Multiple SNP Variations Occur in the Porcine CD34 Gene

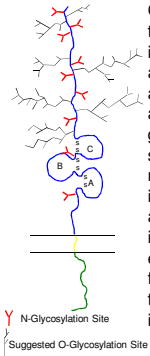
T.W. Jensen[1], D.A. Swanson[1], L.A. Rund[2], L.B. Schook[2]

[1]Institute for Genomic Biology, [2]Department of Animal Sciences, University of Illinois at Urbana-Champaign 61801

Abstract

While stem cell-based therapies to address both hard and soft tissue regeneration have shown great promise in the pig model, the lack of stem cell reagents in the pig model makes it difficult analyze stem cell populations used in this research. CD34 is widely used as a clinical marker for a subset of cells greatly enriched in hematopoietic stem cells. This marker is also associated with other stem and progenitor cells derived from adipose, endothelial progenitor cells, and skeletal muscle satellite cells. While variations in the CD34 sequence have been well studied in human and mouse, little work has been done to characterize variations in porcine CD34. We used mRNA from a single Yorkshire pig to amplify and sequence the open reading frame of porcine CD34. Using genomic DNA, the alternatively spliced Exon X was characterized. Based on sequencing of the first animal, PCR products were generated from the genomic DNA of 8 unrelated animals from each of the Sinclair, Hanford, and Yucatan minipig breeds. The frequency of single nucleotide polymorphisms (SNPs) in this sampling group was determined by sequencing the PCR products. A total of 15 SNPs were identified in exons 2, 3, 7, and 8 from the 25 animals tested. Of these SNPs, 8 resulted in changes in amino acid sequence. Three changes resulted in hydrophobic/polar shifts, one change resulted in a negative charge/polar shift, and one change alters a potential N-linked glycosylation site. Knowledge of potential variations in the translated CD34 protein will aid in the development of porcine stem cell reagents.

Background



CD34 is a highly glycosylated type I transmembrane protein whose function is associated with adhesion and adhesion modulation, although its associated ligand is not well defined and may change due to alterations in glycosylation patterns. An alternative splice variant of the gene in both the mouse and human results from the inclusion of “exon X” between exons 7 and 8. The inclusion of exon X results in a longer mRNA transcript that encodes a shorter peptide due to an in-frame stop codon in exon X that truncates the protein prior to the intracellular signaling domain.

Methods

Animals and tissue collection

Porcine tissues were collected from a Yorkshire cross pig approximately 3 months old. All work was performed with approval from University of Illinois Institutional Animal Care and Use Committee (IACUC). Blood from unrelated Sinclair, Hanford, and Yucatan minipigs was obtained from the Sinclair Research Center, Inc., Columbia, MO.

Methods cont’d

Adipose derived stromal vascular cells (ADSCs) were obtained from a single Yorkshire cross. Cells were incubated at 37ºC, 5% CO_2 in DMEM containing 10% FBS, penicillin/streptomycin, and glutamine. Genomic DNA from the Sinclair, Hanford, and Yucatan pigs was extracted from whole blood using standard techniques.

Cloning and sequencing of the porcine CD34 gene

Reagents were from Sigma Aldrich unless otherwise noted. ADSCs were cultured for 3 days prior to the addition of rhIL-1b (1ng/mL), LPS (1 mg/mL), and PMA (1ng/mL) for 3 hours. Total cellular RNA was isolated using RNeasy kit (Qiagen, Valencia, CA). Single stranded cDNA synthesis was performed using AMV reverse transcriptase (Promega, Madison, WI) and a primer specific to porcine CD34: 5’-CCG GTA CCT CAC AGT TCA GTA TCT GCC ACC-3’ and Oligo dT primers. The region encoding porcine CD34 was amplified by PCR using the forward primer 5’-CCG CTA GCA GGA TGC TCC GCA GG-3’ and reverse primer 5’-CCA AGC TTC AGT CAC AGT TCA GTA TCT GCC A-3’. Individual products were isolated on a 1% agarose TBE gel and were subcloned into pGEM-T Easy vector (Promega) for sequencing. Individual clones were sequenced at the U. of Ill. Core DNA Sequencing Facility.

• As expected, two representing and alternatively spliced variant of porcine CD34 are identified (Fig 1).

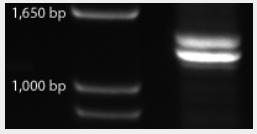


Figure 1. Splice variants in porcine CD34 gene result from the inclusion of “exon X”. (a) Two RT-PCR products are generated using CD34-specific primers spanning the open reading frame. EtBr-stained agarose gel of RT-PCR products show bands of 1,206 and 1,340 bp. (b) Sequencing of each product indicates the inclusion of an additional 144 bp in the larger product after nucleotide 1,020.

The region between porcine CD34 exon 7 and 8 was amplified from genomic DNA using forward primer 5’-GAC CTT GAT TGC ACT GGT CA-3’ and reverse primer 5’-CAC ACT GGC TTT TCC CTG AG-3’ and the amplified product was sequenced.

• Sequencing genomic DNA, the region between Exons 7 and 8 reveals the alternatively spliced “Exon X” analogous to that seen for both human and mouse CD34 (Fig 2).

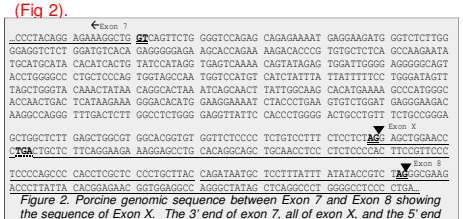
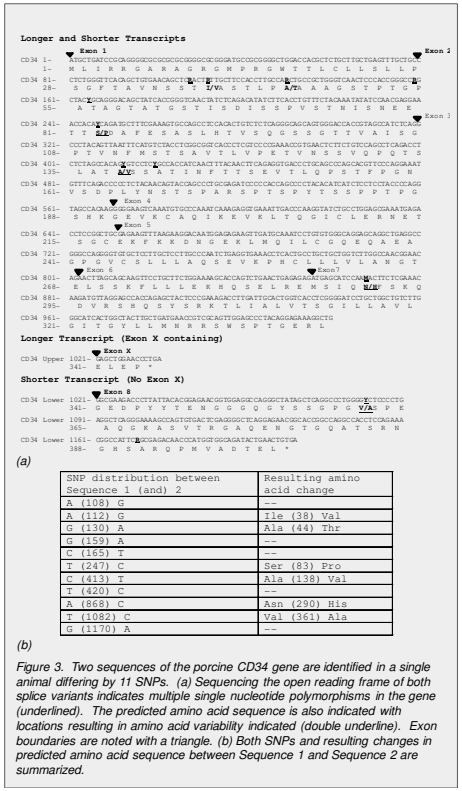


Figure 2. Porcine genomic sequence between Exon 7 and Exon 8 showing the sequence of Exon X. The 3’ end of exon 7, all of exon X, and the 5’ end of exon 8 are underlined. The consensus sequence of the splice sites (GT and AG) are indicated by double underlines. The start of Exon X and Exon 8 are indicated by triangles. The TGA stop codon that results in the truncated form of the translated protein when exon X is transcribed is indicated with a dotted underline.

Results

Individual clones of PCR products from a single Yorkshire cross were subcloning into pGEM-T Easy vector (Promega) and sequenced.

• Two distinct sequences were identified that varied by 11 SNPs (Fig 3).



(a)

SNP distribution between Sequence 1 (and) 2	Resulting amino acid change
A (108) G	--
A (112) G	Ile (38) Val
G (130) A	Ala (44) Thr
G (159) A	--
C (165) T	--
T (247) C	Ser (83) Pro
C (413) T	Ala (138) Val
T (420) C	--
A (868) C	Asn (290) His
T (1082) C	Val (361) Ala
G (1170) A	--

(b)

Figure 3. Two sequences of the porcine CD34 gene are identified in a single animal differing by 11 SNPs. (a) Sequencing the open reading frame of both splice variants indicates multiple single nucleotide polymorphisms in the gene (underlined). The predicted amino acid sequence is also indicated with locations resulting in amino acid variability indicated (double underline). Exon boundaries are noted with a triangle. (b) Both SNPs and resulting changes in predicted amino acid sequence between Sequence 1 and Sequence 2 are summarized.

The prevalence of the SNPs identified from the Yorkshire cross was tested in 3 different minipig breeds: Sinclair, Hanford, and Yucatan. Using genomic DNA, a region spanning Exon 2 and 3 of genomic DNA was PCR amplified using forward primer 5’-TCT GGG TTC ACA GCT GTG AA-3’ and reverse primer 5’-GCT ACC TGG GGT AGG AGG AG-3’. A region spanning Exon 7 and 8 was amplified with forward primer 5’-GCT TCT GGA AAA GCA CCA GT-3’ and reverse primer 5’-AGA GGC AGC ACA CGT TCA G-3’. PCR products were sequenced at the U. of Ill. Core DNA Sequencing Facility.

• All of the previously identified SNPs were also found in the Sinclair, Hanford, and Yucatan breeds. An additional four SNPs were identified that were not noted in the Yorkshire cross (Fig 4).

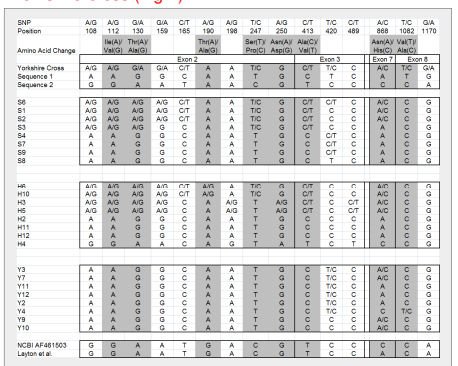


Figure 4. 15 Single nucleotide polymorphisms (SNPs) are identified in the open reading frame of the porcine CD34 gene. Sequencing results identify SNPs from mRNA from a Yorkshire cross as well as from genomic DNA from 8 unrelated minipigs in each of the Sinclair (S), Hanford (H), and Yucatan (Y) breeds. The pCD34 sequence deposited in genbank (AF461503) and that published by Layton et al. (2007) are included for comparison. SNPs resulting in amino acid changes are highlighted.

Conclusions

There are 7 identified SNPs in human CD34 resulting in three amino acid substitutions in the extracellular region of the protein (and 1 frameshift: NCBI Gene ID 947). The frequency of these SNPs is low and alterations in anti-CD34 antibody binding are have been associated with changes in glycosylation patterns and not due to SNP induced protein variations. In the cow, however, Sakurai et al. (Exp Hem, 2006) produced an antibody sensitive to a single amino acid change due to a SNP. Due to the frequency of the SNPs identified here in the porcine CD34 gene, care must be taken in generating reagents for examining porcine CD34+ stem cells. We have since used this information to generate a robust anti-porcine CD34 monoclonal antibody for the study of porcine stem cells.

Acknowledgments

- This work was funded by grants from the USDA (AG 2008-34480-19328 and 538 AG 2009-34480-19875) and by a Fellowship from Carle Foundation Hospital.

ILLINOIS