FOR THE RECORD

The L6 membrane proteins—A new four-transmembrane superfamily

MARK D. WRIGHT,¹ JIAN NI,² AND GEORGE B. RUDY¹

¹The Walter and Eliza Hall Institute of Medical Research, Post Office Royal Melbourne Hospital, 3050 Victoria, Australia

²Human Genome Sciences, 9410 Key West Avenue, Rockville, Maryland 20850

(RECEIVED March 20, 2000; FINAL REVISION May 25, 2000; ACCEPTED May 25, 2000)

Abstract: L6, IL-TMP, and TM4SF5 are cell surface proteins predicted to have four transmembrane domains. Previous sequence analysis led to their assignment as members of the tetraspanin superfamily. In this paper, we identify a new sequence (L6D) that is strikingly similar to L6, IL-TMP, and TM4SF5. Analyses of these four sequences indicate that they are not significantly related to genuine tetraspanins, but instead constitute their own L6 superfamily.

Keywords: L6; superfamily; tetraspanin

L6 is a highly expressed surface protein of human lung, breast, colon, and ovarian carcinomas that has excited considerable interest among tumor immunologists as a possible target for monoclonal antibody based cancer therapy (Hellstrom et al., 1986a, 1986b; Garkavij et al., 1995; Richman et al., 1995). Molecular cloning of the L6 antigen showed it to be a surface protein with four transmembrane domains, and initial analyses of the L6 sequence (Marken et al., 1992) suggested that the protein was a member of the tetraspanin (or transmembrane 4) superfamily. The tetraspanin superfamily comprises cell surface proteins characterized by four highly conserved transmembrane domains (Wright & Tomlinson, 1994; Maecker et al., 1997). The function of these tetraspanin proteins is not currently well understood, although several have been implicated in various signal transduction events mediating cell proliferation and activation (Wright & Tomlinson, 1994; Maecker et al., 1997). They may also control cell motility, tumor cell metastasis and cellular adhesion, possibly via noncovalent molecular interactions with integrins (Hemler et al., 1996). However, subsequent sequence analyses have suggested that L6 is not a bona fide member of the tetraspanin superfamily (Wright & Tomlinson, 1994). This issue has been further confused by the molecular cloning of the intestinal and epithelial protein IL-TMP (Wice & Gordon, 1995) and the prostate cancer expressed protein TM4SF5 (Muller-Pillasch et al., 1998). These proteins, with se-

quences clearly related to L6 (Wice & Gordon, 1995; Muller-Pillasch et al., 1998), were also reported to be members of the tetraspanin superfamily. Here, we report the molecular cloning of a novel membrane protein L6D. Our analyses of its protein sequence reveal highly significant similarity to L6, IL-TMP, and TM4SF5, yet no significant homology to genuine members of the tetraspanin superfamily is found. Thus, we describe the novel L6 membrane protein superfamily, distinct from the previously defined tetraspanin superfamily.

A database containing more than one million ESTs obtained from greater than 650 different cDNA libraries has been generated by Human Genome Sciences, Inc. and The Institute for Genomic Research using high throughput automated DNA sequence analysis of randomly selected human cDNA clones (Adams et al., 1995). Sequence homology comparisons of each EST were performed against the GenBank database using the BLAST family of algorithms (Altschul et al., 1990). ESTs having homology to previously identified sequences (probability >0.01) were collected in a database. A specific homology search using the known amino acid sequence of human L6 against this database revealed several ESTs having greater than 30% homology. Several clones were selected for further investigation. The complete cDNA sequence of L6D was obtained from a human umbilical endothelial cell library.

An alignment of the deduced L6D protein sequence and the other three L6 superfamily members, L6, IL-TMP, and TM4SF5, is shown in Figure 1. For comparison, two typical tetraspanin proteins, CD53 (Amiot, 1990) and CD151 (Fitter et al., 1995), are also included. Superficially, L6, IL-TMP, TM4SF5, and L6D do display a very limited resemblance to the tetraspanin superfamily. They share a similar topology: four transmembrane domains, short cytoplasmic domains at their N- and C-termini, and two extracellular domains, a smaller domain between transmembrane domains 1 and 2 (TM1 and TM2), and a larger domain between TM3 and TM4. This proposed membrane topology has been largely confirmed by monoclonal antibody epitope mapping analyses of both genuine tetraspanin proteins (Tomlinson et al., 1993) and L6 (Edwards et al., 1995).

L6, IL-TMP, TM4SF5, and L6D show a striking degree of overall identity with one another (ranging from 38–50%). This identity is distributed over the entire sequence with the only region that

Reprint requests to: Dr. Mark Wright, The Walter and Eliza Hall Institute of Medical Research, Post Office Royal Melbourne Hospital, 3050 Victoria, Australia; e-mail: wright_m@wehi.edu.au.

Fig. 1. Alignment of the L6 family member sequences, L6 (Marken et al., 1992), IL-TMP (Wice & Gordon, 1995), TM4SF5 (Muller-Pillasch et al., 1998), and L6D; and two typical members of the tetraspanin superfamily, CD53 (Amiot, 1990) and CD151 (Fitter et al., 1995). The alignment shown is a modification of that obtained using PILEUP (Program Manual for the Wisconsin Package, Version 8, Genetics Computer Group, 575 Science Drive, Madison, Wisconsin). Identical residues present in at least three sequences are boxed. Putative transmembrane domains are shaded. Residues strongly conserved in the tetraspanin superfamily are highlighted in the CD53 and CD151 sequences.

shows a greater divergence being the C-terminal cytoplasmic domain. This is in stark contrast to the tetraspanin superfamily, where transmembrane domains are very highly conserved, but cytoplasmic domains and extracellular domains show a much greater degree of divergence (Wright & Tomlinson, 1994). It is therefore notable that the L6 transmembrane domains show only very lim-

ited similarity to those of the genuine tetraspanins (Fig. 1). The length (in amino acid residues) of the transmembrane domains, strongly conserved among genuine tetraspanins, is clearly very different in the L6-like molecules. Moreover, the degree of sequence conservation in the transmembrane domains, comparing the L6-like molecules with the genuine tetraspanins, is not impres-

sive. In particular, the [A/G][F/V]LGC motif (highlighted in Fig. 1), which is present in TM2 of all but the most divergent tetraspanin molecules, is completely absent from the L6 superfamily.

The L6 proteins lack a number of other features shared by almost all tetraspanin proteins and also have characteristics not shared with the tetraspanins. The major extracellular domain of the tetraspanins, although relatively divergent, has four highly conserved cysteine residues contained in the following motifs: CCG, PXSC, EGC (highlighted in Fig. 1). These four cysteines, which participate in disulphide bond formation critical to the correct folding of this domain (Oligino et al., 1988; Levy et al., 1991; Tomlinson et al., 1993), and the CCG motif in particular are absolutely conserved in sequences of bona fide tetraspanin proteins. However, the equivalent domain of the L6 superfamily proteins does not contain these four conserved cysteine residues; rather, there are two cysteine residues, neither of which occurs in the motifs found in the tetraspanins. On this point, there are CCG motifs which do occur in the L6, IL-TMP, and TM4SF5 sequences (but not the L6D sequence) and which have been previously highlighted as sequence which does show homology to the tetraspanin superfamily (Wice & Gordon, 1995). However, it must be pointed out that these sequences do not occur in the extracellular domain. Rather, they are in the cytoplasmic domain between TM2 and TM3; any resemblance here with the tetraspanin superfamily cannot have functional significance and must be coincidental. Other features of the tetraspanin superfamily are the highly conserved charged or polar amino acid residues in TM1 (an asparagine in most family members) and even more notably, the glutamic acid/glutamine residues in TM3 and TM4 (highlighted in Fig. 1). It has been argued that these charged/polar residues may be of great importance to tetraspanin function (Wright & Tomlinson, 1994). They may gate a possible ion channel, or alternatively mediate molecular interactions with other membrane proteins. However, these charged/polar residues are clearly not conserved in the L6 superfamily proteins (Fig. 1). There is no charged or highly polar residue at all in the interior of TM3; however, the L6 superfamily proteins do have two conserved polar residues in TM4, a glutamic acid/glutamine residue (which does not appear to align with the glutamic acid/glutamine residue found in tetraspanins) followed by a glutamine/arginine seven residues later. A final feature of L6 superfamily molecules not found in the tetraspanins is a highly conserved cytoplasmic domain of 23 amino acid residues between TM2 and TM3. In the tetraspanin superfamily, there are only five amino acids separating TM2 and TM3.

It is still possible from the sequence analyses described above that the L6-like molecules may be distantly diverged tetraspanins that have lost many of the features of tetraspanin molecules through evolution. To detect such distant evolutionary relationships, we compared the sequences of L6, IL-TMP, TM4SF5, and L6D to a panel of tetraspanin superfamily proteins using the ALIGN program (Dayhoff et al., 1983) and the 250 PAM matrix (Table 1). This analysis scores pairwise alignments and has been used successfully to analyze the immunoglobulin superfamily (Williams & Barclay, 1988). The ALIGN score is the distance in standard deviations by which an optimal alignment score for two sequences exceeds the average score for pairwise alignments of random permutations of those sequences. Scores of 3.00 (or greater) have less than a 1 in 1,000 probability of occurring by chance and are treated as significant. The 14 tetraspanin sequences used included 10 "canonic" members of the family and four "unusual" family members, the ocular proteins Rom-1 (Bascom et al., 1993) and RDS/

peripherin (Farrar et al., 1991), the sarcoma-expressed molecule SAS (Jankowski et al., 1994), and the bladder epithelial protein uroplakin 1B (Yu et al., 1994). Included in the analyses as controls were 3 four-transmembrane proteins from superfamilies known to be unrelated to the tetraspanin superfamily: CD20 (Tedder et al., 1988), the gap junction protein connexin α -1 (Fishman et al., 1990), and the acetylcholine receptor α protein (Noda et al., 1983). It is clear from the ALIGN analyses (Table 1) that L6, IL-TMP, TM4SF5, and L6D are very closely related to one another with align scores all greater than 18. The degree of similarity within the "canonic" tetraspanin proteins is also apparent, as all scores within this subgroup were >7 . The ability of these analyses to detect distant evolutionary relationships is shown by the analyses of the unusual tetraspanin proteins, RDS and Rom-1, which are particularly similar to one another, and showed seven and five ALIGN scores of >3 , respectively, when aligned to other tetraspanins. In contrast, L6, IL-TMP, TM4SF5, and L6D showed only 1, 0, 1, and 1 ALIGN scores greater than three in alignments to known members of the tetraspanin superfamily. This degree of similarity is no greater than that observed for the unrelated connexin α -1 molecule, which showed a surprising degree of similarity to CD53 (ALIGN score 3.69) in the total absence of similarity to any other tetraspanin protein. These data demonstrate that there is no significant sequence homology between the L6-like molecules and the tetraspanins. Three tetraspanins do give higher align scores with the L6-like molecules than other tetraspanins: CD81, CD9, and Sm23. These values are below what is generally considered to be significant, and it should be noted that CD9 and CD81 are very highly related to one another (ALIGN score 33.07), and have a common ancestor in the hagfish (M.G. Tomlinson, pers. comm.). This very weak similarity between the L6-like molecules and CD81, CD9, and Sm23 occurs in the absence of significant similarity to other tetraspanin proteins and must therefore be among amino acid residues that are not highly conserved throughout the tetraspanin superfamily. This doubtful similarity between a few tetraspanins and the L6-like proteins is of dubious evolutionary significance, and most probably occurs merely by chance.

That the L6-like proteins are distinct from the tetraspanins and constitute their own separate protein superfamily is further supported by an analysis of the gene structure of L6. Proteins that are members of the same protein superfamily are thought to have evolved divergently from a primordial ancestral gene. This common ancestry is reflected not only in observed sequence homology, but also in conserved gene structure. A comparison of the gene structure of L6 and the tetraspanin molecule CD63 is shown diagrammatically in Figure 2, where positions of the introns are mapped onto a diagram showing the likely membrane topology of the expressed proteins. CD63 (Hotta et al., 1988) has a gene structure typical of all the tetraspanins (Wright et al., 1993; Tomlinson & Wright, 1995), as it has six introns whose positions are very highly conserved amongst members of the superfamily. To the best of our knowledge, the only L6 molecule whose gene structure has been determined is hamster L6 (Kurihara et al., 1997). It is clear from Figure 2 that the hamster L6 gene structure bears no resemblance to the conserved gene structures of the tetraspanins. There are only four introns, and the positions of these introns with respect to the coding sequence are entirely different from those of the introns in the tetraspanin genes.

Further confirmation of our conclusion that L6, IL-TMP, TM4SF5, and L6D constitute a separate superfamily distinct from the tetraspanins is provided by phylogenetic analysis using both

[illegible]

^aCD53 (Amiot, 1990), CD37 (Classon et al., 1990), CD81 (Oren et al., 1990), CD63 (Hotta et al., 1988), CD151 (Fitter et al., 1995), CD9 (Boucheix et al., 1991), Sm23 (Wright et al., 1990), CD82 (Gaigitsch et al., 1991), CO/029 (Szala et al., 1990), RDS/peripherin (Travis et al., 1991), SAS (Jankowski et al., 1994), Uroplakin 1B (Yu et al., 1994), Rom-1 (Bascom et al., 1993), and A15 (Emi et al., 1993); four L6 superfamily sequences: L6 (Marken et al., 1992), IL-TMP (Wice & Gordon, 1995), TM4SF5 (Muller-Pillasch et al., 1998), and L6D; and sequences from superfamilies with four transmembrane domains that are not related to the tetraspanins: CD20 (Tedder et al., 1988), connexin α (Zhang & Nicholson, 1989) and acetylcholine receptor α chain (Noda et al., 1983). The four "unusual" tetraspanins are italicized and underlined.

Fig. 2. Gene structures of L6 and of the tetraspanin CD63. A diagram of the predicted protein topology of L6 and CD63 is shown and the position of amino acid residues relative to the lipid bilayer is indicated. The location of introns is indicated by bolded bars.

Fig. 3. Unrooted, distance metric-based (neighbor-joining) phylogenetic tree depicting the relationship of "canonic" and "unusual" tetraspanins (italicized and underlined), the L6 family members (in rectangles), and the unrelated four transmembrane sequences (circled). The analyses were performed using programs available through the Australian National Genomic Information Service (ANGIS) WWW-server (<http://www.angis.org.au>). Alignment was performed using a slow pairwise (full dynamic programming) algorithm, BLOSUM series weight matrices (Henikoff & Henikoff, 1992), and gap opening and extension penalties of 10.00 and 0.10, respectively. Distance measures (maximum likelihood estimates based on the Dayhoff PAM matrix) were computed using EPROTDIST (modified version of PHYLIP version 3.572c PROTDIST). These were in turn analyzed with ENEIGHBOR (modified version of PHYLIP version 3.572c NEIGHBOR) using the neighbor-joining method (Saitou & Nei, 1987). The resultant phylogenetic tree was displayed with TREEVIEW (Page, 1996). Topology of the tree was confirmed by maximum parsimony analysis using EPROTPARS (modified version of PHYLIP version 3.572c PROTPARS). Bootstrap analysis (100 samples) of both methods provided good support for the represented topology, with clustering of all four L6 family members separately from the remaining sequences in 67% (distance metric) and 100% (maximum parsimony) of cases, using ESEQBOOT and ECONSENSE (modified versions of PHYLIP version 3.572c SEQBOOT and CONSENSE, respectively).

distance metric-based and maximum parsimony methods (Fig. 3). Ten "canonic" tetraspanins, four "unusual" tetraspanins, the four putative L6 family members, and the three known-unrelated "control" four-transmembrane sequences were multiply aligned using

ECLUSTALW—part of the original ClustalW distribution (Higgins et al., 1992; Thompson et al., 1994) modified for inclusion in EGCG (Program Manual for the Wisconsin Package, Version 8, Genetics Computer Group, 575 Science Drive, Madison, Wisconsin; Program Manual for the EGCG Package, Peter Rice, The Sanger Centre, Hinxton Hall, Cambridge, CB10 1RQ, England. Additional code by Peter Rice, The Sanger Centre, Hinxton, England and other members of the EGCG team). This alignment was then used to compute a phylogenetic tree (Fig. 3). It is clear from this tree that the four L6 proteins cluster separately and at considerable distance from the genuine tetraspanins. To compare the relative divergence of the L6 proteins with that of the other sequences, we measured the branch distances (brackets in Fig. 3) from a node central to the tetraspanin cluster (indicated by the spot, Fig. 3). The distance of any L6 sequence (IL-TMP is the most proximal) from any of the 10 “canonic” tetraspanins is very much greater than the distance between any of the 10 “canonic” tetraspanins and the 4 “unusual” tetraspanins. The known unrelated acetylcholine α receptor (4007) is actually closer to the center of the tetraspanin cluster than the most proximal L6 sequence (4511). Finally, the most distal L6 sequence (7054) is actually farther away from the central node than any of the known-unrelated connexin α -1, CD20 and acetylcholine α receptor sequences. It seems clear from this phylogenetic analysis that the L6-like sequences show no greater similarity to the tetraspanins than do these three known-unrelated four transmembrane sequences.

In conclusion, we have identified a novel molecule L6D, which shows a striking similarity to the tumor antigen L6, the lung and intestinal epithelial protein IL-TMP, and the pancreatic tumor molecule TM4SF5. These four highly related sequences do not share most of the key features of tetraspanin proteins (Fig. 1) nor do they display significant sequence homology with genuine members of the tetraspanin superfamily (Table 1). The gene structure of L6 shows no similarity to that of the tetraspanins (Fig. 2), and in phylogenetic analyses the L6-like sequences cluster separately and at considerable distance from the genuine tetraspanins (Fig. 3). Future studies of L6, IL-TMP, TM4SF5, and L6D should consider them not as members of the tetraspanin superfamily, but rather as constituting their own L6 superfamily.

Acknowledgments: This work was supported by the Australian National Health and Medical Research Council and the Victorian Anti Cancer Council. We thank Prof. Ken Shortman and Kirsten Puls for the critical reading of this manuscript. The excellent suggestions of the anonymous referee are also gratefully acknowledged.

References

Adams MD, Kerlavage AR, Fleischmann RD, Fuldner RA, Bult CJ, Lee NH, Kirkness EF, Weinstock KG, Gocayne JD, White O, et al. 1995. Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence. *Nature* 377 (Suppl.):3–17.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215:403–410.

Amiot M. 1990. Identification and analysis of cDNA clones encoding CD53. A pan-leukocyte antigen related to membrane transport proteins. *J Immunol* 145:4322–4325.

Bascom RA, Manara S, Collins L, Molday RS, Kalnins VI, McInnes RR. 1993. Cloning of the cDNA for a novel photoreceptor membrane protein (rom-1) identifies a disk rim protein family implicated in human retinopathies. *Neuron* 8:1172–1184.

Boucheix C, Benoit P, Frachet P, Billard M, Worthington RE, Gagnon J, Uzan G. 1991. Molecular cloning of the CD9 antigen. A new family of cell surface proteins. *J Biol Chem* 266:117–122.

Classon BJ, Williams AF, Willis AC, Seed B, Stamenkovic I. 1990. The primary structure of the human leukocyte antigen CD37, a species homologue of the Rat MRC OX-44 antigen. Correction. *J Exp Med* 172:1007.

Dayhoff MO, Barker WC, Hunt LT. 1983. Establishing homologies in protein sequences. *Methods Enzymol* 91:524–545.

Edwards CP, Farr AG, Marken JS, Nelson A, Bajorath J, Hellstrom KE, Hellstrom I, Aruffo A. 1995. Cloning of the murine counterpart of the tumor-associated antigen H-L6: Epitope mapping of the human and murine antigens. *Biochem* 34:12653–12660.

Emi N, Kitaori K, Seto M, Ueda R, Saito H, Takahashi T. 1993. Isolation of a novel cDNA clone showing marked similarity to ME491/CD63 superfamily. *Immunogen* 37:193–198.

Farrar JG, Kenna P, Jordan SA, Kumar-Singh R, Humphries MM, Sharp EM, Sheils DM, Humphries P. 1991. A three-base-pair deletion in the peripherin-RDS gene in one form of retinitis pigmentosa. *Nature* 354:478–480.

Fishman GI, Spray DC, Leinward LA. 1990. Molecular characterization and functional expression of the human cardiac gap junction channel. *J Cell Biol* 111:589–598.

Fitter S, Tetat TJ, Berndt MC, Ashman LK. 1995. Molecular cloning of cDNA encoding a novel platelet-endothelial cell tetra-span antigen PETA-3. *Blood* 86:1348–1355.

Garkavij M, Tennvall J, Strand SE, Norrgren K, Lindgren L, Nilsson R, Sjogren HO. 1995. Enhanced radioimmunotargeting of 125I-labeled L6-biotin monoclonal antibody by combining preloading of cold L6 mAb and subsequent immunoadsorption in rats. *Cancer Res* 55:5874s–5880s.

Gaugitsch HW, Hofer E, Huber NE, Schnabl E, Bamrucker T. 1991. A new superfamily of lymphoid and melanoma cell proteins with extensive homology to *Schistosoma mansoni* antigen Sm23. *Eur J Immunol* 21:377–383.

Hellstrom I, Beaumier PL, Hellstrom KE. 1986a. Antitumor effects of L6, an IgG2a antibody that reacts with most human carcinomas. *Proc Natl Acad Sci USA* 83:7059–7063.

Hellstrom I, Horn D, Linsley P, Brown JP, Brankovan V, Hellstrom KE. 1986b. Monoclonal mouse antibodies raised against human lung carcinoma. *Cancer Res* 46:3917–3923.

Hemler M, Mannion B, Beritchevski F. 1996. Association of TM4SF proteins with integrins: Relevance to cancer. *Biophys Biochem Acta* 1287:67–71.

Henikoff S, Henikoff JG. 1992. Amino acid substitution matrices from protein blocks. *Proc Natl Acad Sci USA* 89:10915–10919.

Higgins DG, Bleasby AJ, Fuchs R. 1992. Clustal V: Improved software for multiple sequence alignment. *Comput Appl Biosci* 8:189–191.

Hotta J, Ross AH, Huebner K, Isobe M, Wendeborn S, Chao MV, Ricciardi RP, Tsujimoto Y, Croce CM, Koprowski H. 1988. Molecular cloning and characterization of an antigen associated with early stages of melanoma tumor progression. *Cancer Res* 48:2955–2962.

Hotta H, Miyamoto H, Hara I, Takahashi N, Homma M. 1992. Genomic structure of the ME491/CD63 antigen gene and functional analysis of the 5′-flanking regulatory sequences. *Biochem Biophys Res Comm* 185:436–442.

Jankowski SA, Mitchell DS, Smith SH, Trent JM, Meltzer PS. 1994. SAS, a gene amplified in human sarcomas, encodes a new member of the transmembrane 4 superfamily of proteins. *Oncogene* 9:1205–1211.

Levy S, Nguyen VQ, Andria ML, Takahashi S. 1991. Structure and membrane topology of TAPA-1. *J Biol Chem* 266:14597–14602.

Maecker HT, Todd SC, Levy S. 1997. The tetraspanin superfamily: Molecular facilitators. *FASEB J* 11:428–442.

Marken JS, Schieven GL, Hellstrom I, Hellstrom KE, Aruffo A. 1992. Cloning and expression of the tumor-associated antigen L6. *Proc Natl Acad Sci USA* 89:3503–3507.

Muller-Pillasch F, Wallrap C, Lacher U, Freiss H, Buchler M, Adler G, Gress TM. 1998. Identification of a new tumor-associated antigen TM4SF5 and its expression in human cancer. *Gene* 208:25–30.

Noda M, Furutani Y, Takahashi H, Toyosato M, Tanabe T, Shimizu S, Kikuyotani S, Hirose T, Inyama S, Numa S. 1983. Cloning and sequence analysis of calf cDNA and human genomic DNA encoding alpha-subunit precursor of muscle acetylcholine receptor. *Nature* 305:818–823.

Oligino LD, Percy AJ, Harn DA. 1988. Purification and immunochemical characterization of a 22 kilodalton surface antigen from *Schistosoma mansoni*. *Mol Biochem Parasitol* 28:95–104.

Oren R, Takahashi S, Doss C, Levy R, Levy S. 1990. TAPA-1, the target of an antiproliferative antibody, defines a new family of transmembrane proteins. *Mol Cell Biol* 10:4007–4015.

Page RD. 1996. TreeView: An application to display phylogenetic trees on personal computers. *Comput Appl Biosci* 12:357–358.

Richman CM, DeNardo SJ, O’Grady LF, DeNardo GL. 1995. Radioimmunotherapy for breast cancer using escalating fractionated doses of 131I-labeled chimeric L6 antibody with peripheral blood progenitor cell transfusions. *Cancer Res* 55:5916s–5920s.

Saitou N, Nei M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425.

- Szala S, Kasai Y, Stepkowski Z, Rodeck U, Koprowski H, Linnebach AJ. 1990. Molecular cloning of cDNA for the human tumor-associated antigen CO-029 and identification of related transmembrane antigens. *Proc Natl Acad Sci USA* 87:6833–6837.
- Tedder TF, Streuli M, Schlossman SF, Saito H. 1988. Isolation and structure of a cDNA encoding the B1 (CD20) cell-surface antigen of human B lymphocytes. *Proc Natl Acad Sci USA* 85:208–212.
- Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680.
- Tomlinson MG, Williams AF, Wright MD. 1993. Epitope mapping of anti-rat CD53 monoclonal antibodies. Implications for the membrane orientation of the transmembrane 4 superfamily. *Eur J Immunol* 23:136–140.
- Travis GH, Brennan MB, Danielson PE, Kozak CA, Sutcliffe JG. 1991. Identification of a photoreceptor-specific mRNA encoded by the gene responsible for retinal degeneration slow (rds). *Nature (London)* 338:70–73.
- Wice BM, Gordon JI. 1995. A tetraspan membrane glycoprotein produced in the human intestinal epithelium and liver can regulate cell density-dependent proliferation. *J Biol Chem* 270:21907–21918.
- Williams AF, Barclay AN. 1988. The immunoglobulin superfamily—Domains for cell surface recognition. *Ann Rev Immunol* 6:381–405.
- Wright MD, Henkle KJ, Mitchell GF. 1990. An immunogenic Mr 23,000 integral membrane protein of *Schistosoma mansoni* worms that closely resembles a human tumor-associated antigen. *J Immunol* 144:3195–3200.
- Wright MD, Tomlinson MG. 1994. The ins and outs of the transmembrane 4 superfamily. *Immunol Today* 15:588–594.
- Yu J, Lin JH, Wu XR, Sun TT. 1994. Uroplakins Ia and Ib, two major differentiation products of bladder epithelium, belong to a family of four transmembrane domain (4TM) proteins. *J Cell Biol* 125:171–182.
- Zhang J-T, Nicholson BJ. 1989. Sequence and tissue distribution of a second protein of hepatic gap junctions, Cx26, as deduced from its cDNA. *J Cell Biol* 109:3391–3401.