



# Evidence for the loss and recovery of SLAMF9 during human evolution: implications on Dollo's law

Maegan K. Murphy<sup>1</sup> · Justin T. Moon<sup>1</sup> · Alexis T. Skolaris<sup>1</sup> · Joseph A. Mikulin<sup>1</sup> · Timothy J. Wilson<sup>1</sup> 

Received: 1 January 2021 / Accepted: 5 February 2021 / Published online: 22 February 2021  
© The Author(s), under exclusive licence to Springer-Verlag GmbH, DE part of Springer Nature 2021

## Abstract

Signaling lymphocyte activation molecule family member 9 (SLAMF9) is a cell surface protein of the CD2/SLAM family of leukocyte surface receptors. It is conserved throughout mammals and has roles in the initiation of inflammatory responses and regulation of plasmacytoid dendritic cell function. Through comparison of reference sequences encoding SLAMF9 in human, mouse, and primate sequences, we have determined that the SLAMF9 gene underwent successive mutation events, resulting in the loss of the protein and subsequent recovery of a less stable version. The mutations included a single base pair deletion in the second exon and a change in the splice acceptor site of that same exon. These changes would have had the effect of creating and later repairing a frameshift in the coding sequence. These events took place since the divergence of the human lineage from the chimpanzee-human last common ancestor and represent the first known case of the functional loss and recovery of a gene within the human lineage.

**Keywords** SLAMF9 · Primate evolution · Receptor structure · Leukocyte

## Introduction

Pathogen-driven selection has been a critical factor during human speciation. Selective pressures applied by ancient pathogens such as *Plasmodium falciparum* and *Mycobacterium tuberculosis* left profound signatures on genes directly or indirectly involved in aspects of the immune response, including inflammation (Barreiro and Quintana-Murci 2010; Domínguez-Andrés and Netea 2019; Karlsson et al. 2014). Indeed, studies have implicated pathogens as a principal driving force of human evolution (Fumagalli et al. 2011). Pathogen-driven selection may select for either gain or loss of function variants. Among the best documented instances of positive selection for gene loss involves the anti-inflammatory regulator caspase-12, the null variant of which is fixed in many current human populations (Barreiro and Quintana-Murci 2010; Karlsson et al. 2014; Xue et al. 2006). Gain of function mutations are also associated with improved outcomes to infection. CEACAM3, which promotes phagocytosis in polymorphonuclear cells, is one of the fastest-evolving genes

in the human genome and has gained the capacity to bind bacterial outer membrane proteins since the divergence of humans and gorillas (Adrian et al. 2019). Although ancient selective pressures tended to promote positive selection of pro-inflammatory gene variants, the selected variations are not necessarily advantageous in all settings. Genes that promote resistance to a certain set of pathogens may be detrimental in the context of another pathogen. TLR4, critical for innate cell sensing of gram-negative bacteria, is exploited for immune evasion by mammalian mammary tumor virus (Wilks et al. 2015). In mice, CEACAM1 facilitates coronavirus binding and infection, which has prompted selection for variants that lack this binding capacity (Peng et al. 2017). Moreover, some alleles associated with pathogen resistance also contribute to autoimmune or inflammatory disorders including inflammatory bowel disease (Cagliani et al. 2013; Jostins et al. 2012). Thus, the specific evolutionary pressures that drive the evolutionary arms race between hosts and pathogens vary depending on the presence or absence of distinct pathogens in a population (Barreiro and Quintana-Murci 2010; Domínguez-Andrés and Netea 2019; Karlsson et al. 2014).

In particular, genes with immunoregulatory functions are sensitive to pathogen-related selective pressures. Immunomodulatory receptors including PILRB, LILRB1, SIRP1- $\alpha$ , killer cell Ig-like receptors (KIRs),

✉ Timothy J. Wilson  
tim.wilson@miamioh.edu; wilso265@miamioh.edu

<sup>1</sup> Department of Microbiology, Miami University, 700 E. High Street, Oxford, OH 45056, USA

and leukocyte receptor complex (LRC) receptors have undergone particularly rapid evolution in the human genome (Barclay and Hatherley 2008; Schwartz et al. 2019; Wilson et al. 2006). Processes including gene conversion, exon shuffling, duplication, and mobile element insertion contributed to the diversification of these receptors in the human genome (Wilson et al. 2006). Novel functions, signaling motifs, and ligands conferred by these mechanisms contribute to the fine-tuning of the immune response and co-evolve with ligands such as MHC molecules. The quickly evolving nature of these and other immunomodulatory receptors may contribute to fitness against pathogenic challenge (Schwartz et al. 2019).

The SLAM family of receptors is implicated in various immunoregulatory roles, enforcing both positive and negative regulation in leukocyte activation. Structurally, the majority of these receptors contain an amino-terminal V-set immunoglobulin superfamily (IgSF) domain, C2 IgSF domain, transmembrane region, and cytoplasmic tail (Cannons et al. 2011; Dragovich and Mor 2018; Wu and Veillette 2016). While the canonical SLAM family members, SLAMF1–7, are expressed predominantly in leukocytes and signal through immunoreceptor tyrosine-based switch motifs (ITSMs), SLAMF9 lacks any recognized intracellular signaling motif in its short, lysine-rich cytoplasmic tail (Fennelly et al. 2001; Volkova et al. 2014; Zhang et al. 2001). SLAMF9 in mice and humans is expressed on macrophages, monocytes, conventional and plasmacytoid dendritic cells, and B1 cells (Dolft et al. 2018; Fennelly et al. 2001; Sever et al. 2019; Wilson et al. 2020; Zeng et al. 2018; Zhang et al. 2001). Functional studies of SLAMF9 revealed its roles in limiting macrophage migration, mediating plasmacytoid development and activation, positively regulating pathogenic and protective inflammatory responses, and clearing *Salmonella enterica* infection in the liver (Dolft et al. 2018; Sever et al. 2019; Wilson et al. 2020; Zeng et al. 2018). Although human and mouse SLAMF9 share similar tissue expression patterns and extensive sequence homology, human SLAMF9 exhibits lower protein expression levels than mouse SLAMF9 (Wilson et al. 2020). Interrogation of differences in the protein sequence between human SLAMF9 and other mammals suggested that intrinsic instability diminished human SLAMF9 expression levels and prompted our investigation of the phylogenetic origins of potentially destabilizing variations. Unexpectedly, our analysis revealed that the human transcript contains two mutations unique among mammals, one disrupting and another restoring the reading frame encoding the N-terminus of the protein, presenting the possibility that SLAMF9 was lost and regained recently in human speciation.

## Materials and methods

### Bioinformatic analysis of SLAMF9 sequences

Sequences for SLAMF9 orthologs were retrieved from NCBI or Ensembl databases (Table 1). Sequence data for Denisovans and Neanderthals was obtained from the Ancient Genome Browser (<https://www.eva.mpg.de/genetics/genome-projects/neandertal/draft-neandertal-genome/data.html>). Nucleic acid and protein sequence alignments were performed either manually or using NCBI and Ensembl BLAST searches. Predictions of RNA secondary structure for minimum free energy structure of positions 101–601 of the *Homo sapiens* and *Pan troglodytes* pre-mRNA transcripts were analyzed using the in silico RNA secondary structure prediction web server, mFold version 3 at [www.unafold.org/mFold/applications/Structure-display-and-free-energy-determination.php](http://www.unafold.org/mFold/applications/Structure-display-and-free-energy-determination.php) (Zuker 2003). Default constraints for a linear structure were applied to the prediction. The ESEfinder3.0 web server was used to search exon 2 cDNA sequences of human and chimpanzee SLAMF9 for ESE motifs. The SRProtein matrix and default settings were applied (Cartegni et al. 2003; Smith et al. 2006). For analysis of signal sequence cleavage site probability, human, mouse, and chimpanzee protein sequences were analyzed on SignalP 5.0 (<http://www.cbs.dtu.dk/services/SignalP/abstract.php#5.0>) (Almagro Armenteros et al. 2019).

### Constructs

The FLAG-tagged human SLAMF9 cDNA sequence was described previously (Wilson et al. 2020). Briefly, SLAMF9 fused with an N-terminal FLAG sequence and leader peptide derived from pFLAG-CMV-3 was ligated into the pLenti-CMV-GFP-Puro vector (Gift of Eric Campeau; Addgene #17448). For generation of the ancestral N-terminal mutant, a cDNA sequence encoding human SLAMF9-FLAG with the N-terminal frameshifts removed and restored to identity

**Table 1** Curated accession numbers used for sequence comparisons

Species	mRNA accession
<i>Homo sapiens</i>	NM_033438.4
<i>Pan troglodytes</i>	NM_001302657.2
<i>Papio anubis</i>	NM_001302659.1
<i>Macaca mulatta</i>	NM_001302666.1
<i>Chlorocebus sabaues</i>	NM_001302660.1
<i>Saimiri boliviensis</i>	NM_001302663.1
<i>Callithrix jacchus</i>	NM_001302667.1
<i>Mus musculus</i>	NM_029612.4
	Genomic accession
<i>Homo sapiens</i>	AL513485.10
<i>Pan troglodytes</i>	Ensembl.org; Pan_tro_3.0

with the chimpanzee sequence was synthesized by Genscript and ligated into the pLenti vector described above.

**Cell lines and transfections**

HEK293T cells (ATCC CRL-3216) were cultured in RPMI supplemented with 5% FCS, 100 units/mL penicillin, 100 units/mL streptomycin, 1 mM sodium pyruvate, 1 mM Glutamax, and 0.00025% 2-mercaptoethanol. Confluent 293T cells were transiently transfected with SLAMF9-expressing vector using 25 kDa-branched polyethylenimine (Sigma, St. Louis, MO) and 2 μg of lentiviral vector encoding human SLAMF9 or ancestral N-terminus SLAMF9. Forty-eight to 72 h after transfection, 293T cells were subjected to flow cytometric analysis.

**Flow cytometry**

Transfected 293T cells were stained for 30 min on ice with 2 μg/mL biotinylated anti-FLAG L5 antibody (BioLegend, San Diego, CA) or isotype control antibody, followed by a 30-min incubation on ice with 0.5 μg/mL APC-conjugated streptavidin and Zombie Aqua viability dye (Biolegend). Sample analysis was conducted on an LSR-II flow cytometer (Becton Dickinson, East Rutherford, NJ). FLAG expression levels on live, single cells were analyzed using FlowJo software (TreeStar, Ashland, OR).

**Results**

**Human SLAMF9 N-terminus is distinct from ancestral sequences**

While attempting to study SLAMF9 protein function, efforts to express soluble or surface-expressed mouse SLAMF9 (mSLAMF9) were considerably more successful than similar strategies for human SLAMF9 (hSLAMF9) expression, with yield of soluble hSLAMF9 consistently too low for functional study. While reduced hSLAMF9 surface expression might

suggest several causes, including enhanced receptor-mediated endocytosis or impaired membrane trafficking, the additional failure to express the soluble ectodomain suggested that hSLAMF9 was intrinsically less stable than mSLAMF9. Because protein instability often results in removal by proteostatic regulatory mechanisms, this hypothesis would account for decreased expression of the human protein (Jackson and Hewitt 2016). Sequence alignments of human and mouse SLAMF9 revealed significant differences in the N-terminus of the protein (Fig. 1). Most strikingly, a twelve-residue sequence encompassing the last three amino acids of the signal peptide and the initial nine amino acids of the N-terminal V-set immunoglobulin domain diverged between the species (Fig. 1). Among the alterations to the N-terminus are the changes of an acidic peptide sequence to a basic one, the addition of bulky non-polar amino acids, and a new cysteine residue. In an effort to trace the evolutionary origins of this divergence, the N-terminal SLAMF9 sequences of old and new-world primates were compared with human and mouse sequences. Surprisingly, the N-terminus of the mouse sequence more closely resembles the primate protein sequences than does the human sequence (Fig. 1). Therefore, the altered N-terminal sequence was concluded to have arisen after the divergence of the human and chimpanzee lineages. In order to assess the implications of the human N-terminal sequence on the efficiency of signal peptide cleavage, the sequences for human, chimpanzee, and mouse SLAMF9 were subjected to analysis on SignalP5 (Almagro Armenteros et al. 2019). The probability of signal peptidase-mediated cleavage of the mouse sequence between Gly<sup>17</sup> and Phe<sup>18</sup> was 0.8006, and the probability of cleavage of the chimpanzee sequence between Gly<sup>18</sup> and Asp<sup>19</sup> was reported as 0.9424. In marked contrast, the highest probability of signal peptidase cleavage in human SLAMF9 occurred between Arg<sup>19</sup> and Arg<sup>20</sup> and had a value of 0.4428. These data suggest that the efficiency of signal peptide cleavage may modulate protein maturation and stability of SLAMF9. Altered splice acceptor site and single base-pair insertion resulted in-frame translation of SLAMF9 second exon.

**Fig. 1** Human SLAMF9 N-terminus diverges from homologous sequences. Alignment of the N-termini of SLAMF9 sequences including mice, new and old-world primates, and humans. The divergent segment of human SLAMF9 is indicated with a red box, while underlined sequences **a** and **b** represent the putative stands of the V-set IgSF domain

Species	Sequence
<i>Mus musculus</i>	MGALLWSLL-LLLQEAKG-FSGDDEDPEEVVGVLQESINLS
<i>Callithrix jacchus</i>	MYAFPWLLLLLLLLLQEAKG-GSGDDVDPEEVVAVLQESISLP
<i>Saimiri boliviensis</i>	MYAFPWLLLFLLLLQEAKG-DSGDDVDPEEVVAVLQESIILP
<i>Chlorocebus sabaeus</i>	MCAFTWLLLLLLLLLQEAKG-DSGDGVDPEEVVAVLHESISLP
<i>Macaca mulatta</i>	MCAFTWLLLLLLLLLQEAKG-DSGDGVDPEEVVAVLHESISLP
<i>Papio anubis</i>	MCAFTWLLLLLLLLLQEAKG-DSGDGMDPEEVVAVLHESISLP
<i>Pan troglodytes</i>	MCAFPWLLLLLLLLLQEAKG-DSGDGVDPEEVVAVLQESISLP
<i>Homo sapiens</i>	MCAFPWLLLLLLLLLQEGSQRRLLWRWCGSEEVVAVLQESISLP

└──────────────────┘ └──────────┘ └──────────┘  
 Signal peptide                                    **a**                                    **b**

To interrogate the genetic and evolutionary origins of the disparate N-terminal protein sequences of human and ancestral SLAMF9, the pre-mRNA sequences of human, Denisovan, Neanderthal, and chimpanzee SLAMF9 primary transcripts were compared. Specifically, the second exons were examined, since this region encodes the altered amino acid residues. Genomic sequence comparison revealed that humans, Denisovans, and Neanderthals shared 100% sequence similarity in the N-terminal region that differs between humans and chimpanzees (Fig. 2). This indicates that a common ancestor of humans, Neanderthals, and Denisovans but not chimpanzees was the origin of the divergent N-terminal sequence. Alignment of human and chimpanzee primary transcripts revealed that the human 5' splice acceptor site for the second exon was shifted 4 positions upstream of the chimpanzee sequence splice acceptor, inducing a frameshift. The sequence variation g.621delT, occurring 32 positions downstream of the ancestral splice acceptor site, results in a second frameshift that restores the translation of the second exon downstream of the deletion mutation, albeit with a more hydrophobic mature N-terminus (Fig. 2).

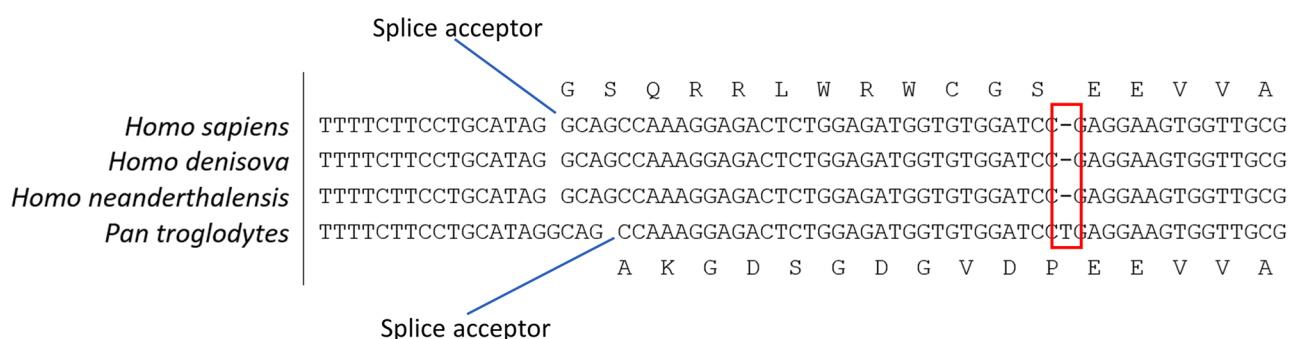
The alteration of the splice acceptor site in conjunction with the g.621delT sequence variation suggests two possible evolutionary origins of hSLAMF9. One explanation entails the loss of SLAMF9 in a human ancestor that emerged after divergence from chimpanzees, followed by the regain of SLAMF9 function by a subsequent mutation in an ancestor that arose prior to the adaptive radiation of humans, Neanderthals, and Denisovans. Alternatively, the g.621delT variation may have disrupted, created, or shifted an exonic splice enhancer element, thereby promoting the usage of a cryptic splice site-4 positions from the ancestral constitutive splice acceptor.

## Variations in SLAMF9 intron 1 result in alterations to intronic primary structure among hominins

To distinguish between these alternate explanations, it was necessary to determine whether differences between SLAMF9 intron 1 in chimpanzees and humans were sufficient to cause perturbations to RNA secondary structure that would potentially influence spliceosome binding and splice acceptor site usage. Within the first intron, human sequence variations from the chimpanzee sequence include the sequence variations IVS1+9G>A, IVS1+39delT, IVS1+41C>A, IVS1+157T>G, IVS1+255G>T, and IVS1+292C>T (Fig. 3). Secondary structure analysis conducted on the mFold web server revealed that the intronic mutations were sufficient to alter the minimum free energy structure of intron 1 (Fig. 3). Importantly, the sequence containing g.621delT was excluded from the structure prediction. Furthermore, when the human and chimpanzee exon 2 sequences were searched for exon splice enhancers on ESEfinder3.0, the sequence flanking the g.621delT sequence variation was not identified as a probable splice enhancer in either species (data not shown). Although further in vitro studies would be required to experimentally rule out the possibility that the g.621delT variation is sufficient to drive cryptic splice site usage, in silico data support the hypothesis that SLAMF9 expression was silenced and subsequently regained in the course of human evolution.

## Replacement of human SLAMF9 N-terminus with the ancestral N-terminus improves surface expression

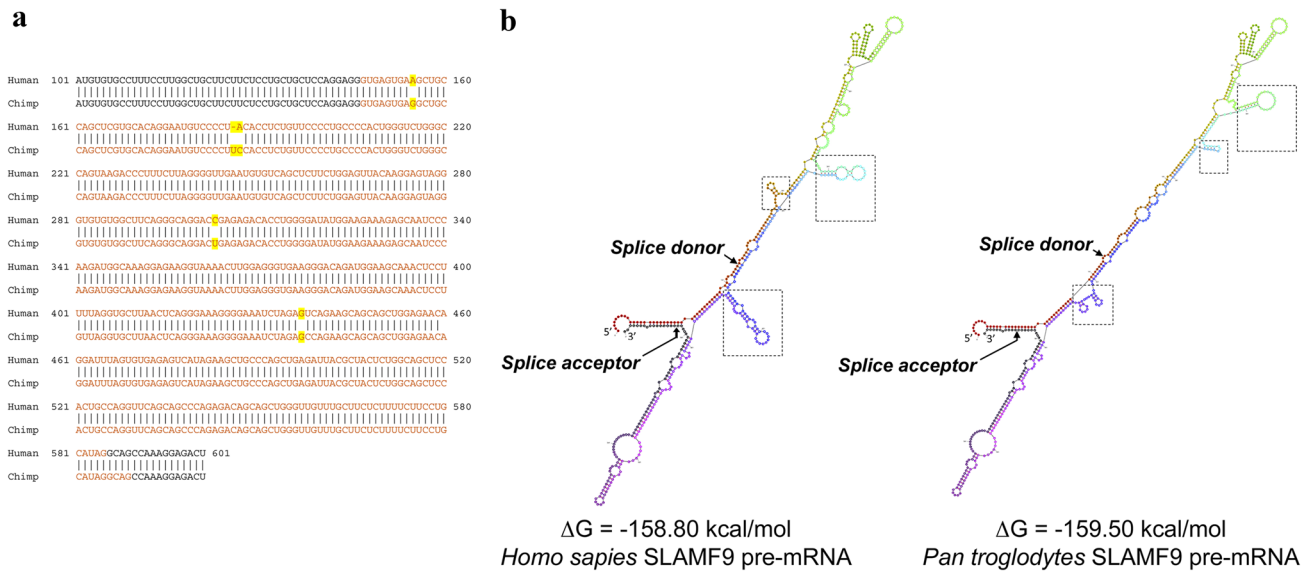
The authors have not found it possible to generate usable amounts of soluble human SLAMF9 for biochemical studies, whereas mouse SLAMF9 can readily be produced in sufficient quantities for biochemical screening and immunization (Wilson



**Fig. 2** *Homo* genus SLAMF9 transcript contains single-nucleotide deletion and de novo splice acceptor site. Sequence alignments of genomic sequence for the 5' end of exon 2 and flanking intronic sequence were compared for the indicated species. Splice acceptor sites are indicated by the gap and blue line at the boundary between the intron and exon sequences. The location of splice sites for human

and chimpanzee are confirmed based on mRNA reference sequences and predicted for *denisova* and *neanderthalensis* based on homology with *sapiens*. Coded amino acids for human and chimpanzee SLAMF9 are indicated above and below the genomic sequences, respectively. The deletion g.621delT is indicated by the red box





**Fig. 3** Intron sequence alignment and secondary structure for SLAMF9 variants. Sequence alignment (a) and secondary structure predictions (b) for positions 101–600 corresponding to the first intron and flanking sequence of the *Homo sapiens* and *Pan troglodytes* pre-mRNA transcripts of *SLAMF9*. Predicted differences in hairpin structures are

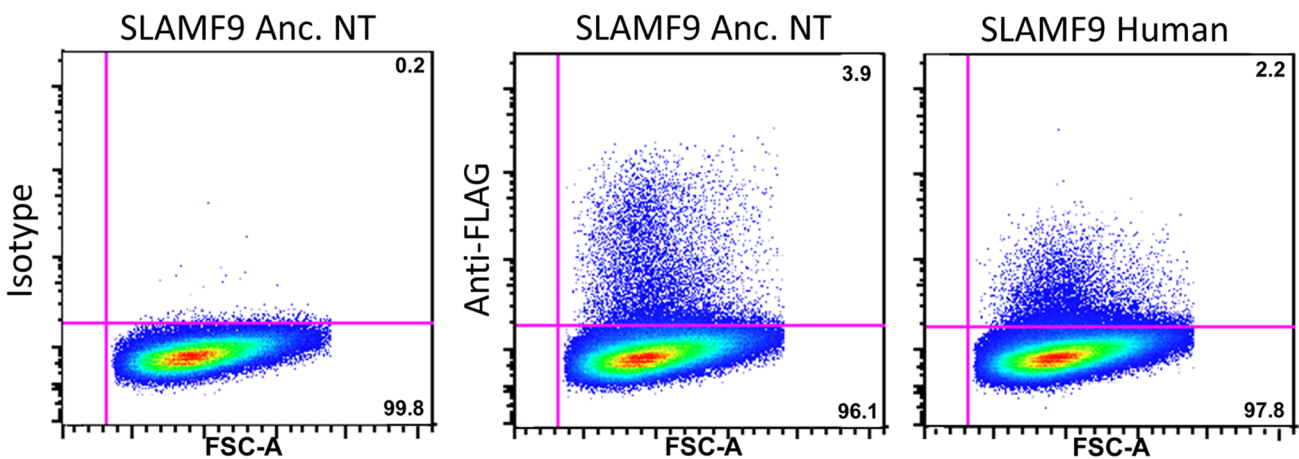
outlined in dashed boxes. Minimum free energy secondary structures of the pre-mRNA transcripts were generated using the mFold web server v. 3 for nucleic acid folding and hybridization prediction

et al. 2020). To determine whether N-terminal sequence variations in human and ancestral SLAMF9 orthologs impact protein stability, we generated a human SLAMF9 mutant construct bearing the ancestrally conserved N-terminal sequence. HEK293T cells were transiently transfected with constructs encoding either conserved hSLAMF9 or mutant hSLAMF9 containing the ancestral N-terminus, each driven by a CMV promoter. When compared with native human SLAMF9 sequence, restoring the N-terminus to an ancestral state was sufficient to improve surface expression of SLAMF9

(Fig. 4). These results indicate that N-terminal sequence variations between human and ancestral SLAMF9 variants have impacted SLAMF9 surface expression in humans.

**Discussion**

Comparison of hominin transcripts of SLAMF9 revealed that the human SLAMF9 N-terminus emerged after the divergence of the *Homo* genus from *Pan troglodytes*. As a



**Fig. 4** Human SLAMF9 expression increases with ancestral N-terminus engineered mutant. Transiently transfected 293T cells were stained for FLAG-tagged SLAMF9. Transfection with a construct encoding the engineered mutant SLAMF9 expressing the frame-shifted ancestral

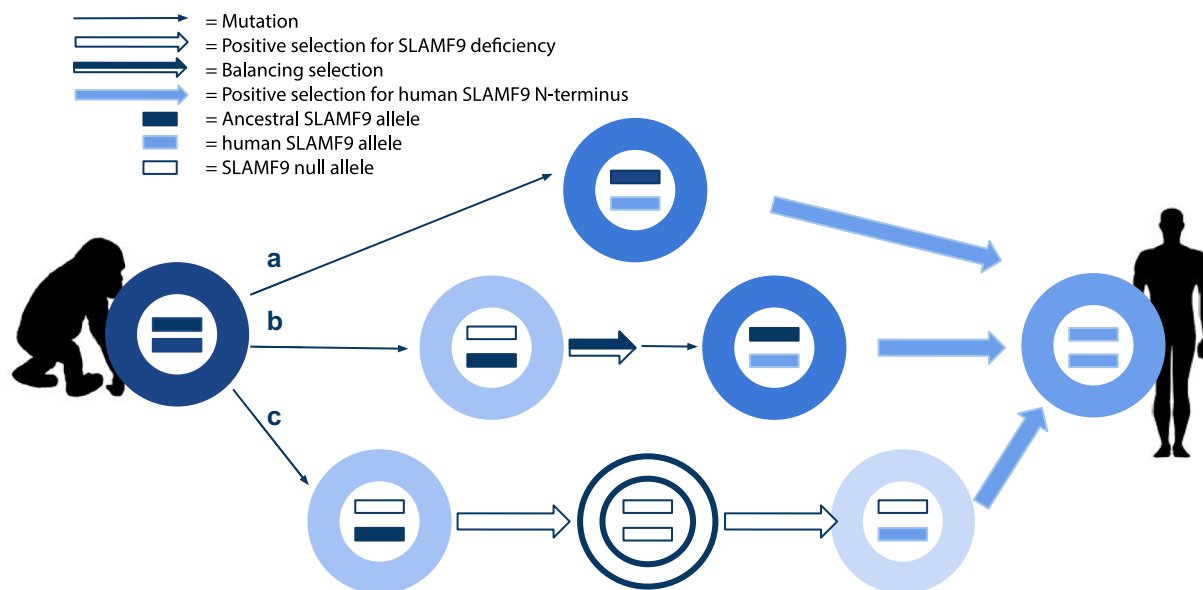
N-terminus (SLAMF9 Anc. NT) results in higher surface SLAMF9 expression compared with the consensus human SLAMF9 expressed in the same vector

result, wild-type human SLAMF9 exhibited lower protein expression levels in transfected cells than did an engineered mutant hSLAMF9 expressing the *Pan troglodytes* N-terminus. While the molecular basis of the diminished expression of human SLAMF9 is not completely clear, there are multiple, potentially destabilizing changes in the human sequence relative to other mammalian homologs. For example, the N-terminus of the mature human protein is N-terminal with respect to the A-strand of the IgSF domain, and therefore is expected to be solvent exposed. As a result of the frameshift mutation, this region now contains two non-polar tryptophan residues and a new cysteine. Together, these have the potential to lead to protein aggregation or formation of improper disulfide bonds. In addition, alterations in the human SLAMF9 signal peptide sequence invite the possibility that either inefficient insertion of SLAMF9 into the ER lumen or signal peptide sequence cleavage regulate the surface expression of SLAMF9. The latter possibility is supported by analysis of cleavage probability, which revealed diminished probability of cleavage of the human N-terminal signal peptide. Although the explanation of signal peptide inefficiency is not mutually exclusive with the potential of insertions downstream of the signal peptide to cause protein misfolding, the inability to express human

SLAMF9 precludes experimental determination of folding parameters that would be necessary to experimentally demonstrate inefficient or improper protein folding.

The in-frame translation of the human SLAMF9 corresponding to the altered N-terminal sequence results from two mutations that each shift the open reading frame of the second exon: a single-nucleotide deletion and the usage of an upstream alternative splice acceptor site. Interestingly, the four base-pair proximity of the human and ancestral SLAMF9 splice sites is closer than previously documented *de novo* splice acceptor sites (Koren et al. 2007). Because no extant species are known to encode a null SLAMF9 allele, it is unclear whether the deletion or splice acceptor site alteration first occurred. Additionally, the causative mutation(s) for the alternative splice acceptor site usage are unclear, as the ancestral AG consensus sequence remains intact in the human transcript.

As previously mentioned, the *g.621delT* variation within the second exon may be sufficient to induce the usage of an upstream cryptic splice site. In such a situation, a selective sweep favoring lower SLAMF9 expression may have favored the *g.621delT* variant (Fig. 5). While *in vitro* characterization of SLAMF9 splicing will be required to definitively confirm or refute this possibility, the positioning of the *g.621delT* variation renders it unlikely to contribute to spliceosome



**Fig. 5** Possible pathways to the evolution of human SLAMF9 from the ancestral chimpanzee sequence include the following: **a** the *g.621delT* mutation drove usage of the alternative splice acceptor site and subsequent selection for the intrinsically unstable variant; **b** an initial nonsense mutation was maintained in the population of an ancestral species by balancing selection prior to a second mutation

that generated the human SLAMF9 variant, which then underwent positive selection; or **c** an initial nonsense mutation was positively selected and SLAMF9 was lost from an ancestral species. A subsequent mutation restored SLAMF9 protein function and underwent positive selection

assembly. Positions proximal to the splice donor site, branch point, polypurine tract, and splice acceptor site tend to be the most relevant to spliceosome assembly, while the g.621*delT* variation occurs +31 positions from the splice acceptor site (Black 2003; Caceres and Hurst 2013; Vořechovský 2006). Moreover, no strong ESE consensus motifs are created or effaced by the g.621*delT* mutation (Caceres and Hurst 2013). Combined with current knowledge concerning ESEs, SLAMF9 intron 1 RNA modeling suggests that a mutation other than g.621*delT* drove the usage of the alternative splice acceptor site. Intronic sequence variations between the human and chimpanzee transcripts are concentrated near the middle of the intron, implicating disruptions of consensus motifs for the branch point or polypurine tract in cryptic splice acceptor site usage. Therefore, the most likely sequence of events is that the human SLAMF9 allele arose from a frameshift reversion in a non-functional SLAMF9 allele in a common ancestor of the *Homo* genus (Fig. 5). This event could be interpreted to represent a violation to Dollo's law in humans.

The resurrection of lost traits is considered highly improbable according to Dollo's law, which states that a trait, once lost, cannot be recovered without significant alterations (Rogozin et al. 2007). At a genetic level, however, statistical analyses suggested that a significant probability exists for the reversion of silenced genes, provided that the timespan between loss and regain is sufficiently short to preclude the accumulation of further inactivating mutations: between 0.5 million and 6 million years (Marshall et al. 1994). Since current estimates place the divergence of humans from chimpanzees as having occurred six to nine million years ago (Rogers and Gibbs 2014), the evolution of human SLAMF9 may fit just within the time range of reactivation of silenced genes prescribed by Marshall et al. A number of studies reported apparent exceptions to Dollo's rule. Potential violations to Dollo's law include the re-evolution of host independence in parasites, shell coiling in limpets, mandibular teeth in tree frogs, oviparity in viviparous reptiles, and ancestral primate muscular structures in hominins (Bochkov and Mironov 2013; Diogo and Wood 2012; Pagel 2004; Recknagel et al. 2018; Wiens 2011; Wu and Veillette 2016). Although these morphological cases of exceptions to Dollo's law have been characterized in depth, fewer individual genes subjected to "reverse evolution" have been fully characterized. Nonetheless, several examples of gene resurrection have been described. At a molecular level, the AN2 transcription factor regulating petunia pigment protein expression underwent successive mutations resulting in a loss and subsequent regain of function (Esfeld et al. 2018; Hopkins 2018). Rotating sex combs in *Drosophila* species are thought to have re-evolved due to successive chromosomal inversions (Seher et al. 2012). To date, there is a paucity of molecularly characterized violations to Dollo's law documented in humans. A rare polymorphism of the

usually silent human KIR gene *KIR3DP1* shows evidence of having developed a novel, active promoter and undergone gene conversion to encode a hydrophilic tail homologous to chimpanzee and rhesus monkey orthologs (Gómez-Lozano et al. 2005). However, the resulting protein represents more of a chimera of multiple proteins rather than the *bona fide* recovery of an ancestrally silenced gene. In contrast, our characterization of SLAMF9 suggests the occurrence of sequential point mutations that resulted in the sequential loss and regain of SLAMF9 in the last common ancestor of the *Homo* genus.

Considering the potential of SLAM family receptors to either positively or negatively regulate immune function in different contexts, SLAMF9 is uniquely poised to have been lost and regained as a consequence of sequential selective sweeps involving opposing selective pressures. The evolutionarily limited time (approximately 6 million years) during which *Homo sapiens* diverged from *Pan troglodytes* may appear to diminish the probability of two successive gene fixation events from occurring to *Slamf9*. However, analysis of evolutionary adaptations to *Plasmodium* species, to which humans have been exposed for under 200,000 years, revealed that this time interval was sufficient to nearly complete fixation of a malaria-resistant variant of *DARC* in sub-Saharan populations to which *Plasmodium vivax* is endemic (Karlsson et al. 2014). Thus, sufficiently strong selective pressures may well have driven the successive fixation of two *Slamf9* alleles in the human species. Recent studies of SLAMF9 function in mice revealed that while SLAMF9 is required for efficient clearance of *Salmonella enterica*, SLAMF9-deficient mice were less susceptible to a murine multiple sclerosis model, experimental autoimmune encephalomyelitis (Sever et al. 2019; Wilson et al. 2020). The contribution of pro-inflammatory roles of SLAMF9 in macrophages and plasmacytoid dendritic cells in response to infection indicates a clear opportunity for pathogen-driven pressure to promote the positive selection of individuals with functional SLAMF9 (Sever et al. 2019; Wilson et al. 2020; Zeng et al. 2018). However, the selective pressure that would promote the previous fixation of a null SLAMF9 allele is not immediately evident, given that prior to human speciation, susceptibility to pathogens would be more likely to reduce fitness than inflammatory disease. A potential answer may be that SLAMF9 is a receptor exploited by a specific pathogen for access to host tissue in some species. Indeed, other SLAM family receptors, including SLAM, SLAMF2 (CD48), and SLAMF6 (NTB-A), are receptors for various viral and bacterial pathogens (Cannons et al. 2011; Wu and Viellette 2016). Therefore, it is tempting to speculate that the presence of a pathogen using SLAMF9 as an entry receptor may have driven the selection of a null variant in the population, only to have receptor expression restored in later hominins under different selection conditions.

Although the data suggest the possibility of violation to Dollo's parsimony, there exists an alternative explanation to the loss and regain of SLAMF9 from an ancestor of the *Homo* genus. In this situation, balancing selection in an ancestral population may have favored the heterozygosity for a defective SLAMF9 allele bearing either the g.621delT variation or the upstream splice acceptor. A subsequent mutation to this allele that resulted in the in-frame translation of the second exon and reduced expression levels could have generated a variant that underwent strong positive selection. Ultimately, this resulted in the fixation of the novel N-terminus allele in the ancestral population (Fig. 5). This possibility would invite speculation that SLAMF9 gene dosage is critical for balancing autoimmunity and pathogen clearance. However, the lack of severe consequences of complete SLAMF9 ablation documented in literature tends to argue against the possibility of selection for low SLAMF9 dosage.

Regardless of the specific sequence of events that led to the evolution of the novel N-terminus during human speciation, selective pressures favored the fixation of either a null or intrinsically unstable SLAMF9 variant in an ancestral population. The selection against ancestral SLAMF9 suggests that high SLAMF9 expression in humans is (or was) detrimental to survival, presumably due to impaired resistance to a particular pathogen or dysregulation of the inflammatory response. However, evidence of impaired clearance of *Salmonella* infection in the liver of SLAMF9-deficient mice provides a clear rationale whereby functional SLAMF9 would confer a selective advantage (Wilson et al. 2020). Although there is insufficient evidence to definitively determine whether the evolution of human SLAMF9 constitutes an exception to Dollo's law, the data suggest that SLAMF9 expression critically impacts fitness against pathogens. These findings emphasize the need for future mechanistic studies probing a role for SLAMF9 in tunably regulating the balance between immunocompetence and autoinflammation.

**Author contribution** M.K.M.: wrote manuscript, conducted research; J.T.M.: conducted research; A.T.S.: conducted research; J.A.M.: conducted research; T.J.W.: wrote and edited manuscript, directed research.

**Funding** This work was supported in part by NIAID R15-AI138184. M.K.M. and J.T.M. were supported through a Miami University Undergraduate Summer Scholarship. M.K.M. was additionally supported by Miami University Dean's Scholarship, DUOS Award, and an HHMI Summer Scholarship.

**Availability of data and material** Data and materials not publicly available can be obtained by contacting the corresponding author.

## Declarations

**Competing interests** The authors declare that they have no competing interests.

## References

- Adrian J, Bonsignore P, Hammer S (2019) Adaptation to host-specific bacterial pathogens drives rapid evolution of a human innate immune receptor. *Curr Biol* 29:616–630
- Almagro Armenteros JJ, Tsirigos KD, Sonderby CK, Petersen TN, Winther O, Brunak S, von Heijne G, Nielsen H (2019) SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol* 37:420–423
- Barclay AN, Hatherley D (2008) The counterbalance theory for evolution and function of paired receptors. *Immunity* 29:675–678
- Barreiro LB, Quintana-Murci L (2010) From evolutionary genetics to human immunology: how selection shapes host defence genes. *Nat Rev Genet* 11:17–30
- Black DL (2003) Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem* 72:291–336
- Bochkov AV, Mironov SV (2013) Is parasitism of Metazoa “a one-way ticket”? *Parazitologiya* 47:212–227
- Caceres EF, Hurst LD (2013) The evolution, impact and properties of exonic splice enhancers. *Genome Biol* 14:R143
- Cagliani R, Pozzoli U, Forni D (2013) Crohn's disease loci are common targets of protozoa-driven selection. *Mol Biol Evol* 30:1077–1087
- Cannons JL, Tangye SG, Schwartzberg PL (2011) SLAM family receptors and SAP adaptors in immunity. *Annu Rev Immunol* 29:665–705
- Cartegni L, Wang J, Zhu Z (2003) ESEfinder: a web resource to identify exonic splicing enhancers. *Nucleic Acids Res* 31:3568–3571
- Diogo R, Wood B (2012) Violation of Dollo's law: Evidence of muscle reversions in primate phylogeny and their implications for the understanding of the ontogeny, evolution, and anatomical variations of modern humans. *Evolution* 66:3267–3276
- Dollt C, Michel J, Kloss L (2018) The novel immunoglobulin super family receptor SLAMF9 identified in TAM of murine and human melanoma influences pro-inflammatory cytokine secretion and migration. *Cell Death Dis* 9
- Domínguez-Andrés J, Netea MG (2019) Impact of historic migrations and evolutionary processes on human immunity. *Trends Immunol* 40:1105–1119
- Dragovich MA, Mor A (2018) The SLAM family receptors: potential therapeutic targets for inflammatory and autoimmune diseases. *Autoimmun Rev* 17:674–682
- Esfeld K, Berardi AE, Moser M (2018) Pseudogenization and resurrection of a speciation gene. *Curr Biol* 28:3776–3786
- Fennelly JA, Tiwari B, Davis SJ, Evans EJ (2001) CD2F-10: a new member of the CD2 subset of the immunoglobulin superfamily. *Immunogenetics* 53:599–602
- Fumagalli M, Sironi M, Pozzoli U (2011) Signatures of environmental genetic adaptation pinpoint pathogens as the main selective pressure through human evolution. *PLoS Genet* 7
- Gómez-Lozano N, Estefania E, Williams F (2005) The silent KIR3DP1 gene (CD158c) is transcribed and might encode a secreted receptor in a minority of humans, in whom the KIR3DP1, KIR2DL4 and KIR3DL1/KIR3DS1 genes are duplicated. *Eur J Immunol* 35:16–24
- Hopkins R (2018) Evolution: flip-flopping flower color defies Dollo's Law. *Curr Biol* 28:R1337–R1339
- Jackson MP, Hewitt EW (2016) Cellular proteostasis: degradation of misfolded proteins by lysosomes. *Essays Biochem* 60:173–180
- Jostins L, Ripke S, Weersma RK (2012) Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* 491:119–124
- Karlsson EK, Kwiatkowski DP, Sabeti PC (2014) Natural selection and infectious disease in human populations. *Nat Rev Genet* 15:379–393
- Koren E, Lev-Maor G, Ast G (2007) The Emergence of Alternative 3' and 5' Splice Site Exons from Constitutive Exons. *PLoS Comput Biol* 3



- Marshall CR, Raff EC, Raff RA (1994) Dollo's law and the death and resurrection of genes (development/gene inactivation/functional constraint/evolutionary reversal/molecular evolution. *Proc Natl Acad Sci USA* 91:12283–12287
- Pagel M (2004) Limpets break Dollo's law. *Trends Ecol Evol* 19:278–280
- Peng G, Yang Y, Pasquarella J (2017) Structural and molecular evidence suggesting coronavirus-driven evolution of mouse receptor. *J Biol Chem* 292:2174–2181
- Recknagel H, Kamenos NA, Elmer KR (2018) Common lizards break Dollo's law of irreversibility: genome-wide phylogenomics support a single origin of viviparity and re-evolution of oviparity. *Mol Phylogenet Evol* 127:579–588
- Rogers J, Gibbs RA (2014) Comparative primate genomics: emerging patterns of genome content and dynamics. *Nat Rev Genet* 15:347–359
- Rogozin IB, Wolf YI, Babenko VN, V KE (2007) Dollo parsimony and the reconstruction of genome evolution. *Parsimony, Phylogeny, and Genomics*.
- Schwartz JC, Sanderson ND, Bickhart DM (2019) The Structure, Evolution, and Gene Expression Within the Caprine Leukocyte Receptor Complex. *Front Immunol* 10
- Seher TD, Ng CS, Signor SA (2012) Genetic basis of a violation of Dollo's law: re-evolution of rotating sex combs in *Drosophila bipectinata*. *Genetics* 192:1465–1475
- Sever L, Radomir L, Strim K (2019) SLAMF9 regulates pDC homeostasis and function in health and disease. *Proc Natl Acad Sci U S A* 116:16489–16496
- Smith PJ, Zhang C, Wang J (2006) An increased specificity score matrix for the prediction of SF2/ASF-specific exonic splicing enhancers. *Hum Mol Genet* 15:2490–2508
- Volkova O, Gusel'nikov S, Mechetina L, Chikaev N, Baranov K, Kulemzin S, Reshetnikova E, Najakshin A, Taranin A (2014) Development and characterization of domain-specific monoclonal antibodies produced against human SLAMF9. *Monoclon Antib Immunodiagn Immunother* 33:209–214
- Vořechovský I (2006) Aberrant 3' splice sites in human disease genes: mutation pattern, nucleotide structure and comparison of computational tools that predict their utilization. *Nucleic Acids Res* 34:4630–4641
- Wiens JJ (2011) Re-evolution of lost mandibular teeth in frogs after more than 200 million years, and re-evaluating Dollo's law. *Evolution* 65:1283–1296
- Wilks J, Lien E, Jacobson AN (2015) Mammalian lipopolysaccharide receptors incorporated into the retroviral envelope augment virus transmission. *Cell Host Microbe* 18:456–462
- Wilson MD, Cheung J, Martindale DW (2006) Comparative analysis of the paired immunoglobulin-like receptor (PILR) locus in six mammalian genomes: duplication, conversion, and the birth of new genes. *Physiol Genomics* 27:201–218
- Wilson TJ, Clare S, Mikulin J (2020) Signalling lymphocyte activation molecule family member 9 is found on select subsets of antigen-presenting cells and promotes resistance to *Salmonella* infection. *Immunology* 159:393–403
- Wu N, Veillette A (2016) SLAM family receptors in normal immunity and immune pathologies. *Curr Opin Immunol* 38:45–51
- Xue Y, Daly A, Yngvadottir B, Liu M, Coop G, Kim Y, Sabeti P, Chen Y, Stalker J, Huckle E, Burton J, Leonard S, Rogers J, Tyler-Smith C (2006) Spread of an inactive form of caspase-12 in humans is due to recent positive selection. *Am J Hum Genet* 78:659–670
- Zeng X, Liu G, Peng W (2018) Combined deficiency of SLAMF8 and SLAMF9 prevents endotoxin-induced liver inflammation by downregulating TLR4 expression on macrophages. *Cell Mol Immunol* 17:153–162
- Zhang W, Wan T, Li N, Yuan Z, He L, Zhu X, Yu M, Cao X (2001) Genetic approach to insight into the immunobiology of human dendritic cells and identification of CD84-H1, a novel CD84 homologue. *Clin Cancer Res* 7:822s–829s
- Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 31:3406–3415

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.