



Published in final edited form as:

Gene. 2006 July 5; 376(1): 95–101. doi:10.1016/j.gene.2006.02.016.

The alleles of PECAM-1

Melanie S. Novinska^a, Bradley C. Pietz^b, Thomas M. Ellis^c, Debra K. Newman^{a,d}, and Peter J. Newman^{a,e,f,g,*}

^a Blood Research Institute, BloodCenter of Wisconsin, Milwaukee, WI 53201 USA

^b Product Development Laboratory, BloodCenter of Wisconsin, Milwaukee, WI 53201 USA

^c Laboratory of Histocompatibility and Immunogenetics, BloodCenter of Wisconsin, Milwaukee, WI 53201 USA

^d Department of Microbiology, Medical College of Wisconsin, Milwaukee, WI 53226 USA

^e Department of Pharmacology, Medical College of Wisconsin, Milwaukee, WI 53226 USA

^f Department of Cellular Biology, Medical College of Wisconsin, Milwaukee, WI 53226 USA

^g Cardiovascular Research Center, Medical College of Wisconsin, Milwaukee, WI 53226 USA

Abstract

Previous studies have reported the existence of eleven different single nucleotide polymorphisms (SNPs) within human PECAM-1 mRNA, several of which have recently been associated with disease. Though SNPs in the *PECAM-1* gene have been known for some time, the genetic background on which they exist, and their association into distinct allelic isoforms has not yet been established. To identify the major allelic isoforms of PECAM-1, we determined the nucleotide sequence of individual full-length cloned cDNAs derived from anonymous, unrelated volunteer individuals. Initial sequence analysis of 34 alleles from 17 individuals confirmed the presence of two distinct human PECAM-1 alleles (L₉₈S₅₃₆R₆₄₃ and V₉₈N₅₃₆G₆₄₃) within the human population. Each of these were found, upon more detailed analysis, to be superimposed on a previously unreported a2479g nucleotide polymorphism within the 3' untranslated region (3'UTR) that occurred on both allelic isoforms - yielding a total of four major alleles. Multiplex Luminex bead analysis of an additional 259 individuals allowed identification of 117 individuals homozygous for either the L₉₈S₅₃₆ or V₉₈N₅₃₆ allele, and sequence analysis around the R643G and a2479g polymorphic sites permitted accurate determination of significant differences in the gene frequencies of LSRa, LSRg, VNGa, and VNGg among Caucasian individuals. Identification of these PECAM-1 allelic isoforms should facilitate future detailed examination of PECAM-1-related disease associations, and may help resolve previously disparate results.

Keywords

single nucleotide polymorphism; disease association; CD31; PECAM-1

*Address correspondence to: Dr. Peter J. Newman, Blood Research Institute, BloodCenter of Wisconsin, P.O. Box 2178, 638 N. 18th Street, Milwaukee, WI 53201, Phone: (414) 937-6237, Fax: (414) 937-6284, peter.newman@bcw.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

PECAM-1 (CD31) is a vascular cell adhesion and signaling receptor that is expressed on the surfaces of platelets, leukocytes, and endothelial cells (Newman et al 1990), and is encoded by a ~70 kb gene near the end of the long arm of chromosome 17 (17q23) (Gumina et al 1996). PECAM-1 exists in mature form as a 130 kDa Type I transmembrane glycoprotein comprised of a 574 amino acid extracellular domain containing six Ig-like homology domains, a 19 amino acid transmembrane domain, and a cytoplasmic tail of varying length due to alternative splicing (Kirschbaum et al 1994). Ig-domain 1 mediates homophilic binding (Liao et al 1997; Newton et al 1997; Sun et al 1996), while Ig-domain 6 binds calcium (Jackson et al 1997a) and has been suggested to participate in cis interactions with integrin $\alpha_v\beta_3$ within the plane of the plasma membrane (Wong et al 2000). The cytoplasmic tail of PECAM-1 possesses two Immunoreceptor Tyrosine Inhibitory Motifs (ITIMs) (Newman 1999) which, upon tyrosine phosphorylation, recruit and activate the protein tyrosine phosphatase, SHP-2 (Jackson et al 1997b; Masuda et al 1997; Sagawa et al 1997). PECAM-1 has been demonstrated to participate in a variety of physiological events, including leukocyte adhesion and migration, angiogenesis, apoptosis, and modulation of Immunoreceptor Tyrosine Activating Motif (ITAM)-mediated cellular activation (for recent reviews, see (Ilan and Madri 2003; Newman 1997; Newman and Newman 2003)).

Like most genes, variations within the nucleotide sequence of the *PECAM-1* gene have been reported, with individual polymorphic residues identified within the 5'UTR, the extracellular and cytoplasmic domains, and the 3'UTR (summarized in Table 1). While the effects of these polymorphisms on PECAM-1-mediated adhesion or signaling have not yet been determined, mismatches at PECAM-1 amino acid residues 98, 536, or 643 have often (Balduini et al 1999; Behar et al 1996; Cavanagh et al 2005; Grumet et al 2001; Maruya et al 1998), though not universally (Nichols et al 1996), been associated with an increased incidence of acute graft-versus-host disease (GVHD), and these and other PECAM-1 polymorphisms have been linked with early onset of atherosclerosis (Elrayess et al 2003), increased risk of cardiovascular disease (Elrayess et al 2004; Fang et al 2005; Listi et al 2004; Sasaoka et al 2001; Song et al 2003; Wei et al 2004), and susceptibility to malarial infection (Kikuchi et al 2001), though the latter is also controversial (Casals-Pascual et al 2001).

While the frequencies of *individual* PECAM-1 polymorphisms have been determined in a limited number of population studies, these polymorphisms have not, to date, been linked into distinct PECAM-1 alleles, hampering efforts to more definitively establish PECAM-1-related disease associations. The purpose of the present investigation, therefore, was to determine the major alleles bearing each of the most commonly reported SNPs within the PECAM-1 gene. This information should not only permit more precise bio-epidemiological associations to be made amongst different human populations, but also enable biochemical and cell biological studies to be performed to investigate whether functional differences between PECAM-1 allelic isoforms might be causally linked to the reported disease associations of PECAM-1 SNPs.

2. Materials and Methods

2.1 Preparation of genomic and cDNA from human whole blood

Human whole blood was obtained from anonymous volunteer blood donors. RNA was isolated using a QIAamp[®] RNA Blood Mini Kit according to manufacturer's instructions (Qiagen, Valencia, CA). cDNA was then prepared from human RNA using the SuperScript[™] First-Strand Synthesis for RT-PCR kit (Invitrogen, Carlsbad CA). Following cDNA synthesis, RNase (1 μ l) (Invitrogen) was added, tubes incubated at 37°C for 20 minutes, and then put on ice or frozen at -20°C for later use. Genomic DNA was isolated from whole blood using a DNA Blood Mini Kit (Qiagen) following lysis of red blood cells with ammonium chloride.

2.2 PCR amplification, cloning, and sequencing

Primers used to generate nested PCR products were designed to amplify a 2520 bp region encompassing all known individual polymorphisms within the human PECAM-1 transcript (including the 5'- and 3'UTRs, and were manufactured by Integrated DNA Technologies, Inc. (Coralville, IA). Primary PCR reactions were carried out using *PfuTurbo*[®] DNA Polymerase (Stratagene, La Jolla, CA) using primers corresponding to human *PECAM-1* 18–35 (sense) 5'-*gccatggctgccattacc-3'* and 2573–2554 (antisense) 5'-*taagaaccggcagcttagcc-3'*. Amplification was performed for 35 cycles (95°C for 30 seconds, 59°C for 30 seconds with a ramping speed of 3°C/second and a 10°C gradient, 72°C for 3 minutes) in an Eppendorf Mastercycler Gradient Thermal Cycler (Brinkmann Instruments, Westbury, NY). Nested PCR reactions were performed for 25 cycles using identical amplification conditions using internally nested primers corresponding to PECAM-1 26–43 (sense) 5'-*tgcattacctgaccagc-3'* and 2545–2528 (antisense) 5'-*tgctgttctgtggag-3'*.

PCR products were separated from primers on a 1.5% agarose gel and then purified using the QIAquick PCR Purification kit (Qiagen) according to manufacturer's instructions. They were then ligated into the pPCR-Script Amp SK(+) cloning vector and transformed into XL10-Gold Kan ultracompetent cells (Stratagene). Plasmids were isolated using the QIAprep Miniprep kit (Qiagen), analyzed by selective restriction enzyme digestion, and their PECAM-1 cDNA inserts fully sequenced in a 96-well plate using an Applied Biosystems Model 3100 Capillary Sequencer using the following primers: 355–372 (sense) 5'-*aagaacctgacctgcag-3'*, 411–392 (antisense) 5'-*aggctgacgtgagaggtgg-3'*, 1177–1194 (sense) 5'-*tttccaagcccgaactg-3'*, 1588–1607 (sense) 5'-*gcggtattcaagacaacc-3'*, 2030–2047 (sense) 5'-*tcggagtgatcattgctc-3'*, and 2545–2528 (antisense) 5'-*tgctgtgtgtggag-3'*. Sequencing reactions were performed for 25 cycles (96°C for 30 seconds, 50°C for 15 seconds, and 60°C for 4 minutes). Sequence analysis of smaller defined regions surrounding the R643G and a2479g polymorphisms was performed by PCR amplification using 5'-*ctagaatttcctgtcactcacc-3'* (sense) and 5'-*gtactctcattgacacatcgct-3'* (antisense), and 5'-*gggcaatctcaatcttgag-3'* (sense) and 5'-*tgggagcagggcaggttcataat-3'* (antisense) primers, respectively, and sequenced using (sense) 5'-*ctagaatttcctgtcactcacc-3'*, (antisense) 5'-*gtactctcattgacacatcgct-3'*, (sense) 5'-*gggcaatctcaatcttgag-3'*.

2.3 Genotyping using Luminex Beads

Large-scale population genotyping for the L98V and S536N polymorphisms was performed using a MultiCode-PLx (EraGen Biosciences, Madison, WI) multiplex assay. Fluorescently labeled allele-specific elongation products were hybridized to Luminex[™] (Austin, TX) microspheres and detected on the Luminex 100 instrument using methods recently described in detail (Pietz et al 2005). Primers used for both locus-specific amplification and allele-specific extension reactions were synthesized by EraGen Biosciences (Madison, WI, see <http://www.era-gen.com/diagnostics/technology.html>). The following primers ('x' denotes the IsoBase residue, isoC) were used in multiplex allele-independent amplification of PECAM-1 around the L98V polymorphism (sense) 5'-*xatctatgactcagggac-3'*, (antisense) 5'-*gtgctcagttccaag-3'*, and around the S536N polymorphic site (sense) 5'-*ttctatcaaatgacctcaaat-3'*, (antisense) 5'-*xaggctgtgcagtaat-3'*. Allele-specific elongation reactions were performed with primers specific to the L98V polymorphism 5'-*caaggactcacttcaccaacag-3'* and 5'-*caaggactcacttcaccaacac-3'*, and primers specific to the S536N polymorphic site 5'-*tggaccaagcagaaggctag-3'* and 5'-*ttggaccaagcagaaggctag-3'*.

2.4 Statistical analysis

Hardy-Weinberg equilibrium analysis was performed, with all deviations from Hardy-Weinberg equilibrium analyzed using a Fisher's Exact Test contingency table. Statistically

significant differences in the gene frequencies of PECAM-1 allelic isoforms were determined by Chi-square analysis.

3. Results and Discussion

3.1 Identification of the primary alleles of human PECAM-1

Single nucleotide polymorphisms within the *PECAM-1* gene, including several that result in amino acid substitutions, were originally identified by comparing the sequences of PECAM-1 cDNAs cloned from different laboratories (Newman et al 1990; Simmons et al 1990; Stockinger et al 1990; Zehnder et al 1992), as well as the single sequence derived thus far from PECAM-1 genomic DNA (Kirschbaum et al 1994). The frequencies for several individual SNPs have since been determined in several small studies, and the results obtained from these are summarized in Table 1, which additionally shows the location of the SNPs relative to the exon in which they are encoded.

Because PECAM-1 SNPs have recently been implicated as risk factors for graft rejection, malarial infection, and cardiovascular disease (Table 2), we undertook identification of linked polymorphisms that might more precisely define PECAM-1 allelic isoforms. In preliminary studies, RNA was isolated from small blood samples derived from seventeen unrelated, anonymous volunteer blood donors, converted to cDNA, and nested PCR was performed to amplify full-length PECAM-1 transcripts (shown schematically in Figures 1A & 1B). The final 2500 bp amplified product (Figure 1C), encompassing all known PECAM-1 polymorphisms, including those within the 5' and 3'UTRs, was subcloned into a plasmid vector, and five clones from each donor were sequenced in their entirety to determine the linkage of SNPs carried on each of the two PECAM-1 alleles present in that individual. Sequence analysis of 34 alleles from 17 individuals revealed the presence of two distinct human PECAM-1 alleles (**L₉₈S₅₃₆R₆₄₃** and **V₉₈N₅₃₆G₆₄₃**), each of which was superimposed on an additional a2479g nucleotide polymorphism within the 3'UTR that occurred on both allelic isoforms – yielding a total of four major alleles, depicted in Figure 2. None of the other reported polymorphisms noted in Table 1 were found in any of the 34 alleles examined, indicating that they are likely to be rare variants of one or more of the four primary alleles. The genetic background on which they occur remains to be established.

To determine the gene frequencies of the four major PECAM-1 alleles - LSRa, LSRg, VNGa, and VNGg, genomic DNA from 259 healthy Caucasian blood donors was subjected to Luminex-based MultiCode-PLx bead hybridization analysis that detected fluorescently labeled allele-specific elongation products hybridized to Luminex™ microspheres in multiplex (Pietz et al 2005). Of these, 49 individuals were found to be L₉₈S₅₃₆ homozygotes and 68 people were V₉₈N₅₃₆ homozygotes. PCR amplification and sequence analysis of the regions surrounding the polymorphisms at R643G and a2479g was performed on these samples, and from them the gene frequencies for LSRa, LSRg, VNGa, and VNGg were calculated, and found to be 0.14, 0.28, 0.27, and 0.31, respectively. Chi-square analysis revealed significant differences in the frequencies of the LSR versus VNG allelic isoforms (p<0.025), and also between the LSRa and LSRg alleles (p<0.01). None of the other gene frequencies (e.g. between VNGa and VNGg) were found to be significantly different.

3.2 The N536S and G643R polymorphisms within exons 8 and 12 travel together as a haplotype block

Haplotype blocks are 10–100 kb regions of the human genome that contain SNPs sufficiently close to each other as to be nearly always inherited together - i.e. in strong linkage disequilibrium (Daly et al 2001; Gabriel et al 2002; Shifman et al 2003). Thus, though few or no recombination events occur within the “blocks” themselves, intervals between haplotype

blocks are commonly sites of recombination, leading to the loss of linkage disequilibrium between blocks. The observation that two major human PECAM-1 alleles, **L₉₈S₅₃₆R₆₄₃** and **V₉₈N₅₃₆G₆₄₃**, are each superimposed upon an additional a2479g nucleotide polymorphism within the 3'UTR, can be explained by the fact that exons 12–16 within the PECAM-1 gene are widely dispersed amongst large introns (more than 26,000 base pairs exist between the R643G polymorphism in exon 12 and the a2479g nucleotide polymorphism within exon 16), and that the middle third of the PECAM-1 gene (e.g. exons 8–12) travels together as a haplotype block. A crossing-over event somewhere between exons 12 and 16 (depicted schematically in Figure 3) was likely responsible for generating the four major alleles of PECAM that we observed in this study. Moreover, though the S536N (exon 8) and R643G (exon 12) polymorphisms have consistently been found in nearly complete linkage disequilibrium ((Elrayess et al 2004; Maruya et al 1998; Sasaoka et al 2001; Wei et al 2004), and this study), evidence exists that the L98V polymorphism within exon 3 can be inherited independently (Cavanagh et al 2005; Listi et al 2004; Maruya et al 1998; Sasaoka et al 2001), and may not be as tightly-linked to the centrally-located haplotype block within the PECAM-1 gene. Thus, a second crossing-over event (depicted with dotted lines in Figure 3) appears to be responsible for forming the less common VSR and LNG allelic isoforms. Future studies will be required to determine whether the frequencies of PECAM-1 alleles might vary within different ethnic groups.

3.3 Potential consequences of the a2479g polymorphism within the 3'UTR of PECAM-1 mRNA

Though no functional consequences of, or disease associations with, the a2479g nucleotide polymorphism have yet been reported, 3'UTRs can influence both mRNA stability (Rousseau et al 2003) and translation efficiency, and have also been associated with disease states. For example, a polymorphism in the 3'UTR of the interleukin-12B gene has been found to associate with late onset of type 1 diabetes mellitus (Windsor et al 2004), while another study demonstrated that a polymorphism within the 3'UTR of the cyclooxygenase-2 gene contributes to lung cancer susceptibility within the Chinese population (Hu et al 2005). Interestingly, the “prioritization” of synthesis of certain mRNAs involved in the synthesis of selenium-containing proteins have been shown to be influenced by sequences within 3'UTRs, and polymorphisms within this region may also regulate expression (Hesketh 2004). Based upon these findings, it may be of future interest to examine whether the PECAM-1 3'UTR polymorphism affects PECAM-1 mRNA and protein expression levels, and thereby PECAM-1-mediated adhesion and/or signaling.

3.4 Disease Associations with PECAM-1 Allelic Isoforms

Our identification of four major allelic isoforms for PECAM-1 within the Caucasian population may permit more detailed and accurate examination of PECAM-1-related disease associations (Table 2). Thus, although Sasaoka and colleagues (Sasaoka et al 2001) found a disproportionate increase in the frequencies of the **L₉₈**, **S₅₃₆**, and **R₆₄₃** forms of PECAM-1 in patients admitted for myocardial infarction (MI), especially males <60 yrs old, Listi et al. (Listi et al 2004) found an increased frequency (38.6% in patients compared with 24.6% in healthy individuals) of the **R₆₄₃** polymorphism in Sicilian MI patients in the *absence* of significant differences in the frequencies of **L₉₈** or **S₅₃₆**. Similarly, Song and colleagues (Song et al 2003) found a significant increase in the frequencies of **V₉₈** and **N₅₃₆** polymorphisms in patients presenting with early onset of coronary artery disease (CAD), while two other studies (Fang et al 2005; Wei et al 2004) found only **V₉₈** associated with CAD. Finally, PECAM-1 serves as an endothelial cell receptor for certain strains of *Plasmodium falciparum*-infected erythrocytes, and there are two reports regarding susceptibility of individuals carrying the **Leu₉₈** polymorphism to malarial infection: one showing an association (Kikuchi et al 2001) and one not showing such a connection (Casals-Pascual et al 2001). It is possible that re-genotyping these populations for PECAM-1 *alleles*, including the newly described 3'UTR polymorphism at nucleotide 2479

(this manuscript), rather than for individual SNPs, might help resolve the seemingly disparate results obtained in each of these studies.

Finally, although the consequences for allogeneic bone marrow transplantation of mismatching individual PECAM-1 polymorphisms has been controversial (Balduini et al 1999; Behar et al 1996; Grumet et al 2001; Maruya et al 1998; Nichols et al 1996), Cavanagh et al. (Cavanagh et al 2005), recently suggested, based on the results of a small study, that genotyping for *linked* PECAM-1 polymorphisms might be a better predictor of both acute graft-versus-host disease ($P=0.004$) and overall patient survival.

3.5 Conclusions

The major finding of the present study is that, rather than existing as a single invariant species, PECAM-1 is actually represented by distinct molecular isoforms that differ in functionally important regions within both the extracellular and cytoplasmic domains. In addition, a polymorphism within the 3' untranslated region further distinguishes these two protein isoforms, yielding four primary human PECAM-1 alleles: LSRa, LSRg, VNGa, and VNGg, with frequencies 0.14, 0.28, 0.27, 0.31, respectively, within the Caucasian population. The middle third of the PECAM-1 gene displays strong linkage disequilibrium, and is inherited as a haplotype block. A crossing over event within the 26 kb region lying *between* the polymorphism in exon 12 and the polymorphism within exon 16 has led to the generation of these four major PECAM-1 alleles, while another recombination event within the first third of the gene led to the generation of at least two additional alleles that are relatively rare, at least in Caucasians of European descent. Because PECAM-1 polymorphisms affect residues within domains of the protein that are involved in adhesion and signal transduction, it may be important in the future to determine whether differences exist in the biochemical and cell biological properties of these distinct isoforms. Finally, identification of the primary alleles of PECAM-1 should provide both investigators and clinicians with the ability to more accurately establish correlation of these alleles with diseases in which PECAM-1 polymorphisms have been implicated, including CAD, MI, and GVHD.

Acknowledgments

The authors thank Daniel B. Rowe, Ph.D., Division of Biostatistics, Medical College of Wisconsin, for his assistance in statistical analysis. This study was funded by grants HL-40126 and Training Grant HL-07209 from the Heart, Lung, and Blood Institute of the National Institutes of Health grant.

Abbreviations

PECAM-1	platelet endothelial cell adhesion molecule-1
SNP	single nucleotide polymorphism
Ig	immunoglobulin
ITIM	immunoreceptor tyrosine inhibitory motif
ITAM	immunoreceptor tyrosine activating motif
GVHD	graft-versus-host disease
PCR	polymerase chain reaction
MI	myocardial infarction
CAD	coronary artery disease

References

- Balduini CL, Noris P, Giorgiani G, Martinetti M, Klersy C, Spedini P, Belletti S, MacCario R, Gusberti L, Locatelli F. Incompatibility for CD31 and human platelet antigens and acute graft-versus-host disease after bone marrow transplantation. *Br J Haematol* 1999;106:723–729. [PubMed: 10468865]
- Behar E, Chao NJ, Hirake DD, Krishnaswamy S, Brown BW, Zehnder JL, Grumet FC. Polymorphism of adhesion molecule CD31 and its role in acute graft-versus-host disease. *N Engl J Med* 1996;334:286–291. [PubMed: 8532023]
- Casals-Pascual C, Allen S, Allen A, Kai O, Lowe B, Pain A, Roberts DJ. Short report: codon 125 polymorphism of CD31 and susceptibility to malaria. *Am J Trop Med Hyg* 2001;65:736–737. [PubMed: 11791967]
- Cavanagh G, Chapman CE, Carter V, Dickinson AM, Middleton PG. Donor CD31 genotype impacts on transplant complications after human leukocyte antigen-matched sibling allogeneic bone marrow transplantation. *Transplantation* 2005;79:602–605. [PubMed: 15753851]
- Daly MJ, Rioux JD, Schaffner SF, Hudson TJ, Lander ES. High-resolution haplotype structure in the human genome. *Nat Genet* 2001;29:229–232. [PubMed: 11586305]
- Elrass MA, Webb KE, Bellingan GJ, Whittall RA, Kabir J, Hawe E, Syvanne M, Taskinen MR, Frick MH, Nieminen MS, Kesaniemi YA, Pasternack A, Miller GJ, Humphries SE. R643G polymorphism in PECAM-1 influences transendothelial migration of monocytes and is associated with progression of CHD and CHD events. *Atherosclerosis* 2004;177:127–135. [PubMed: 15488875]
- Elrass MA, Webb KE, Flavell DM, Syvanne M, Taskinen MR, Frick MH, Nieminen MS, Kesaniemi YA, Pasternack A, Jukema JW, Kastelein JJ, Zwinderman AH, Humphries SE. A novel functional polymorphism in the PECAM-1 gene (53G>A) is associated with progression of atherosclerosis in the LOCAT and REGRESS studies. *Atherosclerosis* 2003;168:131–138. [PubMed: 12732396]
- Fang L, Wei H, Chowdhury SH, Gong N, Song J, Heng CK, Sethi S, Koh TH, Chatterjee S. Association of Leu125Val polymorphism of platelet endothelial cell adhesion molecule-1 (PECAM-1) gene & soluble level of PECAM-1 with coronary artery disease in Asian Indians. *Indian J Med Res* 2005;121:92–99. [PubMed: 15756041]
- Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D. The structure of haplotype blocks in the human genome. *Science* 2002;296:2225–2229. [PubMed: 12029063]
- Grumet FC, Hiraki DD, Brown BW, Zehnder JL, Zacks ES, Draksharapu A, Parnes J, Negrin RS. CD31 mismatching affects marrow transplantation outcome. *Biol Blood Marrow Transplant* 2001;7:503–512. [PubMed: 11669217]
- Gumina RJ, Kirschbaum N, Rao PN, vanTuinen P, Newman PJ. The human PECAM1 gene maps to 17q23. *Genomics* 1996;34:229–232. [PubMed: 8661055]
- Gumina RJ, Kirschbaum NE, Piotrowski K, Newman PJ. Characterization of the human Platelet/Endothelial Cell Adhesion Molecule-1 promoter: Identification of a GATA-2 binding element required for optimal transcriptional activity. *Blood* 1997;89:1260–1269. [PubMed: 9028949]
- Hesketh J. 3'-Untranslated regions are important in mRNA localization and translation: lessons from selenium and metallothionein. *Biochem Soc Trans* 2004;32:990–993. [PubMed: 15506944]
- Hu Z, Miao X, Ma H, Wang X, Tan W, Wei Q, Lin D, Shen H. A common polymorphism in the 3'UTR of cyclooxygenase 2/prostaglandin synthase 2 gene and risk of lung cancer in a Chinese population. *Lung Cancer* 2005;48:11–17. [PubMed: 15777967]
- Ilan N, Madri JA. PECAM-1: old friend, new partners. *Curr Opin Cell Biol* 2003;15:515–524. [PubMed: 14519385]
- Jackson DE, Loo RO, Holyst MT, Newman PJ. Identification and characterization of functional cation coordination sites in Platelet Endothelial Cell Adhesion Molecule-1. *Biochemistry* 1997a;36:9395–9404. [PubMed: 9235983]
- Jackson DE, Ward CM, Wang R, Newman PJ. The protein-tyrosine phosphatase SHP-2 binds PECAM-1 and forms a distinct signaling complex during platelet aggregation. Evidence for a mechanistic link between PECAM-1 and integrin-mediated cellular signaling. *J Biol Chem* 1997b;272:6986–6993. [PubMed: 9054388]

- Kikuchi M, Looareesuwan S, Ubalee R, Tasanor O, Suzuki F, Wattanagoon Y, Na-Bangchang K, Kimura A, Aikawa M, Hirayama K. Association of adhesion molecule PECAM-1/CD31 polymorphism with susceptibility to cerebral malaria in Thais. *Parasitol Int* 2001;50:235–239. [PubMed: 11719109]
- Kirschbaum NE, Gumina RJ, Newman PJ. Organization of the gene for human Platelet/Endothelial Cell Adhesion Molecule-1 (PECAM-1) reveals alternatively spliced isoforms and a functionally complex cytoplasmic domain. *Blood* 1994;84:4028–4037. [PubMed: 7994021]
- Liao F, Ali J, Greene T, Muller WA. Soluble domain 1 of platelet-endothelial cell adhesion molecule (PECAM) is sufficient to block transendothelial migration in vitro and in vivo. *J Exp Med* 1997;185:1349–1357. [PubMed: 9104821]
- Listi F, Candore G, Lio D, Cavallone L, Colonna-Romano G, Caruso M, Hoffmann E, Caruso C. Association between platelet endothelial cellular adhesion molecule 1 (PECAM-1/CD31) polymorphisms and acute myocardial infarction: a study in patients from Sicily. *Eur J Immunogenet* 2004;31:175–178. [PubMed: 15265022]
- Maruya E, Saji H, Seki S, Fujii Y, Kato K, Kai S, Hiraoka A, Kawa K, Hoshi Y, Ito K, Yokoyama S, Juji T. Evidence that CD31, CD49b, and CD62L are immunodominant minor histocompatibility antigens in HLA identical sibling bone marrow transplants. *Blood* 1998;92:2169–2176. [PubMed: 9731077]
- Masuda M, Osawa M, Shigematsu H, Harada N, Fujiwara K. Platelet endothelial cell adhesion molecule-1 is a major SH-PTP2 binding protein in vascular endothelial cells. *Febs Lett* 1997;408:331–336. [PubMed: 9188788]
- Newman PJ. The biology of PECAM-1. *J Clin Invest* 1997;99:3–8. [PubMed: 9011572]
- Newman PJ. Switched at birth: a new family for PECAM-1. *J Clin Invest* 1999;103:5–9. [PubMed: 9884328]
- Newman PJ, Berndt MC, Gorski J, White GC, Lyman S, Paddock C, Muller WA. PECAM-1 (CD31) cloning and relation to adhesion molecules of the immunoglobulin gene superfamily. *Science* 1990;247:1219–1222. [PubMed: 1690453]
- Newman PJ, Newman DK. Signal transduction pathways mediated by PECAM-1. New roles for an old molecule in platelet and vascular cell biology. *Arterioscler Thromb Vasc Biol* 2003;23:953–964. [PubMed: 12689916]
- Newton JP, Buckley CD, Jones EY, Simmons DL. Residues on both faces of the first immunoglobulin fold contribute to homophilic binding sites on PECAM-1/CD31. *J Biol Chem* 1997;272:20555–20563. [PubMed: 9252369]
- Nichols WC, Antin JH, Lunetta KL, Terry VH, Hertel CE, Wheatley MA, Arnold ND, Siemieniak DR, Boehnke M, Ginsburg D. Polymorphism of adhesion molecule CD31 is not a significant risk factor for graft-versus-host disease. *Blood* 1996;88:4429–44344. [PubMed: 8977234]
- Pietz BC, Warden MB, DuChateau BK, Ellis TM. Multiplex genotyping of human minor histocompatibility antigens. *Human Immunology*. 2006 (in press).
- Rousseau P, Le DM, Mouillot G, Marcou C, Carosella ED, Moreau P. The 14 bp deletion-insertion polymorphism in the 3' UT region of the HLA-G gene influences HLA-G mRNA stability. *Hum Immunol* 2003;64:1005–1010. [PubMed: 14602228]
- Sagawa K, Kimura T, Swieter M, Siraganian RP. The protein-tyrosine phosphatase SHP-2 associates with tyrosine-phosphorylated adhesion molecule PECAM-1 (CD31). *J Biol Chem* 1997;272:31086–31091. [PubMed: 9388260]
- Sasaoka T, Kimura A, Hohta SA, Fukuda N, Kurosawa T, Izumi T. Polymorphisms in the platelet-endothelial cell adhesion molecule-1 (PECAM-1) gene, Asn563Ser and Gly670Arg, associated with myocardial infarction in the Japanese. *Ann N Y Acad Sci* 2001;947:259–269. [PubMed: 11795274]
- Shifman S, Kuypers J, Kokoris M, Yakir B, Darvasi A. Linkage disequilibrium patterns of the human genome across populations. *Hum Mol Genet* 2003;12:771–776. [PubMed: 12651872]
- Simmons DL, Walker C, Power C, Pigott R. Molecular cloning of CD31, a putative intercellular adhesion molecule closely related to carcinoembryonic antigen. *J Exp Med* 1990;171:2147–2152. [PubMed: 2351935]
- Song FC, Chen AH, Tang XM, Zhang WX, Qian XX, Li JQ, Lu Q. Association of platelet endothelial cell adhesion molecule-1 gene polymorphism with coronary heart disease. *Di Yi Jun Yi Da Xue Xue Bao* 2003;23:156–158. [PubMed: 12581968]

- Stockinger H, Gadd SJ, Eher R, Majdic O, Schreiber W, Kasinrek W, Strass B, Schnabl E, Knapp W. Molecular characterization and functional analysis of the leukocyte surface protein CD31. *J Immunol* 1990;145:3889–3897. [PubMed: 1700999]
- Sun QH, DeLisser HM, Zukowski MM, Paddock C, Albelda SM, Newman PJ. Individually distinct Ig homology domains in PECAM-1 regulate homophilic binding and modulate receptor affinity. *J Biol Chem* 1996;271:11090–11098. [PubMed: 8626652]
- Wei H, Fang L, Chowdhury SH, Gong N, Xiong Z, Song J, Mak KH, Wu S, Koay E, Sethi S, Lim YL, Chatterjee S. Platelet-endothelial cell adhesion molecule-1 gene polymorphism and its soluble level are associated with severe coronary artery stenosis in Chinese Singaporean. *Clin Biochem* 2004;37:1091–1097. [PubMed: 15589815]
- Windsor L, Morahan G, Huang D, McCann V, Jones T, James I, Christiansen FT, Price P. Alleles of the IL12B 3'UTR associate with late onset of type 1 diabetes. *Hum Immunol* 2004;65:1432–1436. [PubMed: 15603869]
- Wong CW, Wiedle G, Ballestrem C, Wehrle-Haller B, Etteldorf S, Bruckner M, Engelhardt B, Gisler RH, Imhof BA. PECAM-1/CD31 trans-homophilic binding at the intercellular junctions is independent of its cytoplasmic domain; evidence for heterophilic interaction with integrin $\alpha\beta3$ in cis. *Mol Biol Cell* 2000;11:3109–3121. [PubMed: 10982404]
- Zehnder JL, Hirai K, Shatsky M, McGregor JL, Levitt LJ, Leung LLK. The cell adhesion molecule CD31 is phosphorylated after cell activation. Down-regulation of CD31 in activated T lymphocytes. *J Biol Chem* 1992;267:5243–5249. [PubMed: 1544907]

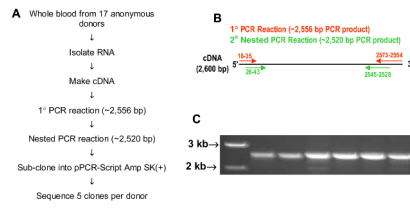


Figure 1. Protocol for determining PECAM-1 alleles

A. Flow diagram of methods used to determine PECAM-1 alleles. **B.** Schematic of primary and nested PCR reactions. **C.** Representative agarose gel of final PCR products.

THE FOUR MAJOR ALLELIC ISOFORMS OF HUMAN PECAM-1

	V ₂₃	N ₄₈	L ₈₈	N ₁₂₄	I ₅₄	D ₂₄₄	H ₄₂₇	S ₂₃₈	R ₄₃₃	² gf
¹ LSR	V ₂₃	N ₄₈	L ₈₈	N ₁₂₄	I ₅₄	D ₂₄₄	H ₄₂₇	S ₂₃₈	R ₄₃₃	0.14 ³
LSRg	V ₂₃	N ₄₈	L ₈₈	N ₁₂₄	I ₅₄	D ₂₄₄	H ₄₂₇	S ₂₃₈	R ₄₃₃	0.28
LSRg	V ₂₃	N ₄₈	V ₁₁₈	N ₁₂₄	I ₅₄	D ₂₄₄	H ₄₂₇	N ₂₃₈	G ₄₃₃	0.27
LSRg	V ₂₃	N ₄₈	V ₁₁₈	N ₁₂₄	I ₅₄	D ₂₄₄	H ₄₂₇	N ₂₃₈	G ₄₃₃	0.31

^{1,2}5'UTR and 3'UTR polymorphisms show nucleotide number based on the transcription start site of PECAM cDNA.

³gf = gene frequency. Frequencies were derived from genotyping 234 alleles from 117 unrelated volunteer blood donors, and were consistent with Hardy-Weinberg equilibrium analysis.

⁴Chi-square analysis revealed that significant differences exist in both the frequencies of the combined LSR versus VNG isoforms (p<0.025), and also between LSRa and LSRg (p<0.01).

Figure 2. The four major allelic isoforms of human PECAM-1
^{1,2}5'UTR and 3'UTR polymorphisms show nucleotide number based on the transcription start site of PECAM cDNA. ³gf = gene frequency. Frequencies were derived from genotyping 234 alleles from 117 unrelated volunteer blood donors, and were consistent with Hardy-Weinberg equilibrium analysis. ⁴Chi-square analysis revealed that significant differences exist in both the frequencies of the combined LSR versus VNG isoforms (p<0.025), and also between LSRa and LSRg (p<0.01).

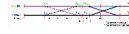


Figure 3. Likely crossing-over events yielding human PECAM-1 alleles

The 16 exons of PECAM-1 are widely dispersed amongst introns, with more than 26 kb between the R643G polymorphism in exon 12 and the a2479g nucleotide polymorphism within exon 16. A crossing-over event is likely to have taken place within this region, resulting in the generation of the four major alleles of PECAM-1. Dotted lines illustrate what was likely a second, more recent crossing-over event that has yielded two additional less common PECAM-1 allelic isoforms, previously observed by (Cavanagh et al 2005; Maruya et al 1998). Exon/Intron spacing is drawn to scale.

Table 1

Single nucleotide polymorphisms within PECAM-1 exons

Exon	1 Nucleotide base substitution	2 Amino acid substitution	3 Amino acid substitution	Frequency	References
1	g53a in 5' UTR	N/A	N/A	G 0.98, A 0.02	Elrayess (a)
3	G442A	V53M	V80M	V 0.99, M 0.01	Maruya
3	A548T	N88I	N115I	Unknown	*
3	C577G	L98V	L125V	L -0.5, V -0.5	Fang, Grumet, Listi, Maruya, Sasaoka, Wei
4	A656G	N124D	N151D	Unknown	*
6	A1329G	I348M	I375M	Unknown	*
6	A1376T	D364V	D391V	Unknown	*
7	T1656C	H457H	H484H	Unknown	*
8	G1892A	S536N	S563N	S -0.5, N -0.5	Fang, Grumet, Listi, Maruya, Sasaoka, Wei
12	A2212G	R643G	R670G	R -0.5, G -0.5	Elrayess (b), Grumet, Listi, Maruya, Sasaoka,
16	a2479g in 3' UTR	N/A	N/A	Unknown	This study

¹Numbering based on PECAM-1 transcription start site (Gumina et al 1997).²Numbering based on the mature protein sequence.³Numbering based on protein sequence including the 27 amino acid signal peptide.

N/A: not applicable – base change is within an untranslated region

(a) Elrayess MA, et al. 2003

(b) Elrayess MA, et al. 2004

* Polymorphisms deduced from differences in the originally-reported PECAM-1 cDNA sequences (Newman et al 1990; Simmons et al 1990; Stockinger et al 1990; Zehnder et al 1992)

Table 2

Diseases associated with PECAM-1 SNPs

[#] Amino acid	Disease association	Reference
g ₅₃ (5'UTR)	atherosclerosis	Elrayess ^(a)
Mismatch at L98V	GVHD	Balduini, Behar, Cavanagh, Grumet
V ₉₈	CHD or CAD	Song, Wei, Fang
N ₅₃₆	CHD or CAD	Song
V ₉₈ + N ₅₃₆	Cerebral malaria	Kikuchi
S ₅₃₆	MI	Sasaoka
Mismatch at S536N	GVHD	Cavanagh, Grumet, Maruya
R ₆₄₃	MI	Sasaoka, Listì, Elrayess ^(b)
G ₆₄₃	CHD or CAD	Elrayess ^(b)
Mismatch at R643G	GVHD	Cavanagh, Grumet, Maruya

[#]Numbering based on mature amino acid sequence

^(a)Elrayess MA, et al 2003

^(b)Elrayess MA, et al 2004