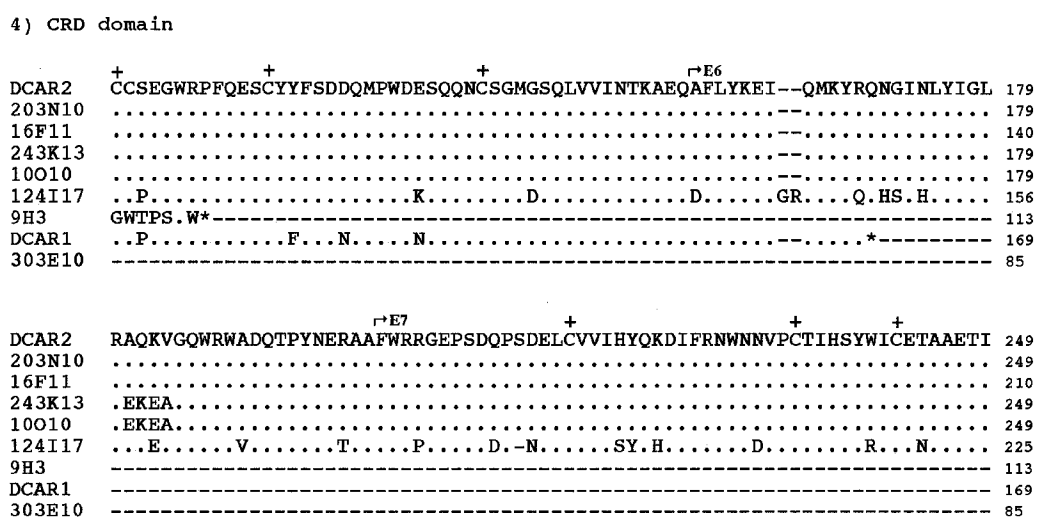


Figure 2-14. Schematic view (A) and sequence alignment (B) of duck DCAR cDNA and its alternatively spliced variants. In (A), the number of each clone was showed in parenthesis. The asterisks indicate the stop codon. In addition to the cDNA (203N10) that corresponded to the duck DCAR2 gene, five other transcript variants (16F11, 124I17, 9H3, 10O10 and 303E10) were also identified in the cDNA library. Their structures are depicted here and detailed description is given in the results.

is highly possible that these two isoforms were transcribed from the DCAR2 allele. 16F11 represents the second group, which is considered an alternative splicing form of DCAR2. 16F11 displays 100% nucleotide identity to the DCAR2 transcript, except for a 60 bp long incorrectly incorporated intron insertion at its 5 UTR. This unexpected insertion introduces an inframe stop codon, and results in a putative polypeptide lacking all of the cytoplasmic domain and part of the transmembrane domain.

Group three contains three transcripts, 124I17, 9H3 and 303E10, and featured more sequence discrepancy from the sequence of DCAR2. All these three isoforms lacked exon IV, resulting in transcripts with half of the original stalk region. In addition to one exon deletion, 124I17 has a two amino acid insertion, a one amino acid deletion in the CRD domain, as well as 33 amino acid mismatches. 9H3 was a DCAR-like transcript, identified from the previous EST project. As shown in the alignment (Figure 2-14B), its cDNA sequence was more similar to 124I17. Different from 124I17, 9H3 contains a 187 bp long incorrectly incorporated intron region between Exon VI and VII. It also showed a 16 bp deletion at the beginning of Exon III, which caused the frame-shift of the reading frame. This frame-shift subsequently leads to an early stop codon near the end of Exon V, which makes the polypeptide lose two CRD encoding exons. The transcript stopped in the middle of Exon VII, followed by a poly-A tail. 303E10 shared the highest nucleotide sequence identity to 124I17 (90%). Except for an Exon IV deletion, it contained a 146 bp intron region between Exon III and IV,

which included an in-frame stop codon and gives rise to the early termination of translation at the beginning of Exon IV.

Table 2-2 summarized the location, length and exon-intron boundary sequences of the deletion or insertion of four DCAR2-like transcripts. The table clearly shows that four out of the six deletion/insertion events happened at exon 3 and 4 region.

3.10 Expression construct of DCIR for binding assay

To facilitate characterization of the function of DCIR, a eukaryotic expression vector was constructed. dDCIR 14B7.2 ORF was subcloned into the eukaryotic vector pCI-neo. The insertion was confirmed by DNA sequencing. However, due to the fact that there is still no ligand identified for mammalian DCIR and DCAR, as well as difficulty of duck cell classification, the originally proposed binding assay was put on hold at this time.

clone number	splicing events	Position	length (bp)	sequence identity (%)	donor sequence	acceptor sequence	effect
16F11	insertion	1/3 of Intron1	60	100 to DCAR2	ag/gtgagt	cag/g	in frame stop codon
303E10	insertion	1/10 of intron 3	146	94.4 to DCAR1	ag/gtaagt	---/c	intron insertion, in frame stop codon
	deletion	exon 4	93	-	-	cag/t	exon4 deletion
124I17	deletion	exon 4	93	-	-	cag/t	exon4 deletion
9H3	deletion	exon 4	93	-	-	cag/t	exon4 deletion
	deletion	n1-16 of Exon3	16	100 to DCAR2	-	cag/g	17bp missing, frameshift
	insertion	3/4 of Intron6	187	93.6 to DCAR1	-	---/g	intron insertion

Table 2-2. A summary of the deletion or insertion events in DCAR2-like transcripts.

CHAPTER IV. DISCUSSION

In this work, we identified the gene structure and expression profile of two C-type lectin receptors in the duck, DCIR and DCAR2. Both receptors contain a conserved CRD domain, belonging to the type II C-type lectin family. DCIR has an ITIM motif in its cytoplasmic domain, which is presumed to confer an inhibitory role. Whereas DCAR does not show any significant motifs except for the presence of a positively charged amino acid residue (Lysine) in its transmembrane domain, which was believed to elicit immune responses through non-covalent binding to an adaptor protein. The gene location, gene structure and immune and mucosal tissue expression profile of the two novel genes, all indicated their roles in regulating the activity of antigen presenting cells and mediating downstream immune reactions.

4.1 Organization of the duck APLEC genomic region

We obtained four C-type lectin containing genomic clones, of which the clone 110O1 was fully sequenced, revealing one DCIR and two DCAR genes. The other three positive fosmid clones, 140L23, 205H3 and 260J15 may come from the same or the other allele as 110O1. This genomic locus represents the APLEC C-type lectin gene cluster in ducks.

DCIR and DCAR were previously identified in rodents. So far, four DCIRs and two DCARs were found in both mouse and rat (Flornes *et al.*, 2004). Rat Dcar2 was proposed to be a pseudogene or a gene fragment, because only two exons were detected. In contrast to the rodents, only one DCIR gene was identified in the human genome and no human DCAR gene was discovered (Bates *et al.*, 1999; Flornes *et al.*, 2004). In the case of chickens, the closest relatives of ducks, no DCIR and DCAR genes were identified in the EST database or genome, indicating that chickens may have lost these genes in evolution or they has not been well characterized yet.

In parallel to the genomic sequencing study, a previous genomic Southern blot using DCIR and DCAR random priming probes showed 1 to 3 hybridizing bands to DCIR and 1 to 5 bands to DCAR (Branton, 2004). According to the genomic sequence data presented here, both the DCIR and the DCAR probe used in his genomic Southern blot were not specific. In addition, the multiple hybridizing bands may come from other gene loci or arise from the other allele.

The sequence information on the other three genomic clones may help us identify new gene locus flanking the 110O1 genomic fragment. As shown in Fig.7, both 140L23 and 205H3 contain a long fragment more than 12 kb length, upstream of 110O1. Southern blots showed that this long fragment hybridized to neither the DCIR nor DCAR probe. At the downstream side of the 110O1 fragment, genomic clone 260J15 had a longer DCAR2 containing fragment with a 1.4kb short fragment attached to it.

The gene orientation of the APLEC region in the duck, with the DCIR and two DCARs in the same transcriptional orientation is also conserved in rodents (Flornes *et al.*, 2004). However, with regard to the order of genes, different patterns were observed. In rodents that have multiple DCIR and DCAR loci, the genes are arranged in the format as Dcir4-Dcir3-Dcir2-Dcar2-Dcir2-Dcar1 (Flornes *et al.*, 2004). DCAR genes are not continuous in this format and it is believed that a recent gene duplication involving the chromosomal region of Dcir1 and Dcar1 contributed to this pattern. Nevertheless, in the case of duck, at least two DCAR genes were found head to tail. The variation on gene organization suggests that the genes in the APLEC cluster of duck underwent a different mechanism of gene duplication, with no DCIR locus involved. The scenario in humans provides a more diversified example, showing only one DCIR and no DCAR in the APLEC region. It remains to be determined what forces drove the duplication of this region in some species. Possibly the receptors are evolving in response to different pathogens.

With the assistance of computer based searching (TESS), we identified the transcription factor binding motifs that were located at the 5 flanking region of both DCIR and DCAR2 genes. Strikingly, there is no TATA box found in the promoter region of both DCIR and DCAR2.

The promoter activity of another mammalian APLEC lectin, Dectin-2, was intensively studied taking the advantage of the luciferase expression assay. In the upstream region of Dectin-2, the -123 to -36 bp region of the promoter had the

highest activity for transcriptional regulation and cell expression specificity (Bonkobara *et al.*, 2001). The binding sites for interferon regulatory elements (IREs), GM-CSF, Ap-1 were detected in this region, which is in agreement with our promoter analysis of duck lectins. For Dectin-2, the stimuli from above factors were only effective in certain DC subsets, inferring that the activity of regulatory elements is cell specific. In addition, the promoter sequence of a macrophage restricted C-type lectin, MCL, was also analyzed with a computer searching program. Similar to duck lectin promoters, the binding motifs for AML, IRF, c-Ets and C/EBP were detected (Balch *et al.*, 2002b). No TATA box was found in MCL either.

Future studies could include the following aspects. First, to test the activity of regulatory regions, we would clone fragments of the known 5 flanking region of DCIR or DCAR2 genes into a construct for the luciferase reporter system. The location of enhancer and repressor regions might be found in this way. Secondly, based on the predicted transcription factor binding sites, we may use the corresponding stimuli on transfected cells and test the expression level of receptor transcripts. In addition, we could determine the potential transcriptional initiation sites with the help of a primer extension experiment.

4.2 Structure of duck DCIR and DCAR relevant to function

The genomic structure of DCIR was obtained by aligning genomic DCIR with its cDNA transcript 14B7. From the results, we found that duck DCIR has six exons, with the first three encoding the cytoplasmic, transmembrane and stalk region, respectively. Moreover, the last three exons encode the CRD domain, which shares over 40 % homology to human and rodent DCIR. The cytoplasmic domain of duck DCIR is relatively long, carrying a canonical ITIM motif and is believed to confer inhibitory signals. DCIR is the only identified APC expressed C-type lectin bearing an ITIM motif. An EPS motif has been found at the conserved site in the CRD domain, which may mediate sugar recognition by the receptor.

Thus far, one DCIR gene was identified in humans and four in rodents, with human DCIR showing over 50% similarity to rodent (Kanazawa *et al.*, 2002). The sequence and location of the ITIM motif are conserved in human and rodent DCIR1 and 2, but diversified in mouse and rat DCIR3 and 4 (Flornes *et al.*, 2004). The exon number, distribution and size of mammalian DCIR are similar to duck. The EPS motif is conserved in most mammalian and duck DCIR, except in rodent DCIR2, while an EPN motif was found at the conserved site.

In an attempt to investigate the inhibitory function of mammalian DCIR, a chimeric receptor was constructed by fusing the cytoplasmic and transmembrane domain of DCIR to the extracellular domain of the Fc receptor. After co-ligation with BCR using anti-Fc antibody, this chimeric receptor was found to inhibit BCR-mediated activation signals (Kanazawa *et al.*, 2002). Further studies have

shown that the DCIR ITIM motif may take its effect by recruiting two tyrosine phosphatases SHP-1 (Huang *et al.*, 2001) and SHP-2 (Kanazawa *et al.*, 2002), which remove the phosphate group from the triggered tyrosine and terminate the subsequent activating signaling pathway. Since duck DCIR has a similar inhibitory motif, we predict it may have a similar function.

Unlike duck DCIR and all other APLEC lectins, the functional duck DCAR2 had seven exons, with an extra exon (III) encoding a longer stalk region. The CRD domain of duck DCAR shares 36 % homology to rodent DCARs. The cytoplasmic domain of duck DCAR is very short and lacked any signaling motifs. Its transmembrane domain contained a positively charged lysine residue, which is believed to form a non-covalent bond with the adaptor proteins. Similar to DCIR, an EPS motif was found at the conserved site in its CRD domain.

So far, most of the known APLEC lectins have the same exon numbers and distribution, with exon I encoding the cytoplasmic region, exon II encoding the transmembrane domain around 25aa in length, and exon III encoding the stalk region that is usually less than 40aa. However, duck DCAR2 provides the only exception, displaying a 63 aa long stalk region, which is encoded by two exons (exon III and IV). The amino acid sequence of DCAR2 Exon IV is similar to DCIR Exon III. The function of this long stalk region is unknown. Common to all known APLEC lectins, the CRD domain is encoded by the last three exons, usually Exon IV-VI.

It is worth notice that a lysine (K) residue appears in the duck DCAR2 transmembrane domain, which is consistent with the finding of another positively charged residue, arginine (R), in rodent DCARs. It has been determined that this arginine residue was responsible for the binding of the receptor with an ITAM containing adaptor protein, FcR γchain. This adaptor protein functions to significantly increase the membrane expression of DCAR, as well as to trigger the BCR mediated activating signals (Kanazawa *et al.*, 2003). In the case of duck DCAR2, it is still to be determined whether the lysine residue in duck DCAR2 can confer a similar activating role. In addition, an immunoprecipitation experiment may help identify the adaptor protein(s).

4.3 Sugar binding structure

Duck DCIR and DCAR 2 shared over 80% sequence identity in their CRD domain, and conservation of their calcium/sugar binding residues. Even though the ligands of both receptors are still unknown, their highly similar CRD domains indicated that they might recognize the same or similar ligands.

Although not much information concerning the structure of C-type lectins is available, the criterion concluded from the well-studied ones such as MBP will aid us to predict the basic structure of duck DCIR and DCAR. A previous C-type lectin model showed that calcium binding is necessary for carbohydrate binding since it acts as a bridge between the receptor and carbohydrate (Weis and

Drickamer, 1996). The receptor, calcium and sugar form a complex through the coordination bonds formed between calcium and carbohydrate, as well as the hydrogen bond formed between calcium and the side chain of receptor amino acid (Weis and Drickamer, 1996). For example, the presence of five conserved amino acids in the CRD domain, two asparagines, two glutamic acids and one aspartic acid, are necessary for mannose binding. Whereas two amino acid residues are distinct for galactose-specific binding, one asparagine position was replaced by an aspartic acid and one glutamic acid is replaced by a glutamine (Drickamer, 1993). A more striking model described recently is TL14, a lectin from a tunicate, *Polyandrocarpa misakiensis,* in which only four calcium binding amino acid residues are required. In an EPS motif, the serine hydroxyl and a water molecule form hydrogen bonds with sugar hydroxyl groups (Drickamer, 1999; Poget *et al.*, 1999). Interestingly, no matter which amino acid residues are involved in the complex, three continuous amino acid residues always constitute the sugar-binding region, in which a proline must be present in the middle.

In duck DCIR and DCAR, we found an EPS motif located at the established carbohydrate binding position. The glutamic acid within this motif, together with the other two asparagines and one aspartic acid, may be involved in the formation of the receptor-calcium-sugar complex. Moreover, a serine and a water molecule may interact with sugar directly. However, the EPS motif is not conserved across species. The alignment of CRD domains of DCIR and DCAR orthologs of human, rat and mouse showed that the serine was replaced by an

asparagine in rodent DCAR1 and DCIR1. It is still unknown whether this change will alter the sugar binding specificity. Nevertheless, the sequence of the carbohydrate binding site suggests that the duck DCIR and DCAR may recognize the sugar structure similar to D-Galactose, the ligand of tunicate C-type lectin TL14.

4.4 The significance of alternative splicing of DCIR

In the duck spleen cDNA library, five DCIR positive transcripts were detected and sequenced. One of them was identical to the previously identified duck DCIR gene, whereas the other four represent the same alternative splicing form of the gene, with the transmembrane domain deleted.

Alternative splicing forms have also been reported in human DCIR, with three detected variants, which lacked exon II, exon III or both (Bates *et al.*, 1999; Huang *et al.*, 2001; Thibault *et al.*, 2003). With regard to the tissue distribution, the isoform with a transmembrane (Exon II) deletion had similar tissue expression profile as the full size transcript (Bates *et al.*, 1999). The isoform that lacks the stalk region (Exon III) was also detected by northern blot in neutrophils (Thibault *et al.*, 2003).

Since the duck DCIR variant lacks the exon encoding the transmembrane domain, it should not be expressed on the surface of the cell. To test this hypothesis, we could construct a cell expression system for both the full length

and variant form of duck DCIR. In order to efficiently detect their expression using antibodies, it is necessary to attach an in-frame expression tag. Whether there is surface expression of either form would be determined by flow cytometry.

A binding assay that tests the specific binding of DCIR transfected cells with various duck cells, also could be conducted in the future. Does the ITIM in the soluble form work in the same way as in membrane bound form? Can the soluble form recognize some cytosolic antigens? The answers of these questions will expand our understanding of DCIR-mediated immunoregulatory mechanisms.

4.5 Duck DCAR2-like transcripts

Compared to DCIR, duck DCAR2 has a complicated pool of isoforms. In addition to 203N10, the transcript corresponding to DCAR2, we identified six other DCAR2-like transcripts in the duck cDNA library. According to the sequence identity and gene structures, we categorized the six isoforms into three groups. The first group (243K13 and 10O10) might be derived from the other DCAR2 allele, since they displayed only minor amino acid mismatches and the same gene structure. Group two (16F11) is an alternative splicing form of DCAR2, showing 100% nucleotide sequence identity, with only one mis-spliced intron region. The three members in groups three, 124I17, 9H3 and 303E10, might be the products of other unidentified DCAR genes. Two pieces of evidence

are present to support this conclusion. First, these three members share relatively low sequence identity to DCAR2 at nucleotide level, and second, gaps were found within these sequences.

Non-classical exon-intron boundary sequences may contribute to alternative splicing. From the genomic sequence, we found that the sequence at the intron 2 acceptor site is $(c/t)_n$cag/T. In comparison, a more classical splice site appeared early in exon III of 9H3, $(c/t)_n$cag/G, and resulted in the deletion of the first 16 nucleotides. Likewise, the irregular acceptor sequence of intron 3 was $(c/t)_n$cag/T, which may be the reason for Exon IV deletion in three transcripts. The other three intron insertions could have resulted from the classical exon-intron boundary sequence present in the intron region. This alternative splicing issue is still quite open, since the genomic sequences corresponding to 124I17, 9H3 and 303E10 have not been examined.

In rodents, one DCAR alternatively spliced form was identified, in which the whole stalk region was deleted (Kanazawa *et al.*, 2003). This result is consistent with what we have found in duck, where three out of four transcripts skip the stalk region. It is interesting to question why the stalk region is subject to such highly variable splicing. One postulation is that the stalk region of DCAR hosts some important functions, for instance forming a dimer, and the alternative splicing of this region may change or eliminate this function. It is also possible that the DCAR2 transcript (203N10) itself is an alternative spliced form, since it contains two exons encoding an exceptionally long stalk region. Further studies of

other DCAR like genes may provide more clues about the functional significance of the stalk region.

In addition to DCIR and DCAR, alternative splicing is ubiquitous among members of the C-type lectin family. The other two APLEC genes, Dectin-2 (Ariizumi *et al.*, 2000a) (Gavino *et al.*, 2005) and BDCA-2 (Dzionek *et al.*, 2001), as well as Dectin-1 (Willment *et al.*, 2001), Langerin (JID) and MCL (Arce *et al.*, 2004), all have alternatively spliced forms. Interestingly, for most of the above C-type lectin receptors, at least one isoform skips either the transmembrane or stalk region, with rare cases of the deletion of CRD encoding exons. One explanation to this is that the CRD domain encodes the region that is responsible for the maintenance of the extracellular structure and the recognition of ligands, any mis-splicing in this region may make the whole molecule inactive. In contrast, the change at transmembrane and stalk region may only change the cellular localization feature of the molecule, while keeping the ligand recognition function.

REFERENCES

Adachi, Y., Ishii, T., Ikeda, Y., Hoshino, A., Tamura, H., Aketagawa, J., Tanaka, S. and Ohno, N. (2004). Characterization of beta-glucan recognition site on C-type lectin, Dectin 1. Infection and Immunity. **72** (7): 4159-4171.

Alvarez, C. P., Lasala, F., Carrillo, J., Muniz, O., Corbi, A. L. and Delgado, R. (2002). C-type lectins DC-SIGN and L-SIGN mediate cellular entry by Ebola virus in cis and in trans. Journal of Virology. **76** (13): 6841-6844.

Aragane, Y., Maeda, A., Schwarz, A., Tezuka, T., Ariizumi, K. and Schwarz, T. (2003). Involvement of dectin-2 in ultraviolet radiation-induced tolerance. Journal of Immunology. **171** (7): 3801-3807.

Arce, I., Martinez-Munoz, L., Roda-Navarro, P. and Fernandez-Ruiz, E. (2004). The human C-type lectin CLECSF8 is a novel monocyte/macrophage endocytic receptor. European Journal of Immunology. **34** (1): 210-220.

Ariizumi, K., Shen, G. L., Shikano, S., Ritter, R., Zukas, P., Edelbaum, D., Morita, A. and Takashima, A. (2000a). Cloning of a second dendritic cell-associated C-type lectin (Dectin-2) and its alternatively spliced isoforms. Journal of Biological Chemistry. **275** (16): 11957-11963.

Ariizumi, K., Shen, G. L., Shikano, S., Xu, S., Ritter, R., Kumamoto, T., Edelbaum, D., Morita, A., Bergstresser, P. R. and Takashima, A. (2000b). Identification of a novel, dendritic cell-associated molecule, dectin-1, by subtractive cDNA cloning. Journal of Biological Chemistry. **275** (26): 20157-20167.

Baekkevold, E. S., Yamanaka, T., Palframan, R. T., Carlsen, H. S., Reinholt, F. P., von Andrian, U. H., Brandtzaeg, P. and Haraldsen, G. (2001). The CCR7

ligand ELC (CCL19) is transcytosed in high endothelial venules and mediates T cell recruitment. Journal of Experimental Medicine. **193** (9): 1105-1112.

Balch, S. G., Greaves, D. R., Gordon, S. and McKnight, A. J. (2002). Organization of the mouse macrophage C-type lectin (Mcl) gene and identification of a subgroup of related lectin molecules. European Journal of Immunogenetics. **29** (1): 61-64.

Balch, S. G., McKnight, A. J., Seldin, M. F. and Gordon, S. (1998). Cloning of a novel C-type lectin expressed by murine macrophages. Journal of Biological Chemistry. **273** (29): 18656-18664.

Banchereau, J., Briere, F., Caux, C., Davoust, J., Lebecque, S., Liu, Y.-J., Pulendran, B. and Palucka, K. (2000). Immunobiology of dendritic cells. Annual Review of Immunology. **18** (1): 767-811.

Barth, H., Ulsenheimer, A., Pape, G. R., Diepolder, H. M., Hoffmann, M., Neumann-Haefelin, C., Thimme, R., Henneke, P., Klein, R., Paranhos-Baccala, G., Depla, E., Liang, T. J., Blum, H. E. and Baumert, T. F. (2005). Uptake and presentation of hepatitis C virus-like particles by human dendritic cells. Blood. **105** (9): 3605-3614.

Bates, E. E. M., Fournier, N., Garcia, E., Valladeau, J., Durand, I., Pin, J. J., Zurawski, S. M., Patel, S., Abrams, J. S., Lebecque, S., Garrone, P. and Saeland, S. (1999). APCs express DCIR, a novel C-type lectin surface receptor containing an immunoreceptor tyrosine-based inhibitory motif. Journal of Immunology. **163** (4): 1973-1983.

Biessen, E. A. L., van Teijlingen, M., Vietsch, H., Barrett-Bergshoeff, M. M., Bijsterbosch, M. K., Rijken, D. C., van Berkel, T. J. C. and Kuiper, J. (1997). Antagonists of the Mannose receptor and the LDL receptor related protein

dramatically delay the clearance of tissue plasminogen activator. Circulation. **95** (1): 46-52.

Bonifaz, L. C., Bonnyay, D. P., Charalambous, A., Darguste, D. I., Fujii, S.-I., Soares, H., Brimnes, M. K., Moltedo, B., Moran, T. M. and Steinman, R. M. (2004). In vivo targeting of antigens to maturing dendritic cells via the DEC-205 receptor improves T cell vaccination. Journal of Experimental Medicine. **199** (6): 815-824.

Bonkobara, M., Hoshino, M., Yagihara, H., Tamura, K., Isotani, M., Tanaka, Y., Washizu, T. and Ariizumi, K. (2006). Identification and gene expression of bovine C-type lectin dectin-2. Veterinary Immunology and Immunopathology. **110** (1-2): 179-186.

Bonkobara, M., Zukas, P. K., Shikano, S., Nakamura, S., Cruz, P. D. and Ariizumi, K. (2001). Epidermal Langerhans cell-targeted gene expression by a dectin-2 promoter. Journal of Immunology. **167** (12): 6893-6900.

Brown, G. D. and Gordon, S. (2001). Immune recognition - A new receptor for beta-glucans. Nature. **413** (6851): 36-37.

Coggeshall, K. M. (1998). Inhibitory signaling by B cell Fc[gamma]RIIb. Current Opinion in Immunology. **10**(3): 306-312.

Colonna, M., Samaridis., J. and Angman, L. (2000). Molecular characterization of two novel C-type lectin-like receptors, one of which is selectively expressed in human dendritic cells. European Journal of Immunology. **30** (2): 697-704.

Cyster, J. G. (1999). Chemokines and the Homing of Dendritic Cells to the T Cell Areas of Lymphoid Organs. Journal of Experimental Medicine. **189** (3): 447-450.

Daeron, M., Latour, S., Malbec, O., Espinosa, E., Pina, P., Pasmans, S. and Fridman, W. H. (1995). The same tyrosine-based inhibition motif, in the intra-

cytoplasmic domain of Fc[gamma]RIIB, regulates negatively BCR-, TCR-, and FcR-dependent cell activation. Immunity. **3** (5): 635-646.

Dissen, E., Ryan, J., Seaman, W. and Fossum, S. (1996). An autosomal dominant locus, Nka, mapping to the Ly-49 region of a rat natural killer (NK) gene complex, controls NK cell lysis of allogeneic lymphocytes. Journal of Experimental Medicine. **183** (5): 2197-2207.

Douillard, P., Stoitzner, P., Tripp, C. H., Clair-Moninot, V., Ait-Yahia, S., McLellan, A. D., Eggert, A., Romani, N. and Saeland, S. (2005). Mouse lymphoid tissue contains distinct subsets of Langerin/CD207+ dendritic cells, only one of which represents epidermal-derived Langerhans cells. Journal of Investigative Dermatology. **125** (5): 983-994.

Drickamer, K. (1988). Two distinct classes of carbohydrate-recognition domains in animal lectins. Journal of Biological Chemistry. **263** (20): 9557-9560.

Drickamer, K. (1993). Ca2+-dependent carbohydrate-recognition domains in animal proteins. Current Opinion in Structural Biology. **3** (3): 393-400.

Drickamer, K. (1999). C-type lectin-like domains. Current Opinion in Structural Biology. **9** (5): 585-590.

Drickamer, K. and Taylor, M. E. (1993). Biology of animal lectins. Annual Review of Cell Biology. **9** (1): 237-264.

Dzionek, A., Fuchs, A., Schmidt, P., Cremer, S., Zysk, M., Miltenyi, S., Buck, D. W. and Schmitz, J. (2000). BDCA-2, BDCA-3, and BDCA-4: Three markers for distinct subsets of dendritic cells in human peripheral blood. Journal of Immunology. **165** (11): 6037-6046.

Dzionek, A., Sohma, Y., Nagafune, J., Cella, M., Colonna, M., Facchetti, F., Gunther, G., Johnston, I., Lanzavecchia, A., Nagasaka, T., Okada, T., Vermi, W.,

Winkels, G., Yamamoto, T., Zysk, M., Yamaguchi, Y. and Schmitz, J. (2001). BDCA-2, a novel plasmacytoid dendritic cell-specific type II C-type lectin, mediates antigen capture and is a potent inhibitor of interferon α/β induction. Journal of Experimental Medicine. **194** (12): 1823-1834.

East, L. and Isacke, C. M. (2002a). The mannose receptor family. Biochimica et Biophysica Acta (BBA) - General Subjects. **1572** (2-3): 364-386.

East, L., Rushton, S., Taylor, M. E. and Isacke, C. M. (2002b). Characterization of sugar binding by the mannose receptor family member, Endo180. Journal of Biological Chemistry.**277** (52): 50469-50475.

Eichbaum, Q., Heney, D., Raveh, D., Chung, M., Davidson, M., Epstein, J. and Ezekowitz, R. A. B. (1997). Murine macrophage mannose receptor promoter is regulated by the transcription factors PU.1 and SP1. Blood. **90** (10): 4135-4143.

Fanger, N., Wardwell, K., Shen, L., Tedder, T. and Guyre, P. (1996). Type I (CD64) and type II (CD32) Fc gamma receptor-mediated phagocytosis by human blood dendritic cells. Journal of Immunology. **157** (2): 541-548.

Fernandes, M. J., Finnegan, A. A., Siracusa, L. D., Brenner, C., Iscove, N. N. and Calabretta, B. (1999). Characterization of a novel receptor that maps near the natural killer gene complex: Demonstration of carbohydrate binding and expression in hematopoietic cells. Cancer Research. **59** (11): 2709-2717.

Figdor, C. G., van Kooyk, Y., Adema, G.J. (2002). C-Type lectin receptors on dendritic cells and langerhans cells. Nature Reviews Immunology. **2** (2): 74-84.

Flornes, L. M., Bryceson, Y. T., Spurkland, A., Lorentzen, J. C., Dissen, E. and Fossum, S. (2004). Identification of lectin-like receptors expressed by antigen presenting cells and neutrophils and their mapping to a novel gene complex. Immunogenetics. **56** (7): 506-517.

Fraser, I. P., Takahashi, K., Koziel, H., Fardin, B., Harmsen, A. and Ezekowitz, R. A. B. (2000). Pneumocystis carinii enhancessoluble mannose receptor production by macrophages. Microbes and Infection. **2** (11): 1305-1310.

Gantner, B. N., Simmons, R. M., Canavera, S. J., Akira, S. and Underhill, D. M. (2003). Collaborative induction of inflammatory responses by Dectin-1 and Toll-like receptor 2. Journal of Experimental Medicine. **197** (9): 1107-1117.

Gavino, A. C. P., Chung, J. S., Sato, K., Ariizumi, K. and Cruz, P. D. (2005). Identification and expression profiling of a human C-type lectin, structurally homologous to mouse dectin-2. Experimental Dermatology. **14** (4): 281-288.

Geijtenbeek, T. B. H., Krooshoop, D., Bleijs, D. A., van Vliet, S. J., van Duijnhoven, G. C. F., Grabovsky, V., Alon, R., Figdor, C. G. and van Kooyk, Y. (2000a). DC-SIGN-ICAM-2 interaction mediates dendritic cell trafficking. Nature Immunology. **1** (4): 353-357.

Geijtenbeek, T. B. H., Kwon, D. S., Torensma, R., van Vliet, S. J., van Duijnhoven, G. C. F., Middel, J., Cornelissen, I., Nottet, H., KewalRamani, V. N., Littman, D. R., Figdor, C. G. and van Kooyk, Y. (2000b). DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. Cell. **100** (5): 587-597.

Geijtenbeek, T. B. H., Torensma, R., van Vliet, S. J., van Duijnhoven, G. C. F., Adema, G. J., van Kooyk, Y. and Figdor, C. G. (2000c). Identification of DC-SIGN, a novel dendritic cell-specific ICAM-3 receptor that supports primary immune responses. Cell. **100** (5): 575-585.

Granelli-Piperno, A., Pritsker, A., Pack, M., Shimeliovich, I., Arrighi, J.-F., Park, C. G., Trumpfheller, C., Piguet, V., Moran, T. M. and Steinman, R. M. (2005). Dendritic Cell-Specific Intercellular adhesion molecule 3-Grabbing Nonintegrin/CD209 is abundant on macrophages in the normal human lymph

node and is not required for dendritic cell stimulation of the mixed leukocyte reaction. Journal of Immunology. **175** (7): 4265-4273.

Gunn, M. D., Tangemann, K., Tam, C., Cyster, J. G., Rosen, S. D. and Williams, L. T. (1998). A chemokine expressed in lymphoid high endothelial venules promotes the adhesion and chemotaxis of naive T lymphocytes. Proceedings of the National Academy of Sciences of the United States of America. **95** (1): 258-263.

Halary, F., Amara, A., Lortat-Jacob, H., Messerle, M., Delaunay, T., Houles, C., Fieschi, F., Arenzana-Seisdedos, F., Moreau, J. F. and Dechanet-Merville, J. (2002). Human cytomegalovirus binding to DC-SIGN is required for dendritic cell infection and target cell trans-infection. Immunity. **17** (5): 653-664.

Hanasaki, K., Yokota, Y., Ishizaki, J., Itoh, T. and Arita, H. (1997). Resistance to endotoxic shock in Phospholipase A2 receptor-deficient mice. Journal of Biological CHemistry. **272** (52): 32792-32797.

Huang, X., Yuan, Z., Chen, G., Zhang, M., Zhang, W., Yu, Y. and Cao, X. (2001). Cloning and characterization of a novel ITIM containing lectin-like immunoreceptor LLIR and its two transmembrane region deletion variants,. Biochemical and Biophysical Research Communications. **281** (1): 131-140.

Hulse-Post, D. J., Sturm-Ramirez, K. M., Humberd, J., Seiler, P., Govorkova, E. A., Krauss, S., Scholtissek, C., Puthavathana, P., Buranathai, C., Nguyen, T. D., Long, H. T., Naipospos, T. S. P., Chen, H., Ellis, T. M., Guan, Y., Peiris, J. S. M. and Webster, R. G. (2005). Role of domestic ducks in the propagation and biological evolution of highly pathogenic H5N1 influenza viruses in Asia. Proceedings of the National Academy of Sciences of the United States of America. **102** (30): 10682-10687.

Ignacio Arce, L. M.-M., Pedro Roda-Navarro, Elena FernÆndez-Ruiz, (2004). The human C-type lectin CLECSF8 is a novel monocyte/macrophage endocytic receptor. European Journal of Immunology. **34** (1): 210-220.

Inaba, K., Witmer-Pack, M., Inaba, M., Hathcock, K., Sakuta, H., Azuma, M., Yagita, H., Okumura, K., Linsley, P. and Ikehara, S. (1994). The tissue distribution of the B7-2 costimulator in mice: abundant expression on dendritic cells in situ and during maturation in vitro. Journal of Experimental Medicine. **180** (5): 1849-1860.

Iobst, S. and Drickamer, K. (1994). Binding of sugar ligands to Ca(2+)-dependent animal lectins. II. Generation of high-affinity galactose binding by site-directed mutagenesis. Journal of Biological Chemistry . **269** (22): 15512-15519.

Isakov, N. (1997). Immunoreceptor tyrosine-based activation motif (ITAM), a unique module linking antigen and Fc receptors to their signaling cascades. Journal of Leukocyte Biology. **61** (1): 6-16.

Jacob, J. P., Milne, S., Beck, S. and Kaufman, J. (2000). The major and a minor class II ˆ -chain (B-LB) gene flank the Tapasin gene in the B-F/B-L region of the chicken major histocompatibility complex. Immunogenetics. **51** (2): 138-147.

Kaiser, A., Donnadieu, E., Abastado, J.-P., Trautmann, A. and Nardin, A. (2005). CC Chemokine ligand 19 secreted by mature dendritic cells increases naive T Cell scanning behavior and their response to rare cognate antigen. Journal of Immunology. **175** (4): 2349-2356.

Kanazawa, N., Okazaki, T., Nishimura, H., Tashiro, K., Inaba, K. and Miyachi, Y. (2002). DCIR acts as an inhibitory receptor depending on its immunoreceptor tyrosine-based inhibitory motif. Journal of Investigative Dermatology. **118** (2): 261-266.

Kanazawa, N., Tashiro, K., Inaba, K., Lutz, M. B. and Miyachi, Y. (2004a). Molecular cloning of human dectin-2. Journal of Investigative Dermatology. **122** (6): 1522-1524.

Kanazawa, N., Tashiro, K., Inaba, K. and Miyachi, Y. (2003). Dendritic cell immunoactivating receptor, a novel C-type lectin immunoreceptor, acts as an activating receptor through association with Fc receptor gamma chain. Journal of Biological Chemistry. **278** (35): 32645-32652.

Kanazawa, N., Tashiro, K. and Miyachi, Y. (2004b). Signaling and immune regulatory role of the dendritic cell immunoreceptor (DCIR) family lectins: DCIR, DCAR, dectin-2 and BDCA-2. Immunobiology. **209** (1-2): 179-190.

Kang, P. B., Azad, A. K., Torrelles, J. B., Kaufman, T. M., Beharka, A., Tibesar, E., DesJardin, L. E. and Schlesinger, L. S. (2005). The human macrophage mannose receptor directs Mycobacterium tuberculosis lipoarabinomannan-mediated phagosome biogenesis. Journal of Experimental Medicine. **202** (7): 987-999.

Kato, M., Neil, T. K., Clark, G. J., Morris, C. M., Sorg, R. V. and Hart, D. N. J. (1998). cDNA cloning of human DEC-205, a putative antigen-uptake receptor on dendritic cells. Immunogenetics. **47** (6): 442-450.

Kaufman, J., Milne, S., Gobel, T. W. F., Walker, B. A., Jacob, J. P., Auffray, C., Zoorob, R. and Beck, S. (1999). The chicken B locus is a minimal essential major histocompatibility complex. Nature. **401** (6756): 923-925.

Kolatkar, A. R. and Weis, W. I. (1996). Structural basis of Galactose recognition by C-type animal lectins. Journal of Biological Chemistry. **271** (12): 6679-6685.

Lanier, L. L. (1998). NK cell receptors. Annual Review of Immunology. **16** (1): 359-393.

Largent, B., Walton, K., Hoppe, C., Lee, Y. and Schnaar, R. (1984). Carbohydrate-specific adhesion of alveolar macrophages to mannose- derivatized surfaces. Journal of Biological Chemistry. **259** (3): 1764-1769.

Lee, S. J., Evers, S., Roeder, D., Parlow, A. F., Risteli, J., Risteli, L., Lee, Y. C., Feizi, T., Langen, H. and Nussenzweig, M. C. (2002). Mannose receptor-mediated regulation of serum glycoprotein homeostasis. Science. **295** (5561): 1898-1901.

Leteux, C., Chai, W., Loveless, R. W., Yuen, C.-T., Uhlin-Hansen, L., Combarnous, Y., Jankovic, M., Maric, S. C., Misulovin, Z., Nussenzweig, M. C. and Ten Feizi (2000). The Cysteine-rich domain of the macrophage mannose receptor is a multispecific lectin that recognizes Chondroitin Sulfates A and B and sulfated Oligosaccharides of blood group Lewisa and Lewisx types in addition to the sulfated N-Glycans of Lutropin. Journal of Experimental Medicine. **191** (7): 1117-1126.

Linehan, S. A., Martinez-Pomares, L., Stahl, P. D. and Gordon, S. (1999). Mannose receptor and its putative ligands in normal murine lymphoid and nonlymphoid organs: In situ expression of mannose receptor by selected macrophages, endothelial cells, perivascular microglia, and mesangial cells, but not dendritic cells. Journal of Experimental Medicine **189** (12): 1961-1972.

Ludwig, I. S., Lekkerkerker, A. N., Depla, E., Bosman, F., Musters, R. J. P., Depraetere, S., van Kooyk, Y. and Geijtenbeek, T. B. H. (2004). Hepatitis C virus targets DC-SIGN and L-SIGN to escape lysosomal degradation. Journal of Virology **78** (15): 8322-8332.

Mahnke, K., Guo, M., Lee, S., Sepulveda, H., Swain, S. L., Nussenzweig, M. and Steinman, R. M. (2000). The dendritic cell receptor for endocytosis, DEC-205, can recycle and enhance antigen presentation via Major Histocompatibility

Complex Class II-positive lysosomal compartments. Journal of Cell Biology **151** (3): 673-684.

Marodi, L., Schreiber, S., Anderson, D. C., Macdermott, R. P., Korchak, H. M. and Johnston, R. B. (1993). Enhancement of macrophage candidacidal activity by Interferon-Gamma - Increased phagocytosis, killing, and calcium signal mediated by a decreased number of mannose receptors. Journal of Clinical Investigation. **91** (6): 2596-2601.

Martinez-Pomares, L., Reid, D. M., Brown, G. D., Taylor, P. R., Stillion, R. J., Linehan, S. A., Zamze, S., Gordon, S. and Wong, S. Y. C. (2003). Analysis of mannose receptor regulation by IL-4, IL-10, and proteolytic processing using novel monoclonal antibodies. Journal of Leukocyte Biology. **73** (5): 604-613.

Matsumoto, M., Tanaka, T., Kaisho, T., Sanjo, H., Copeland, N. G., Gilbert, D. J., Jenkins, N. A. and Akira, S. (1999). A novel LPS-inducible C-type lectin is a transcriptional target of NF-IL6 in macrophages. Journal of Immunology. **163** (9): 5039-5048.

McGreal, E. P., Rosas, M., Brown, G. D., Zamze, S., Wong, S. Y. C., Gordon, S., Martinez-Pomares, L. and Taylor, P. R. (2006). The carbohydrate recognition domain of Dectin-2 is a C-type lectin with specificity for high-mannose. glycobiology. Glycobiology. 16 (5):422-30.

Miller, M. M., Goto, R. M., Taylor, R. L., Jr., Zoorob, R., Auffray, C., Briles, R. W., Briles, W. E. and Bloom, S. E. (1996). Assignment of Rfp-Y to the chicken major histocompatibility complex/NOR microchromosome and evidence for high-frequency recombination associated with the nucleolar organizer region. Proceedings of the National Academy of Sciences of the United States of America. **93** (9): 3958-3962.

Mitchell, D. A., Fadden, A. J. and Drickamer, K. (2001). A novel mechanism of carbohydrate recognition by the C-type lectins DC-SIGN and DC-SIGNR. subunit organization and binding to multivalent ligands. Journal of Biological Chemistry. **276** (31): 28939-28945.

Moon, D. A. and Magor, K. E. (2004). Construction and characterization of a fosmid library for comparative analysis of the duck genome. Animal Genetics. **35** (5): 417-418.

Napper, C.E., Drickamer, K., and Taylor, M.E. (2006). Collagen binding by the mannose receptor mediated through the fibronectin. Biochemical Journal. 395 (3): 579-86.

Morita, A., Ariizumi, K., Ritter, R., Jester, J. V., Kumamoto, T., Johnston, S. A. and Takashima, A. (2001). Development of a Langerhans cell-targeted gene therapy format using a dendritic cell-specific promoter. Gene Therapy. **8** (22): 1729-1737.

Poget, S. F., Legge, G. B., Proctor, M. R., Butler, P. J. G., Bycroft, M. and Williams, R. L. (1999). The structure of a tunicate C-type lectin from polyandrocarpa misakiensis complexed with -galactose. Journal of Molecular Biology. **290** (4): 867-879.

Poget, S. F., Legge, G. B., Proctor, M. R., Butler, P. J. G., Bycroft, M. and Williams, R. L. (1999). The structure of a tunicate C-type lectin from Polyandrocarpa misakiensis complexed with D-galactose. Journal of Molecular Biology. **290** (4): 867-879.

Rappocciolo, G., Jenkins, F. J., Hensler, H. R., Piazza, P., Jais, M., Borowski, L., Watkins, S. C. and Rinaldo, C. R. (2006). DC-SIGN is a receptor for human herpesvirus 8 on dendritic cells and macrophages. Journal of Immunology. **176** (3): 1741-1749.

Reid, D. M., Montoya, M., Taylor, P. R., Borrow, P., Gordon, S., Brown, G. D. and Wong, S. Y. C. (2004). Expression of the {beta}-glucan receptor, Dectin-1, on murine leukocytes in situ correlates with its function in pathogen recognition and reveals potential roles in leukocyte interactions. Journal of Leukocyte Biology. **76** (1): 86-94.

Renedo, M. n., Arce, I., Rodrˆ›guez, A., Carretero, M., Lanier, L. L., Lopez-Botet, M. and Fernˆ¡ndez-Ruiz, E. (1997). The human natural killer gene complex is located on chromosome 12p12-p13. Immunogenetics. **46** (4): 307-311.

Richard, M., Thibault, N., Veilleux, P., Breton, R. and Beaulieu, A. D. (2003). The ITIM-bearing CLECSF6 (DCIR) is down-modulated in neutrophils by neutrophil activating agents. Biochemical and Biophysical Research Communications. **310** (3): 767-773.

Richard, M., Veilleux, P., Rouleau, M., Paquin, R. and Beaulieu, A. D. (2002). The expression pattern of the ITIM-bearing lectin CLECSF6 in neutrophils suggests a key role in the control of inflammation. Journal of Leukocyte Biology. **71** (5): 871-880.

Riedl, E., Tada, Y. and Udey, M. C. (2003). Identification and characterization of an alternatively spliced isoform of mouse Langerin/CD207. Journal of Investigative Dermatology. **123** (1): 78-86.

Rogers, N. C., Slack, E. C., Edwards, A. D., Nolte, M. A., Schulz, O., Schweighoffer, E., Williams, D. L., Gordon, S., Tybulewicz, V. L., Brown, G. D. and Sousa, C. R. E. (2005). Syk-dependent cytokine induction by Dectin-1 reveals a novel pattern recognition pathway for C type lectins. Immunity. **22** (4): 507-517.

Rogers, S., Shaw, I., Ross, N., Nair, V., Rothwell, L., Kaufman, J. and Kaiser, P. (2003). Analysis of part of the chicken Rfp-Y region reveals two novel lectin

genes, the first complete genomic sequence of a class I a-chain gene, a truncated class II ^ -chain gene, and a large CR1 repeat. Immunogenetics. **55** (2): 100-108.

Rogers, S. L., Gobel, T. W., Viertlboeck, B. C., Milne, S., Beck, S. and Kaufman, J. (2005). Characterization of the chicken C-type lectin-like receptors B-NK and B-lec suggests that the NK complex and the MHC share a common ancestral region. Journal of Immunology. **174** (6): 3475-3483.

Sallusto, F., Cella, M., Danieli, C. and Lanzavecchia, A. (1995). Dendritic cells use macropinocytosis and the mannose receptor to concentrate macromolecules in the major histocompatibility complex class II compartment: downregulation by cytokines and bacterial products. Journal of Experimental Medicine. **182** (2): 389-400.

Sallusto, F. and Lanzavecchia, A. (1994). Efficient presentation of soluble antigen by cultured human dendritic cells is maintained by granulocyte/macrophage colony-stimulating factor plus interleukin 4 and downregulated by tumor necrosis factor alpha. Journal of Experimental Medicine. **179** (4): 1109-1118.

Small, M. and Kraal, G. (2003). In vitro evidence for participation of DEC-205 expressed by thymic cortical epithelial cells in clearance of apoptotic thymocytes. International Immunology. **15** (2): 197-203.

Sobanov, Y., Bernreiter, A., Derdak, S., Mechtcheriakova, D., Schweighofer, B., Duchler, M., Kalthoff, F. and Hofer, E. (2001). A novel cluster of lectin-like receptor genes expressed in monocytic, dendritic and endothelial cells maps close to the NK receptor genes in the human NK gene complex. European Journal of Immunology. **31** (12): 3493-3503.

Sturm-Ramirez, K. M., Ellis, T., Bousfield, B., Bissett, L., Dyrting, K., Rehg, J. E., Poon, L., Guan, Y., Peiris, M. and Webster, R. G. (2004). Reemerging H5N1

Influenza viruses in Hong Kong in 2002 are highly pathogenic to ducks. Journal of Virology. **78**(9): 4892-4901.

Taylor, M., Bezouska, K. and Drickamer, K. (1992). Contribution to ligand binding by multiple carbohydrate-recognition domains in the macrophage mannose receptor. Journal of Biological Chemistry. **267** (3): 1719-1726.

Taylor, P. R., Brown, G. D., Reid, D. M., Willment, J. A., Martinez-Pomares, L., Gordon, S. and Wong, S. Y. C. (2002). The beta-glucan receptor, dectin-1, is predominantly expressed on the, surface of cells of the monocyte/macrophage and neutrophil lineages. Journal of Immunology. **169** (7): 3876-3882.

Taylor, P. R., Reid, D. M., Heinsbroek, S. E. M., Brown, G. D., Gordon, S. and Wong, S. Y. C. (2005). Dectin-2 is predominantly myeloid restricted and exhibits unique activation-dependent expression on maturing inflammatory monocytes elicited in vivo. European Journal of Immunology. **35** (7): 2163-2174.

Thibault, N., Richard, M., Veilleux, P. and Beaulieu, A. D. (2003). Immunoreceptor tyrosine-based inhibition motif of CLECSF6 (DCIR) on resting neutrophils bind specifically protein tyrosine phosphatase SHP-2. Faseb Journal. **17** (5): A1036-A1036.

Thomas, E. K., Nakamura, M., Wienke, D., Isacke, C. M., Pozzi, A. and Liang, P. (2005). Endo180 binds to the C-terminal region of type I collagen. Journal of Biological Chemistry. **280** (24): 22596-22605.

Thoraval, P., Afanassieff, M., Bouret, D. l., Luneau, G., Esnault, E., Goto, R. M., Chauss, A.-M., Zoorob, R., Soubieux, D., Miller, M. M. and Dambrine, G. (2003). Role of nonclassical class I genes of the chicken major histocompatibility complex Rfp-Y locus in transplantation immunity. Immunogenetics. **5 5**(9): 647-651.

Trowsdale, J. (2001). Genetic and functional relationships between MHC and NK receptor genes. Immunity. **15** (3): 363-374.

Uccini, S., Sirianni, M., Vincenzi, L., Topino, S., Stoppacciaro, A., Lesnoni La Parola, I., Capuano, M., Masini, C., Cerimele, D., Cella, M., Lanzavecchia, A., Allavena, P., Mantovani, A., Baroni, C. and Ruco, L. (1997). Kaposi's sarcoma cells express the macrophage-associated antigen mannose receptor and develop in peripheral blood cultures of Kaposi's sarcoma patients. American Journal of Pathology. **150** (3): 929-938.

Valladeau, J., Clair-Moninot, V., Dezutter-Dambuyant, C., Pin, J.-J., Kissenpfennig, A., Mattei, M.-G., Ait-Yahia, S., Bates, E. E. M., Malissen, B., Koch, F., Fossiez, F., Romani, N., Lebecque, S. and Saeland, S. (2002). Identification of mouse Langerin/CD207 in Langerhans cells and some dendritic cells of lymphoid tissues. Journal of Immunology. **168** (2): 782-792.

Valladeau, J., Ravel, O., Dezutter-Dambuyant, C., Moore, K., Kleijmeer, M., Liu, Y., Duvert-Frances, V., Vincent, C., Schmitt, D. and Davoust, J. (2000). Langerin, a novel C-type lectin specific to Langerhans cells, is an endocytic receptor that induces the formation of Birbeck granules. Immunity. **12** (1): 71-81.

van Gisbergen, K., Sanchez-Hernandez, M., Geijtenbeek, T. B. H. and van Kooyk, Y. (2005). Neutrophils mediate immune modulation of dendritic cells through glycosylation-dependent interactions between Mac-1 and DC-SIGN. Journal of Experimental Medicine. **201** (8): 1281-1292.

Verdijk, P., Dijkman, R., Plasmeijer, E. I., Mulder, A. A., Zoutman, W. H., Mieke Mommaas, A. and Tensen, C. P. (2005). A Lack of Birbeck Granules in Langerhans Cells Is Associated with a Naturally Occurring Point Mutation in the Human Langerin Gene. Journal of Investigative Dermatology. **124** (4): 714-717.

Vigerust, D. J., Egan, B. S. and Shepherd, V. L. (2005). HIV-1 Nef mediates post-translational down-regulation and redistribution of the mannose receptor. Journal of Leukocyte Biology. **77** (4): 522-534.

Weis, W. I. and Drickamer, K. (1996). Structural basis of lectin-carbohydrate recognition. Annual Review of Biochemistry. **65**: 441-473.

Weis, W. I., Kahn, R. and al, e. (1991). Structure of the calcium-dependent lectin domain from a rat mannose-binding protein determined by MAD phasing. Science. **254**(5038): 1608.

Willment, J. A., Gordon, S. and Brown, G. D. (2001). Characterization of the human beta -glucan receptor and its alternatively spliced isoforms. Journal of Biological Chemistry. **276** (47): 43818-43823.

Willment, J. A., Lin, H. H., Reid, D. M., Taylor, P. R., Williams, D. L., Wong, S. Y. C., Gordon, S. and Brown, G. D. (2003). Dectin-1 expression and function are enhanced on alternatively activated and GM-CSF-treated macrophages and are negatively regulated by IL-10, dexamethasone, and lipopolysaccharide. Journal of Immunology. **171** (11): 6297-6297.

Winzler, C., Rovere, P., Rescigno, M., Granucci, F., Penna, G., Adorini, L., Zimmermann, V. S., Davoust, J. and Ricciardi-Castagnoli, P. (1997). Maturation stages of mouse dendritic cells in Growth Factor-dependent long-term cultures. Journal of Experimental Medicine. **185** (2): 317-328.

Witmer-Pack, M. D., Swiggard, W. J., Mirza, A., Inaba, K. and Steinman, R. M. (1995). Tissue distribution of the DEC-205 protein that is detected by the monoclonal antibody NLDC-145: II. Expression in situ in lymphoid and nonlymphoid tissues. Cellular Immunology. **163** (1): 157-162.

Yanagihara, S., Komura, E., Nagafune, J., Watarai, H. and Yamaguchi, Y. (1998). EBI1/CCR7 is a new member of dendritic cell Chemokine receptor that is pp-regulated upon maturation. Journal of Immunology. **161** (6): 3096-3102.

Yokoyama, W., Ryan, J., Hunter, J., Smith, H., Stark, M. and Seaman, W. (1991). cDNA cloning of mouse NKR-P1 and genetic linkage with LY-49. Identification of a natural killer cell gene complex on mouse chromosome 6. Journal of Immunology. **147** (9): 3229-3236.