Cutting Edge: DC-SIGN; a Related Gene, DC-SIGNR; and CD23 Form a Cluster on 19p13

Elizabeth J. Soilleux, Roland Barten and John Trowsdale

*J Immunol* 2000; 165:2937-2942; doi: 10.4049/jimmunol.165.6.2937

Why *The JI*?

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

*average

**References** This article cites 33 articles, 13 of which you can access for free at:

http://www.jimmunol.org/content/165/6/2937.full#ref-list-1

**Subscription** Information about subscribing to *The Journal of Immunology* is online at:

http://jimmunol.org/subscription

**Permissions** Submit copyright permission requests at:

http://www.aai.org/About/Publications/JI/copyright.html

**Email Alerts** Receive free email-alerts when new articles cite this article. Sign up at:

http://jimmunol.org/alerts
Cutting Edge: DC-SIGN; a Related Gene, DC-SIGNR; and CD23 Form a Cluster on 19p13\(^1,2\)

Elizabeth J. Soilleux, \(^3\) Roland Barten, and John Trowsdale

DC-SIGN is a C-type lectin, expressed on a dendritic cell subset. It is able to bind ICAM3 and HIV gp120 in a calcium-dependent manner. Here we report the genomic organization of DC-SIGN and map it to chromosome 19p13 adjacent to the C-type lectin CD23 (FcεRII). We also report a novel, closely linked gene, DC-SIGNR, which shows 73% identity to DC-SIGN at the nucleic acid level and a similar genomic organization. Proteins encoded by both genes have tracts of repeats of 23 aa, predicted to form a coiled coil neck region. They also possess motifs that are known to bind mannose in a calcium-dependent fashion. We show concomitant expression of the two genes in endometrium, placenta, and stimulated KG1 cells (phenotypically similar to monocyte-derived dendritic cells). The existence of a DC-SIGN-related gene calls for reinterpretation of the HIV data to consider possible DC-SIGN/DC-SIGNR hetero-oligomerization. The Journal of Immunology, 2000, 165: 2937–2942.

C-SIGN, originally described in 1992 as a C-type lectin able to bind the HIV surface protein, gp120 (1), has been shown to be important for efficient infection with HIV (2). The DC-SIGN molecule is used by HIV to attach to dendritic cells in the genital tract and rectum (3). EGTA and mannan can inhibit this binding (2). Geitjenbeek et al. (2) suggest that dendritic cells then carry HIV particles to lymph nodes, where the infection of T lymphocytes via receptors such as CD4 and CCR5 may occur. DC-SIGN also binds the highly glycosylated molecule, ICAM3, found on T lymphocytes, enhancing the interaction of dendritic cells with T lymphocytes (3). A partial cDNA of a second, closely related gene was identified in 1999 by Yokoyama-Kobayashi et al. (4). In this study we investigated the mapping, genomic organization, and expression of DC-SIGN and a closely related gene, DC-SIGNR.

Materials and Methods

PCR and cloning of full genes and cDNAs

End primers (DC-SIGN-F (CTAAAGCAGGAGTTCTGGAC), DC-SIGN-R (CTAAAGGTTGCAAAGGTGGAG), DC-SIGNR-F (AACATCTGGGA CAGCG), and DC-SIGNR-R (GCAGTTACAACATTTCGACTCT)) were designed from published cDNA sequences (1, 4). Genes were amplified from genomic DNA using the Promega Taq DNA polymerase system (Promega, Southampton, U.K.), dNTPs (2.0 mM), magnesium (2.0 mM), and cycle conditions of 92°C for 1 min, 60°C for 1 min, and 72°C for 2–6 min. PCR products were cloned with a TOPO XL cloning kit (Invitrogen, San Diego, CA). cDNAs representing the entire coding region were cloned from placental cDNA as described above. Clones were sequenced as described previously (5). Predicted protein sequences were analyzed using Fip (http://menu.hgmp.mrc.ac.uk/menu-bin/Fip/) and ExPasy (http://www.expasy.ch/).

Mapping of the genes

Mapping was conducted by PCR using a radiation hybrid panel (HGMP-MRC) (6). A 366-bp sequence-tagged site from exon 4 of DC-SIGN was amplified as described with primers Lizo188 (CCCGATCTACACAGACCTG) and Lizo91 (TCTCGGTAGATCTCCTGCAT). The map position was confirmed with an RhyMe program (http://menu.hgmp.mrc.ac.uk/ menu-bin/RhyMe/).

PI1 artificial chromosome (PAC)\(^4\) identification

Gridded human RPCI 1–5 PAC libraries (7) were hybridized for 20 h with 50 ng of \(^{3}P\)dCTP-labeled full-length DC-SIGN cDNA. Membranes were washed to a stringency of 1× SSC/0.1% SDS at 65°C and exposed to x-ray film at −70°C for 24 h (8–10). Positive clones were provided by the Human Genome Mapping Project Resource Center. DNA was extracted using standard techniques (9). PAC DNA (150 ng) was digested with Nael (New England Biolabs, Beverly, MA) and separated on a pulse-field gel, with ramped switch times from 1–13 s at 200 V for 16 h. PCR screening confirmed the presence of DC-SIGN, DC-SIGNR, or CD23 on the PACs. Results were confirmed by sequencing (5).

DNA blot analysis

PAC DNA (20 μg) was digested with the restriction enzymes EcoRI (New England Biolabs) and PstI (New England Biolabs), followed by electrophoresis on 0.8% agarose-TBE gel. DNA was transferred to a Hybond N\(^+\) nylon membrane. Membranes were prehybridized for 1 h in hybridization buffer (0.001 M EDTA, 0.5 M sodium phosphate (pH 7.2), and 7% SDS). Fifty nanograms of a 140-bp probe across exon 5 of DC-SIGN was labeled with \(^{32}P\)dCTP and hybridized for 18 h at 65°C (7). Membranes were washed to a stringency of 0.1% SSC with 0.1% SDS at 65°C and exposed to Biomax MR film (Eastman Kodak, Rochester, NY) for 12 h.

\(^1\) This work was supported by grants from the Wellcome Trust, the Medical Research Council, a Medical Research Council Clinical Training Fellowship (to E.J.S.), and the Sackler Foundation (to E.J.S.).

\(^2\) The following accession numbers have been deposited in GenBank: AF209479 (DC-SIGN gene), AF209480 (DC-SIGNR gene, exons 1 and 2), AF209481 (DC-SIGNR gene, exons 3–8), and AF245219 (DC-SIGNR cDNA).

\(^3\) Address correspondence and reprint requests to Dr. Elizabeth J. Soilleux, Department of Pathology, Cambridge University, Tennis Court Road, Cambridge, U.K. CB2 1QP. E-mail address: ejs17@mole.bio.cam.ac.uk.

\(^4\) Abbreviations used in this paper: PAC, P1 artificial chromosome; CRD, carbohydrate recognition domain.
Cell culture and cDNA

Human YT, U937, Ind, 293T, and HCMED1-E6, were cultured in RPMI 1640, 10% FCS, 100 U/mL penicillin, and 100 U/mL streptomycin. Human Jurkat, T47D, MRC5, and C33a were cultured in DMEM, 10% FCS, 100 U/mL penicillin/streptomycin, and 2 mM l-glutamine. HT29 cells were cultured in DMEM/F12 (1/1), 10% FCS, and 100 U/mL penicillin/streptomycin. KG1 cells were cultured in IMDM, 20% FCS, 100 mM l-glutamine, and 100 U/mL penicillin/streptomycin and stimulated with PMA (10 ng/ml) and ionomycin (100 ng/ml; Sigma, St. Louis, MO) (11). Lange-rhans-type dendritic cells (12, 13) were donated by Paul Lehner (Cambridge University, Cambridge, UK.). Total RNA was isolated using th RNA-easy minikit (Qiagen, Crawley, U.K.). Placental and endometrial cDNA were donated by Amanda Evans (Cambridge University). Mono-cyte-derived dendritic cell cDNA (14) was donated by Jason Caulfield (Kings College, London, U.K.).

RT-PCR

First-strand cDNA was prepared using polyT and the Superscript kit (Life Technologies, Paisley, U.K.). cDNA synthesis was controlled using GAPDH primers (forward, ACCACACTTATGTCACACTC; reverse, TC-CACCCACCTTGTTCGTGTA). PCR for DC-SIGN and DC-SIGNR was performed with: DC-SIGN: Lizzol125, TGGACTCTCCTCTCTCATC; Lizzol202, CTGCGTGCTGGCCCTGAT; and DC-SIGNR: Lizzol126, TGGACCTCCTCCTCTCATC; Lizzol201, CTGGACGCGGTGAT. PCR conditions were 92°C for 1 min, 61°C for 1 min, and 72°C for 2 min for 35 cycles, with magnesium at 1.5 mM. PCR was controlled using DC-SIGN and DC-SIGNR cDNA clones. PCR products were sequenced (5).

Results

Cloning and sequencing of cDNAs

Using the GenBank database sequence for DC-SIGN (1) and a closely related partial cDNA sequence, AB015629 (4), end primers were de-signed, and full-length cDNAs were cloned from human placenta. Comparison of our full-length AB015629 clone, now termed DC-SIGN, with the previously published sequence revealed two extra exons. Additionally a 330-bp 3' intron (Fig. 4A) previously described in AB015629 cDNA had been spliced out of our cDNA sequence.

Both DC-SIGN and DC-SIGNR DNA sequences encode type II integral membrane proteins. The two genes show 73% identity at the nucleic acid level and 77% identity at the amino acid level. Predicted protein sequences from the full-length cDNAs are compared in Fig. 1. The cytoplasmic tails of both genes contained a di-leucine motif, which is a recognized internalization sequence (15). Exon 2 of DC-SIGN, but not that of DC-SIGNR, encodes a YXXL motif, as a further potential internalization signal (16). The putative transmembrane domains of DC-SIGN and DC-SIGNR consist of approximately 18 and 22 aa, respectively. An N-linked glycosylation sequence is found immediately after the transmembrane domains of both molecules (17), followed by a neck, encoded by a single exon, containing seven repeats of the 23 aa sequence KAAVGELExEKSKxQEIYQELTxL. The carbohydrate recognition domains (CRDs) are encoded by three separate exons, as described for CD23 (18) and the asialoglycoprotein receptors (19). These CRDs contain all the residues previously shown to be required for calcium-dependent binding of mannose (Fig. 2) (20).

Mapping of DC-SIGN and DC-SIGNR

Mapping of DC-SIGN and DC-SIGNR was conducted by performing PCR across exon 4 of DC-SIGN on a radiation hybrid panel (HGMP) (6). Results were consistent with a localization on chromosome 19p13.3. The RPCI PAC libraries (7) were screened with a probe consisting of the full-length sequence of DC-SIGN. DC-SIGN and DC-SIGNR were found on overlapping PACs, the sizes of which were determined by pulse field gel electrophoresis (Fig. 3). PACs were screened by PCR using exon primers for DC-SIGN and DC-SIGNR as well as CD23, a C-type lectin known to map to 19p13 (21), revealing the gene order shown in Fig. 3. The smallest PAC containing DC-SIGN, DC-SIGNR, and CD23 has an insert size of 105 kb. These data concur with preliminary high throughput genomic sequence data from The Lawrence Livermore center and with the localization of the CD23 gene (18). Therefore, these three C-type lectin genes, which have analogous genomic structures (see below), form a tight cluster on human chromosome 19p13. We probed the PAC clones to search for additional related
genes using an exon 5 probe from DC-SIGN, which shows 93% identity to exon 5 of DC-SIGNR and 25% identity to CD23 at the nucleic acid level. Only two equimolar bands corresponding to DC-SIGN and DC-SIGNR were obtained on the PACs already identified (Fig. 3). Although we cannot rule out the presence of further closely related genes elsewhere in the genome, no additional EST sequences are currently identifiable, and high throughput genomic sequence data have revealed no evidence for this. It is therefore most likely that only two genes exist in the DC-SIGN family at 19p13.

**Gene structure of DC-SIGN and DC-SIGNR**

The complete genes for DC-SIGN and DC-SIGNR were cloned by PCR from genomic DNA derived from a B cell lymphoma line and sequenced. Coding regions were predicted by alignment to the sequences obtained from placental cDNA. Both genomic sequences showed 100% identity to the corresponding cDNAs over the coding regions. Exon/intron structures of the two genes are shown in Fig. 4. All intron-exon boundaries follow the ag/tg rule (Fig. 4). The genes have very similar structures, except that DC-SIGNR has a longer 3' untranslated region, spanning two exons, compared with one in DC-SIGN. DC-SIGNR also contains an insert of approximately 1400 bp between exons 2 and 3, making intron 2 approximately 2000 bp, compared with 626 bp in DC-SIGN.

**Expression pattern**

Tissue specificity of expression was investigated by RT-PCR using gene-specific primers for DC-SIGN and DC-SIGNR (Fig. 5). In agreement with previous results (3), expression of DC-SIGN was restricted to endometrium, placenta, and stimulated KG1 cells (phenotypically similar to myeloid dendritic cells) (11). Although DC-SIGNR showed a lower level of expression, it was consistently detected in placenta with a very low level of expression in endometrium and stimulated KG1 cells. Cultured dendritic cells with a Langerhans cell-type phenotype (22) were negative for both molecules. We demonstrated a low level of expression of both DC-SIGN and DC-SIGNR by RT-PCR from monocyte-derived dendritic cells (data not shown) and subsequently cloned full-length cDNAs corresponding to both transcripts.
Discussion

This study describes the comparative sequences, genomic organization, mapping to chromosome 19p13, and expression analysis of two related C-type lectin genes, DC-SIGN and a novel, closely related gene, DC-SIGNR. These two genes and CD23, the low affinity receptor for IgE, are all encoded within a 105-kb region of 19p13. The DC-SIGN, DC-SIGNR, and CD23 genes all possess CRDs encoded by three separate exons. The close linkage and similar gene structures suggest that these three genes may have arisen via duplication of an ancestral gene.

The cytoplasmic tails of both DC-SIGN and DC-SIGNR contain di-leucine motifs, which may mediate internalization, thus indicating that these molecules may act as carbohydrate receptors (23). However, work by Geijtenbeek et al. suggests that endocytosis via DCSIGN is extremely inefficient in dendritic cells (2). A YXXL motif in the cytoplasmic tail of DC-SIGN (Fig. 1) may provide a further internalization motif (16) or, alternatively, a site for tyrosine phosphorylation (24).

CD23 has a neck consisting of three leucine-rich repeats in man and two in rat and mouse. This forms an α-helical coiled coil structure, which mediates trimerization (25, 26). Each of the repeat units is 21 aa long and is encoded by a separate exon. The close linkage and similar gene structures suggest that these three genes may have arisen via duplication of an ancestral gene.

The cytoplasmic tails of both DC-SIGN and DC-SIGNR contain di-leucine motifs, which may mediate internalization, thus indicating that these molecules may act as carbohydrate receptors (Fig. 1) (23). However, work by Geijtenbeek et al. suggests that endocytosis via DCSIGN is extremely inefficient in dendritic cells (2). A YXXL motif in the cytoplasmic tail of DC-SIGN (Fig. 1) may provide a further internalization motif (16) or, alternatively, a site for tyrosine phosphorylation (24).

CD23 has a neck consisting of three leucine-rich repeats in man and two in rat and mouse. This forms an α-helical coiled coil structure, which mediates trimerization (25, 26). Each of the repeat units is 21 aa long and is encoded by a separate exon. This contrasts with DC-SIGN and DC-SIGNR, which each have 7 repeats of a 23-aa sequence, encoded by a single exon (Fig. 4). There is very high sequence identity between the repeat units, within each protein, and between DC-SIGN and DC-SIGNR (Fig. 1). By analogy to other lectin receptors, such as the asialoglycoprotein receptors and CD23 (26–29), we suggest that this domain could mediate oligomerization, forming an α-helical coiled coil. Oligomerization may serve to increase the avidity of ligand binding. The human

---

FIGURE 3. A, Alignment of PAC clones hybridizing with DC-SIGN. The RPCI 1–5 PAC libraries were screened with the full-length cDNA for DC-SIGN. Subsequent PCR confirmed that seven PACs contained DC-SIGNR, of which six contained DC-SIGN and three contained CD23 by PCR, consistent with the gene order shown. The smallest PACs containing all three genes are 105 kb. PAC sizes are as follows: 1098P12, 105 kb; 84B17, 105 kb; 1191N20, 110 kb; 362C12, 97 kb; 1103M8, 100 kb; 729E9, 108 kb; 765L11, 108 kb. B, Southern blot of PACs. A 300-bp probe was used to a conserved region using primers flanking exon 5 of the DC-SIGN gene. Digestion of all the DC-SIGNR-positive PACs was conducted with EcoRI and PstI, which did not cut within the probe sequence. There was no evidence for additional closely related genes on these PACs.

FIGURE 4. A, DC-SIGN and DC-SIGNR gene structures. These were assembled by comparison of the full genomic sequence of DC-SIGN and partial sequence of DC-SIGNR. The two genes have very similar gene structures, except for the presence of an additional 3′ untranslated exon of DC-SIGNR. This is present in cDNA. Sequence data across intron 2 of DC-SIGNR are incomplete. Sizes are shown in nucleotides. 5′ and 3′ untranslated regions are shown in italics. Exons are written in capital letters. Introns are written in lowercase letters. AG/GT splice donor and acceptor sites are shown in bold type.

B, Intron-exon boundaries of DC-SIGN and DC-SIGNR. All intron-exon boundaries follow the ag/gt rule. Initiating ATGs are underlined. 5′ and 3′ untranslated regions are shown in italics. Exons are written in capital letters. Intron-exon boundaries of DC-SIGN and DC-SIGNR are shown in bold type. DC-SIGN has a short 3′ untranslated region, encoded by only one exon and consisting of 59 bases, while the DC-SIGNRs 289-bp 3′ untranslated region is encoded by two exons.
asialglycoprotein receptors, H1 and H2, have similar coiled coil neck structures and have been shown to form noncovalently associated heterotetramers with a stoichiometry of 2:2 (28). Like DC-SIGN and DC-SIGNR, the H1 and H2 genes are linked, but are found on chromosome 17p (30).

The sequences of the CRDs of DC-SIGN and DC-SIGNR show great identity to the human asialglycoprotein receptors (41 and 34% at the amino acid level, respectively) and rat CD23 (both 33% at the amino acid level; Fig. 2). Consistent with previous work (1–3), DC-SIGN, shows features of a mannose binding lectin, as opposed to the features of a protein-binding NK cell lectin (Fig. 2) (31). DC-SIGNR shows 77% identity to DC-SIGN at the amino acid level and also possesses all the residues shown to be required for the binding of mannose (Fig. 2) (31). The high level of homology between DC-SIGN and DC-SIGNR and their concomitant expression in placenta, endometrium, and a subset of dendritic cells suggest that DC-SIGNR may function in a similar manner to DC-SIGN, binding HIV gp120, ICAM-3 and perhaps other mannosylated proteins. Although it has been shown that the binding of DC-SIGN to ICAM3 and gp120 can be inhibited by mannan (1, 3) and may therefore not involve direct protein-protein interactions, the presence of the residues required for mannose binding must be interpreted cautiously. These residues are largely conserved in human CD23 and are completely conserved in murine and rat CD23 (Fig. 2), although the primary ligand of these molecules is thought to be IgE (18). Data suggest that mannan can partially inhibit the binding of human CD23 to certain of its ligands (32, 33). Therefore, DC-SIGN and DC-SIGNR may also bind a protein ligand, and experiments are underway to investigate this possibility.

The existence of a DC-SIGN-related gene (DC-SIGNR) with a similar pattern of expression is intriguing. Besides the formation of homo-oligomers, it is possible that hetero-oligomers of the two polypeptides may form, as described for the human asialglycoprotein receptors (28). Given the heterogeneity of dendritic cell phenotypes (34), further investigation of the differential expression patterns of the two genes across these phenotypes and at different anatomical sites is warranted. A further question is posed by the high level of expression in placenta, which may not be accounted for entirely by the presence of cells with a dendritic cell phenotype. The expression of DC-SIGN or DC-SIGNR on placenta may provide the key to explaining the mechanism of vertical transmission of HIV and may therefore give valuable insights into its prevention.

Acknowledgments
We acknowledge Mary Carrington and Kurt Drickamer for helpful discussion. We thank Amanda Evans, Jason Caulfield, Paul Macary, and Paul Lehner for provision of DNA, and Rachel Allen for careful reading of the manuscript.

References


