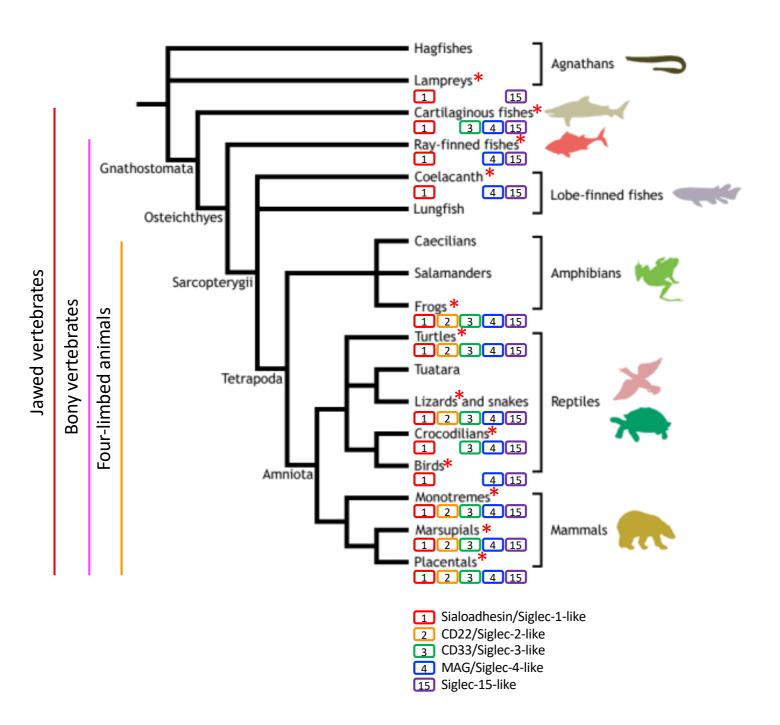
*	$\mathbf{\nabla}$
HsSig15 -MEKSIWLLAC-LAWVLPTCSFVRTKIDTTENLLNTEVHSSPAQRWSMQVPPEV	SAEAGDAAVLPCTFTHPHRHYDG 75
HsSig2 MHLLGPW HIVEY A	YAWEGACVWIPCTYRALDGDL-ESFI 57
HSSig4MIFLTALPIFWI-MISASRGGHWGAWMPSSI HSSig1MGFLPKILLLASFFP-A	SAFEGTCVSIPCRFDFPDELRPA 53
HsSig1MGFLPKIIIIIASFFP-AGQASWGVSSPQDV	OGVKGSCLLIPCIFSFPADVEVPDG- 54
HSSIgI0MLEPHILSSILG-GSQAMDGREWIRVQESV	MVPEGLCISVPCSFSYPRQDWTGS-T 54
	PVPEGLCVIVSCNLSIPRDGWDES-T 55
	TWOFCLOVINDCSESYDWRSW_VSSP 54
HsSig14MEPLITEPLIWG-GSISIOEKPVVELOVOKSV	TWOEGLEVILVPCSFSYPWRSW-YSSP 54
HsSig3MPLILLLPLLWA-GALALAMDPNEWLOVQESV	TVOEGLCVLVPCTFFHPIPYY-DKNS 54
HsSig6AUPHILLPLEWA-GALALAQERRFOLEGPESL	TVQEGLCVLVPCRLPTTLPASY 50
HsSig8MLLLLLLPLLWG-TKGMEGDRQYGDGYLLQVQELV	TVQEGLCVHVPCSFSYPQDGWTDSD- 60
HsSig1MGFLPKLLLASFFP-A	TVQEGMCVHVRCSFSYPVDSQTDSD- 59
HsSig9MUILLPINWG-RERAEGQTSKLITMQSSV	TVQEGLCVHVPCSFSYPSHGWIYPGP 55
HSSig15 PITAIWRAGEPYAGPOWFRCAAARGSELCQTALSLHGRERILGNPRRNDLSL	
HSSIGIS PETATMRAGEPIAGPONERCAAARGSELCGIALSLIGRIRULGNPRRNELSL HSSig2 LFHNPEYNKNTSKFDGT-RLYESTKDGKVPSEQKRVQFLGDK-NKNCTL	RVERIALADDRRTCRVEFAGDVHDR 155
HsSig1 -ITAIWYYDYSGOROVVSHSADPKLVEARFRGRTEFMGNPEHRVCNL	LLKDLOPEDSGSYNFRFEISEVNRWS 126
HsSig10 PAYGYWFKAVTETTKGAPVATNHQSREVEMSTRGRFQLTCDPAKGNCSL	VIRDAQMQDESQYFFRVERGSYVRYN 129
HsSig11 AAYGYWFKGRTSPKTGAPVATNNQSREVEMSTRDRFQLTGDPGKGSCSL	VIRDAQRE <mark>DEAWYFFR</mark> VERGSRVRHS 130
HsSig16 AAYGYWFKGRTSPKTGAPVATNNQSREVAMSTRDRFQLTGDPGKGSCSL	VIRDAQREDEAWYFFRVERGSRVRHS 130
HsSig5 PLYVYWFRDGEIPYYAEVVATNNPDRRVKPETQGRFRLLGDVQKKNCSL	SIGDARMEDTGSYFFRVERGRDVKYS 129
HSSig4 VVHEVWFNSPYPKNYPPVVFKSRTQVVHESFQGRSRLLGDLGLRNCTL HSSig1 - TTAIWYYDYSGQRQVVSHSADPKLVEARFQGREFQFGNPEHRVCNL HSSig10 PAYGYWFKAVTETTKGAPVATNHQSREVEMSTRGRFQLTGDPAKGNCSL HSSig11 AAYGYWFKGRTSPKTGAPVATNNQSREVEMSTRDRFQLTGDPGKGSCSL HSSig16 AAYGYWFKGRTSPKTGAPVATNNQSREVEMSTRDRFQLTGDPGKGSCSL HSSig5 PLYVYWFRDGEIPYYAEVVATNNPDRRVKPETQGRFRLLGDVQKKNCSL HSSig14 PLYVYWFRDGEIPYYAEVVATNNPDRRVKPETQGRFRLLGDVQKKNCSL HSSig3 PVHCYWFRGAISGDSPVATNNPDRRVKPETQGRFRLLGDVQKKNCSL HSSig4 PLYVWFRDGEIPYYAEVVATNNPDRRVKPETQGRFRLLGDVQKKNCSL	SHGDARMEDTGSYFFRVERGRDVKYS 129
HSSIGS PVHGYWRELGAIISGDSPVAINRLDQEVQEEIRQGRABDGDPSRNNCSL HSSig6 YGYGYWFLEGADVPVAINDPDEEVQEEIRGREHLLWDPRRKNCSL	STRDARRRDNGSITTRMERGS-IKIS 120
HSSig6 YCYCYWFLEGADVPVATNDPDEEVQEETRGRHLLWDPRKNCSI HSSig8 PVHCYWFRAGDRPYQDAPVATNNPDREVQAETQGRFOLLGDIWSNDCSI HSSig7 PVHCYWFRAGNDISWKAPVATNNPAWAVQEETRDRFHLLGDPQTKNCTI	STRDARKRDKGSYFFRLERGS-MKWS 134
HsSig7 PVHGYWFRAGNDISWKAPVATNNPAWAVQEETRDRFHLLGDPQTKNCTL	SIRDARMSDAGRYFFRMEKGN-IKWN 133
HsSig9 VVHGYWFREGANTDQDAPVATNNPARAVWEETRDRFHLLGDPHTKNCTL	SIRDARRSDAGRYFFRMEK <mark>G</mark> S-IKWN 129
· · · · · · · · · · · · · · · · · · ·	
HsSig15 YESRHGVRLHVTAAPRIVNISVLPSPAHAFRALCTÄEGEPPPA	LAWSGPALGNSLAAVRSPREG 217
HsSig2 ERIHINVSERPFPPHIQLPPEIQESQEVTITCLLNISCYGYPI-Q	LQWLLEGVPMRQAAVTSTSLTIKSVF 199
HSS1g4 FSEHSVLD VNTPNIVVPPEVVAGTEVENSCMVPDNCPELRP-E	LSWLCHECLGEPAVLGRLREDEGTWV 197
HSSIGI DVRGIDVIVILEPRVPITASPVELLEGIEVDFNOSIFNVELQEQV-R	FSWTGAALSSOCTK-PTTS 195
HSSig11 FLSNAFFLKVTALTKKPDVYIPETLEPGOPVTVICVFNWAFKKCPAPS	FSWTGAALSPRRTR-PSTS 196
HsSig16 FLSNAFFLKVTALTQKPDVYIPETLEPCQPVTVICVFNWAFKKCPAPS	FSWTGAALSPRRTR-PSTS 196
HsSig5 YQQNKLNLEVTALIEKPDIHFLEPLESCRPTRLSCSLPGSCEAGPPLT	F <mark>SW</mark> T <mark>G</mark> NALSPLDPETT 193
HsSig14 YQQNKLNLEVTALIEKPDIHFLEPLESGRPTRLSCSLPGSCEAGPPLT	FSWTGNALSPLDPETT 193
HsSig3 YKSPQLSVHVTDLTHRPKILIPGTLEPGHSKNUTCSVSWACEQGTPPI	FSWLSAAPTSLGPRTT 192
HSSig6 YTSSKLSVRVMALTHRPNISIPGTLESCHPSNLTCSVPWVCEQGTPPI	FSWMSAAPTSLGPRIT 184
HSSig15 YESRHGVRLHVTAAPRIVNISVLPSPAHAFRALCTAEGEPPA HSSig2 ERHGVRLHVTAAPRIVNISVLPSPAHAFRALCTAEGEPPA HSSig2 FSEHSVLDIVNTPNIVVPPEVAGTEVEVSCMVPDNCPELRP-E HSSig1 DV	ITSWIGISVSPLDPSTT 197
HsSig15 HGHLVTAELPALTHDGRYTCTAANSLGRSEASVYLFRFHCASGAST 263	
HsSig2 TRSELKFSPQWSHHGKIVTCQLQDADGKFLSNDTV-QLNVK 239	
HSSig4 QVSLLHFVPTREANGHRLGCQASFPNTTLQFFGYA-SMDVK 237 HSSig1 HLETLHMAMSWODHGRILRCQLSVANHRAQSFIHL-QVK 235	 — Signal peptide
HSSig1 HLETLHMAMSWODHGRI <mark>LRCO</mark> ISVANHRAQSEIHL-QVK 235 HSSig10 HF <mark>SVLSFTPRPQDHNTDLTCHVDFSRKGVSAQRTV-RLRV</mark> A 235	
HSSig11 HFSVLSFTPSPQDHDTDLTCHVDFSRKGVSAQRTV-RLRVA 236	— lg1
HSSig16 HFSVLSFTPSPQDHDTDLTCHVDFSRKGVSAQRTV-RURVA 236	•
HsSig5 RSSELTLTPRPEDHGTNLTCOMKRQGAQVTTERTV-QLNVS 233	— lg2
HsSig14 RS <mark>SELT</mark> LTPRPEDHGTNLTCQVKRQGAQVTTERTV-QLNVS 233	* Eccontial Ara
HSSig5 RSSELTLTBRPEDHGTNLTCOMKRQGAQVTTBRTV-QLNVS 233 HSSig14 RSSELTLTBRPEDHGTNLTCOVKRQGAQVTTBRTV-QLNVS 233 HSSig3 HSSVLIITBRPODHGTNLTCOVKFAGAGVTTBRTT-QLNVT 232	 Essential Arg
HSSIG6 QSSVLTITPRPODHSTALTCOVTFPGAGVTMERTI-QLAVS 224	 Aromatic amino acid
HSSig8 RSSVLTLTPKPODHGTSLTCOVTLPGTGVTTTSTV-RLDVS 244	
HSSig7 RSSVLTLIPOPOHHGTSLTCOVTLPGAGVTTNRTI-QLNVS 237 HSSig9 RSSVLTLIPOPOHGTSLTCOVTFPGASVTTNKTV-HLNVS 233	interacting with Sia
WORLAN WOMMENTERSENTONOMONDATICGODMITERSENTERSENTONO	•
	Cys involved in inter-
	•
	domain disulfide bond

Supplementary Figure 1. Alignment of human Siglecs.

Amino acid sequences of the N-terminal part (signal peptide – Ig1 – Ig2) of human Siglecs were aligned with ClustalO, and conserved residues (>70% consensus) are highlighted. Black shade: same amino acid; Gray shade: similar amino acids.



Supplementary Figure 2. Tree of vertebrate life and presence of Siglecs.

The phylogenetic tree of vertebrates is based on Meyer and Zardoya (2003) *Annu. Rev. Ecol. Evol. Syst.* 34:311–338, and was downloaded from University College London website (https://www.ucl.ac.uk/museums-static/obl4he/vertebratediversity/index.html). One or more of the species belonging to the clade indicated with red asterisk (*) were subjected to genomic DNA sequence survey for Siglec V-set (Ig1) domain-like sequence and mentioned in the main text.

Licensed under the Creative Commons Attribution Non-commercial Share Alike 3.0 License (CC BY-NC-SA 3.0).

Α

HsSig1 PmSig1L	* MGFLPKLLLLASFFPAGQASWGVSSPQDVQGVKGSCLLIPCIFSFPADVEVPDGITAIWYYDYSGQRQV MHSKIQDLAVRSWIIVSCFFISLHSSQAESEWGVTYGQEVIASPGSSVILPCTFKYPEKGKLISGPVWMKDSDYSISNIV	
HsSig1 PmSig1L	* VSHSADPKLVEARFRGRTEFMGNPEHRVCNLLLKDLQPEDSGSYNFRFEISEVNRWSDVKGTLVTVTEEPRVPT-IASPV YSTKADEIVEEYRGRTSLVGDLDANNCSLKIIHAEKSDGRKYFFRFMTHDSYTGIAGIELKVQDESEWGVMYGQEV	
HsSigl PmSiglL	ELLEGTEVDFNCSTPYVCLQEQVRLQWQGQDPARSVTFNSQKFEPTGVGHLETLHMAMSWQDHGRILRCQLSVANHRAQS KTWPGGSAILPCSFKYIDEGKRVTWAGWMKNKNFPYDTIVYSTSGEVMQEYS-GRTSLVGDLYARDCSLKIINTSKDD	
HsSig1 PmSig1L	EIHLQVK 235 GSRYFFRFTTKSHWTGNVGIQLTVH 258	

В

* HsSig15MEKSIWLLACLAWVLPTGSFVRTKIDTTENLLNTEVHSSPAQRWSMQVPPEVSAEAGDAAVLPCTFTHPHRHY PmSig15L MRLSPMDKMIIIQLLFTITFMLTLFGLAASEFDCQNHGRWSICVPKSVTGLKTKEATLLCNFTYPRGARNM	73 71
* HsSig15 DGPLTAIWRAGEPYAGPQVFRCAAARGSELCQTALSLHGRFRLLGNPRRND-LSLRVERLALADDRRYFCRVEFAGDVHD PmSig15L DTKIQVIWKKNILFEKNASYYIFNRATNTTRFDFNGRVSLVGDPENHSTASIQIRELKYSDSDNYYCRFEIEK	
	226 224
HSSig15 PALTHDGRYTCTAANSLGRSEASVYLFRFHGASGASTVALLLGALGFKALLLLGVLAARAARRRP 291 PmSig15L CGELPTGAYTCRAINQHGEDSTSHDVEPPPTTSTAILIGASAATFFLCTLFAFITFWHCRIQ- 286	

- Signal peptide
- lg1
- lg2
- Essential Arg
- Aromatic amino acid interacting with Sia
- Cys involved in interdomain disulfide bond

Supplementary Figure 3. Alignment of putative lamprey Siglecs with human homologs.

Amino acid sequences of the N-terminal part (signal peptide – Ig1 – Ig2) of (A) human Siglec-1 (HsSig1) and lamprey homolog (PmSig1L), and (B) human Siglec-15 (HsSig15) and lamprey homolog (PmSig15L) were aligned with ClustalO, and conserved residues are highlighted. Black shade: same amino acid; Gray shade: similar amino acids. GenBank accession number for Lamprey Siglec-1-like sequence: XM_032693083; Siglec-15-like sequence: XM_032963052.

HsSiglec15 HsHAVCR2 HsHAVCR1 HsTIMD4	MEKSIWLLACLAWVLPTGSFVRTKIDTTENLLNTEVHSSPAQRWSMQVPPEVSAEAGDAAVLPCTFTHPHRHYDGP MFSHLPFDCVLLLLLLLLTRSSEVEYRAEVGQNAYLPCFYTPAAPGNLVP -MHPQVVILSLILHLADSVAGSVKVGGEAGPSVTLPCHYSGAVTS -MSKEPLILWLMIEFWWLYLT	50 44
HsSiglec15 HsHAVCR2 HsHAVCR1 HsTIMD4	LTAIWRAGEPYAGPQVFRCAAARG-SELCQTALSLHGRFRLLGNPRRNDLSLRVERLALADDRRYFCRVEFAGDVHDRYE VCWGKGACPVFECGNVVLRTDERDVNYWT-SRYWLNGDFRKGDVSLTIENVTLADSGIYCCRIQIPGIMNDEK- MCWNRGSCSLFTCQNGIVWTNGTHVTYRKDTRYKLLGDLSRRDVSLTIENTAVSDSGVYCCRVEHRGWFNDMK- MCWGKDQCPYSGCKEALIRTDGMRVTSRKSAKYRLQGTIPRGDVSLTILNPSESDSGVYCCRIEVPGWFNDVK-	122 117
HsSiglec15	SRHGVRLHVT- 165	

HSSIGLECIS	SRHGVRLHVT-	102
HsHAVCR2	FNLKLVIKP	131
HsHAVCR1	ITVSLEIVP	126
HsTIMD4	INVRLNLQR	133

Supplementary Figure 4. Alignment of human Siglec-15 and related IgSF proteins.

Amino acid sequences of the signal peptide – Ig1 of human Siglec-15, TIM-1 (*HAVCR1*), TIM-3 (*HAVCR2*), TIM-4 (*TIMD4*) were aligned with ClustalO, and conserved residues (>70% consensus) are highlighted.

Black shade: same amino acid; Gray shade: similar amino acids.