

Providing Standardized Gene Nomenclature for Domestic Ungulate Scavenger Receptor Genes

Jessica Cruz¹, Fiona McCarthy^{2§}

¹University of Arizona, Tucson, Arizona, United States

²Animal Comparative and Biomedical Sciences, University of Arizona

[§]To whom correspondence should be addressed: fionamcc@arizona.edu

Abstract

Scavenger Receptor (SCAR) genes are a heterogeneous group involved in innate immunity and immunomodulation. While standardized gene nomenclature is provided for model organism SCAR genes, ungulates have novel SCAR genes which have not been systematically classified. This study compares SCAR genes for five representative ungulate species which are selected based on their agricultural importance and representation. We identify and classify twenty-seven SCAR genes common to all species and provide standardized gene nomenclature for genes novel to these ungulate species. Notably, we provide standardized gene nomenclature for the previously studied Workshop Cluster 1 (WC1) genes which are absent in model organisms.

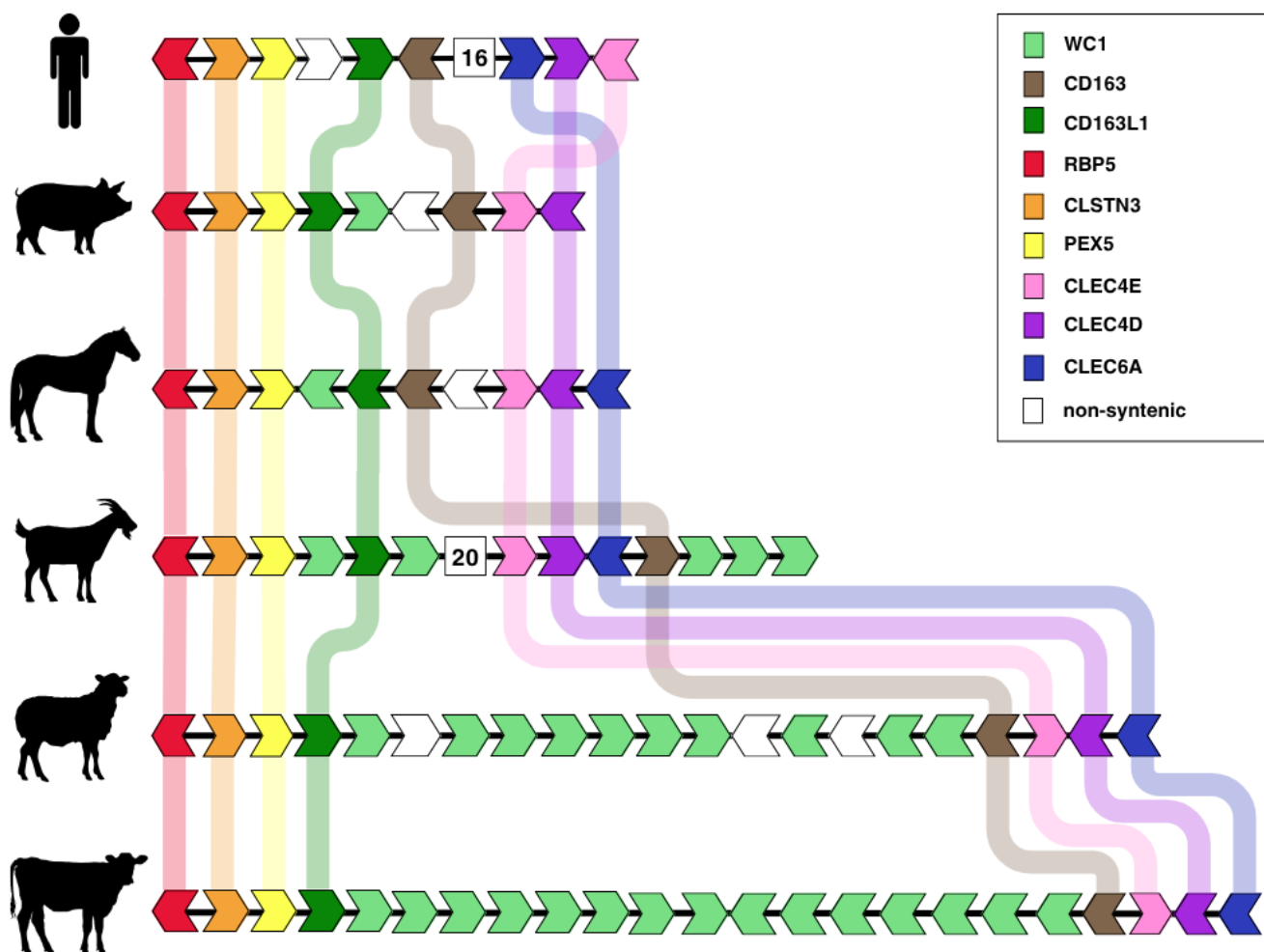


Figure 1. Comparison of gene order for the SCAR1 (WC1) gene cluster across key vertebrate species.:

Species listed from top to bottom include *Homo sapien*, *Sus scrofa*, *Equus caballus*, *Capra hircus*, *Ovis aries*, and *Bos taurus*. Genes are shown as arrows representing orientation. Orthologous genes are shown in the same color, with corresponding

shaded lines connecting them across species. Flanking genes (RBP5, CLSTN3 and PEX5 on the left, CLEC4E, CLEC4D and CLEC6A on the right) surrounding the cluster are included to demonstrate the synteny around this gene expansion. *Sus scrofa*'s gene order is based upon the Ensembl annotation, while in other species there was no difference in the gene annotation provided by Ensembl and NCBI. White arrows indicate genes that are unrelated to this study.

Description

Scavenger Receptor (SCAR) genes are a heterogeneous group of cell-surface receptor genes involved in innate immunity and immunomodulation (Alquraini & Khoury, 2020). In humans, scavenger receptors are further classified into twelve subgroups based on structure and function (Alquraini & Khoury, 2020; PrabhuDas et al., 2017). However, SCAR genes are present across a wide range of animals, and nonhuman vertebrates have an expanded repertoire of some of these receptors (Herzig et al., 2010). This project provides standardized gene nomenclature for previously unassigned SCAR genes found across multiple ungulate species. We selected cattle, horses, pigs, sheep and goats as specific of agricultural importance whose SCAR genes have been previously studied. Our experimental approach is to identify the SCAR gene repertoire in these species and provide standardized gene nomenclature which reflects what is known about the function of these genes, particularly SCAR genes not found in humans. We used sequence similarity, motif analysis and phylogenetic analysis to identify SCAR genes in cattle, horses, pigs, sheep and goats. Gene synteny and phylogenetic analysis were used to determine relationships between mammalian SCAR genes while standardized gene nomenclature was assigned using VGNC guidelines (Jones et al., 2023). For SCAR genes not found in humans we reviewed publications to determine function and assess names which have previously been used by the research community. The proposed names for these novel genes better reflect the function of their gene products and their probable roles in innate immune responses.

We identified 27 genes orthologous to human SCAR genes. In addition, we identified a novel group of SCAR genes found in ungulate species, but absent in humans, that have previously been described in cattle as Workshop Cluster 1 (WC1) genes (Morrison & Davis, 1991). WC1 genes are conserved across multiple ungulate genomes but have different copy numbers (16 in cattle, 13 in sheep, 8 in goat, 9 in pig and 1 in horse). Sequence searches of vertebrate proteins demonstrated that WC1 proteins are found in all mammals except primates and rodents. These genes were previously determined to be structurally and functionally related to human CD163 and CD163L1 (Herzig et al., 2010). WC1 genes are expressed on gamma delta T cells as transmembrane co-stimulatory receptors involved in pathogen pattern recognition and inflammatory response initiation (Hedges et al., 2003; Wang et al., 2011). Figure 1 shows the conserved WC1 cluster in ungulates, depicting this gene expansion which is absent in the human genome. In addition to this cluster of WC1 genes in the ungulates, cattle, goat and pig have additional WC1 genes which are located on unplaced contigs, which can be placed within this WC1 cluster in their respective genomes.

While most of the species in this study show synteny with the human genes in this region, the downstream gene order in the current goat genome is not well conserved, with the WC1 cluster split by the conserved CLEC and CD163 genes. In addition, three goat WC1 genes are found on unplaced contigs within the goat genome (*Capra hircus* ARS1.2). The unplaced sequences in the WC1 cluster indicate poor assembly in this region, which may also impact the observed lack of gene synteny, particularly in comparison to the gene synteny in the closely related sheep genome. Likewise, the gene synteny in pigs does not match the observed gene order when compared to other ungulates. Pig CD163 is annotated downstream of the single placed WC1 gene while CD163L1 is annotated upstream, and this section of the genome appears to be flipped in the NCBI annotation. Both the NCBI and Ensembl annotations of the pig genome do not include CLEC6A. While only one WC1 gene is placed in this region, we identified seven additional WC1 genes on unplaced contigs, and this is in agreement with previous studies (Le Page et al., 2018; Le Page et al., 2019). This serves as further evidence that this region has assembly errors in the pig genome (*Sus scrofa* Sscrofa11.1). We note that the gene order in the Ensembl assembly (Sscrofa11.1) is accurate when compared to our observed genomes, as shown in Figure 1.

WC1 is expressed in a subpopulation of CD4⁺ CD8⁺ T cells called gamma delta T cells (Morrison & Davis, 1991). In humans, most T cells develop in the thymus and have a T cell receptor (TCR) composed of alpha beta chains (Fozza et al., 2017). However, a small population of gamma delta T cells localized to peripheral tissues (skin, mucosa, intestine) was identified in humans, differentiated in receptor structure via gamma and delta chains (Bank et al., 1986; Brenner et al., 1986; Hu et al., 2023). The gamma delta TCR can rapidly and directly interact with various pathogens without prior antigen exposure (Tanaka et al., 1995), exhibiting characteristics of both innate and adaptive immunity. Much like the alpha beta TCR, the gamma delta TCR signals through the CD3 complex, which binds to antigen-presenting non-classical MHC and related molecules (Luoma et al., 2013). However, the gamma delta TCR may also recognize pathogens directly, without the use of antigen-presenting molecules, and may initiate an immune response on its own (Chien & Konigshofer, 2007; Morath & Schamel, 2020). WC1 genes contribute to immunity as gamma delta T cell-specific co-receptors involved in varied pathogen recognition and cell activation accompanied by the CD3 complex (Hsu et al., 2015). Vertebrate species with an expanded repertoire of WC1 genes

have a larger subpopulation of gamma delta T cells in the peripheral blood, ranging from 30-50% of T cells especially in young ruminants (Hein & Mackay, 1991). In primates and rodents which naturally lack WC1 genes, CD163L1 or SCART1 respectively act as functional equivalents to WC1 (Herzig et al., 2010).

Based upon scavenger receptor classification, we propose that WC1 genes be labelled as scavenger receptor class I member (SCAR1). We propose gene nomenclature in line with the recommended protein nomenclature (PrabhuDas et al., 2017). We applied this numbering system, and numbered the additional ungulate genes according to gene synteny and phylogenetic inference. This system of standardized nomenclature agrees with VGNC guidelines (Jones et al., 2023) and feedback from community experts, and is extensible to genes in other species.

Methods

Human SCAR genes were defined using the HGNC gene group set. We identified orthologous SCAR genes between human and representative domestic ungulates (*Equus caballus* EquCab3.0, *Sus scrofa* Sscrofa11.1, *Bos taurus* ARS-UCD2.0, *Ovis aries* ARS-UI_Ramb_v3.0, and *Capra hircus* ARS1.2) using NCBI and Ensembl ortholog prediction resources. In cases with a direct 1:1 ortholog, standardized gene nomenclature was applied to the ungulate gene based upon the corresponding human gene nomenclature. Additional names used for these genes in the published literature were collected by manual review of the literature associated with each gene.

Additional ungulate SCAR genes with no human orthologs were identified by sequence similarity searches using human SCAR proteins to search the NCBI species-specific protein databases, followed by manual review to identify proteins with scavenger receptor motifs. Ungulate SCAR genes with no human ortholog were assigned standardized nomenclature based upon gene synteny, phylogenetic analysis, and review of associated literature. The gene order of WC1 genes were manually reviewed, using three conserved mammalian genes on either side of the gene cluster; CLEC6A, CLEC4D, CLEC4E and PEX5, CLSTN3, RBP5.

Nomenclature standardization from amongst ungulate species is based upon phylogenetic analysis and review of gene order. The longest protein sequence was selected as the representative protein sequence for phylogenetic analysis. Protein sequences were imported into Geneious Prime 2024.0.7 and aligned using Geneious Alignment (with free end gaps and matrix Blosum62). Phylogenetic trees were created from the alignment through the Geneious Tree Builder (Jukes-Cantor, Neighbor-Joining) and 100 bootstrap replicates. Bovine WC1 genes were numbered in line with established protein nomenclature (PrabhuDas et al., 2017), with numbers not identified in the current bovine annotation reserved for future use. Additional bovine WC1/SCAR1 genes were then numbered based upon their gene order and annotation. Following the assignment of gene order in bovine, phylogenetic analysis was performed to determine likely orthologs in the remaining ungulates. Strict orthologs were assigned standardized names, while the remaining genes were numbered consecutively according to gene order along equine chromosome 6, swine chromosome 5, ovine chromosome 3 and caprine chromosome 5. Genes found on unassigned sequences were named lastly in respective assemblies.

Acknowledgements: JC was supported by the University of Arizona and Place-Based Health Informatics Research Education (PHIRE), which is part of the Undergraduate Biology Research Program. The authors would like to thank Drs. Elspeth Bruford, Cynthia Baldwin and Janice Telfer for their expertise on naming the WC1 genes.

Extended Data

Description: This dataset presents the orthologs of the SCAR and related WC1 genes across genome assemblies described within this study. Provided are the NCBI and Ensembl gene IDs, current gene nomenclature at the time of this study, our proposed nomenclature, and known aliases.. Resource Type: Dataset. File: [SCAR Orthologs.txt](#). DOI: [10.22002/fbz4s-mdw82](#)

References

- Alquraini A, El Khoury J. 2020. Scavenger receptors. *Current Biology* 30: R790-R795. DOI: [10.1016/j.cub.2020.05.051](#)
- Bank I, DePinho RA, Brenner MB, Cassimeris J, Alt FW, Chess L. 1986. A functional T3 molecule associated with a novel heterodimer on the surface of immature human thymocytes. *Nature* 322: 179-181. DOI: [10.1038/322179a0](#)
- Brenner MB, McLean J, Dialynas DP, Strominger JL, Smith JA, Owen FL, et al., Krangel. 1986. Identification of a putative second T-cell receptor. *Nature* 322: 145-149. DOI: [10.1038/322145a0](#)
- Chien Yh, Konigshofer Y. 2007. Antigen recognition by $\gamma\delta$ T cells. *Immunological Reviews* 215: 46-58. DOI: [10.1111/j.1600-065X.2006.00470.x](#)

- Fozza C, Barraqueddu F, Corda G, Contini S, Viridis P, Dore F, Bonfigli S, Longinotti M. 2017. Study of the T-cell receptor repertoire by CDR3 spectratyping. *Journal of Immunological Methods* 440: 1-11. DOI: [10.1016/j.jim.2016.11.001](https://doi.org/10.1016/j.jim.2016.11.001)
- Hedges JF, Cockrell D, Jackiw L, Meissner N, Jutila MA. 2003. Differential mRNA expression in circulating $\gamma\delta$ T lymphocyte subsets defines unique tissue-specific functions. *Journal of Leukocyte Biology* 73: 306-314. DOI: [10.1189/jlb.0902453](https://doi.org/10.1189/jlb.0902453)
- Hein WR, Mackay CR. 1991. Prominence of $\gamma\delta$ T cells in the ruminant immune system. *Immunology Today* 12: 30-34. DOI: [10.1016/0167-5699\(91\)90109-7](https://doi.org/10.1016/0167-5699(91)90109-7)
- Herzig CT, Waters RW, Baldwin CL, Telfer JC. 2010. Evolution of the CD163 family and its relationship to the bovine gamma delta T cell co-receptor WC1. *BMC Evolutionary Biology* 10: 10.1186/1471-2148-10-181. DOI: [10.1186/1471-2148-10-181](https://doi.org/10.1186/1471-2148-10-181)
- Hsu H, Chen C, Nenninger A, Holz L, Baldwin CL, Telfer JC. 2015. WC1 Is a Hybrid $\gamma\delta$ TCR Coreceptor and Pattern Recognition Receptor for Pathogenic Bacteria. *The Journal of Immunology* 194: 2280-2288. DOI: [10.4049/jimmunol.1402021](https://doi.org/10.4049/jimmunol.1402021)
- Hu Y, Hu Q, Li Y, Lu L, Xiang Z, Yin Z, Kabelitz D, Wu Y. 2023. $\gamma\delta$ T cells: origin and fate, subsets, diseases and immunotherapy. *Signal Transduction and Targeted Therapy* 8: 10.1038/s41392-023-01653-8. DOI: [10.1038/s41392-023-01653-8](https://doi.org/10.1038/s41392-023-01653-8)
- Jones TEM, Yates B, Braschi B, Gray K, Tweedie S, Seal RL, Bruford EA. 2023. The VGNC: expanding standardized vertebrate gene nomenclature. *Genome Biology* 24: 10.1186/s13059-023-02957-2. DOI: [10.1186/s13059-023-02957-2](https://doi.org/10.1186/s13059-023-02957-2)
- Le Page LA, Buck J, Boisvert N, Gillespie A, Hudgens E, Hsu H, Baldwin CL, Telfer JC. 2018. Classification the WC1 gene family in *Sus scrofa* and evaluation of individual SRCR domain affinity for *Mycobacterium bovis* and *Leptospira spp.*. *The Journal of Immunology* 200: 59.18-59.18. DOI: [10.4049/jimmunol.200.Supp.59.18](https://doi.org/10.4049/jimmunol.200.Supp.59.18)
- Le Page LA, Buck J, Boisvert N, Gillespie A, Yirsaw A, Hudgens E, et al., Telfer. 2019. Annotation of the WC1 gene family in *Sus scrofa* and evaluation of individual SRCR domain affinity for *Mycobacterium bovis* and *Leptospira spp.*. *The Journal of Immunology* 202: 73.23-73.23. DOI: [10.4049/jimmunol.202.Supp.73.23](https://doi.org/10.4049/jimmunol.202.Supp.73.23)
- Luoma AM, Castro CD, Mayassi T, Bembinster LA, Bai L, Picard D, et al., Adams. 2013. Crystal Structure of V δ 1 T Cell Receptor in Complex with CD1d-Sulfatide Shows MHC-like Recognition of a Self-Lipid by Human $\gamma\delta$ T Cells. *Immunity* 39: 1032-1042. DOI: [10.1016/j.immuni.2013.11.001](https://doi.org/10.1016/j.immuni.2013.11.001)
- Morath A, Schamel WW. 2020. $\alpha\beta$ and $\gamma\delta$ T cell receptors: Similar but different. *Journal of Leukocyte Biology* 107: 1045-1055. DOI: [10.1002/JLB.2MR1219-233R](https://doi.org/10.1002/JLB.2MR1219-233R)
- Morrison WI, Davis WC. 1991. Differentiation antigens expressed predominantly on CD4⁺ CD8⁺ T lymphocytes (WC1, WC2). *Veterinary Immunology and Immunopathology* 27: 71-76. DOI: [10.1016/0165-2427\(91\)90082-N](https://doi.org/10.1016/0165-2427(91)90082-N)
- PrabhuDas MR, Baldwin CL, Bollyky PL, Bowdish DME, Drickamer K, Febbraio M, et al., El Khoury. 2017. A Consensus Definitive Classification of Scavenger Receptors and Their Roles in Health and Disease. *The Journal of Immunology* 198: 3775-3789. DOI: [10.4049/jimmunol.1700373](https://doi.org/10.4049/jimmunol.1700373)
- Tanaka Y, Morita CT, Tanaka Y, Nieves E, Brenner MB, Bloom BR. 1995. Natural and synthetic non-peptide antigens recognized by human $\gamma\delta$ T cells. *Nature* 375: 155-158. DOI: [10.1038/375155a0](https://doi.org/10.1038/375155a0)
- Wang F, Herzig CTA, Chen C, Hsu H, Baldwin CL, Telfer JC. 2011. Scavenger receptor WC1 contributes to the $\gamma\delta$ T cell response to *Leptospira*. *Molecular Immunology* 48: 801-809. DOI: [10.1016/j.molimm.2010.12.001](https://doi.org/10.1016/j.molimm.2010.12.001)

Funding: Research reported in this publication was supported by the National Library of Medicine (NLM) of the National Institutes of Health under award number R25LM01422. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author Contributions: Jessica Cruz: investigation, writing - original draft, writing - review editing, data curation, methodology. Fiona McCarthy: conceptualization, project administration, writing - review editing, visualization.

Reviewed By: Anonymous

History: Received March 22, 2025 **Revision Received** June 20, 2025 **Accepted** June 20, 2025 **Published Online** June 24, 2025 **Indexed** July 8, 2025

Copyright: © 2025 by the authors. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

6/24/2025 - *Open Access*

Citation: Cruz J, McCarthy F. 2025. Providing Standardized Gene Nomenclature for Domestic Ungulate Scavenger Receptor Genes. microPublication Biology. [10.17912/micropub.biology.001584](https://doi.org/10.17912/micropub.biology.001584)