

Supporting Information

Li et al. 10.1073/pnas.1314540110

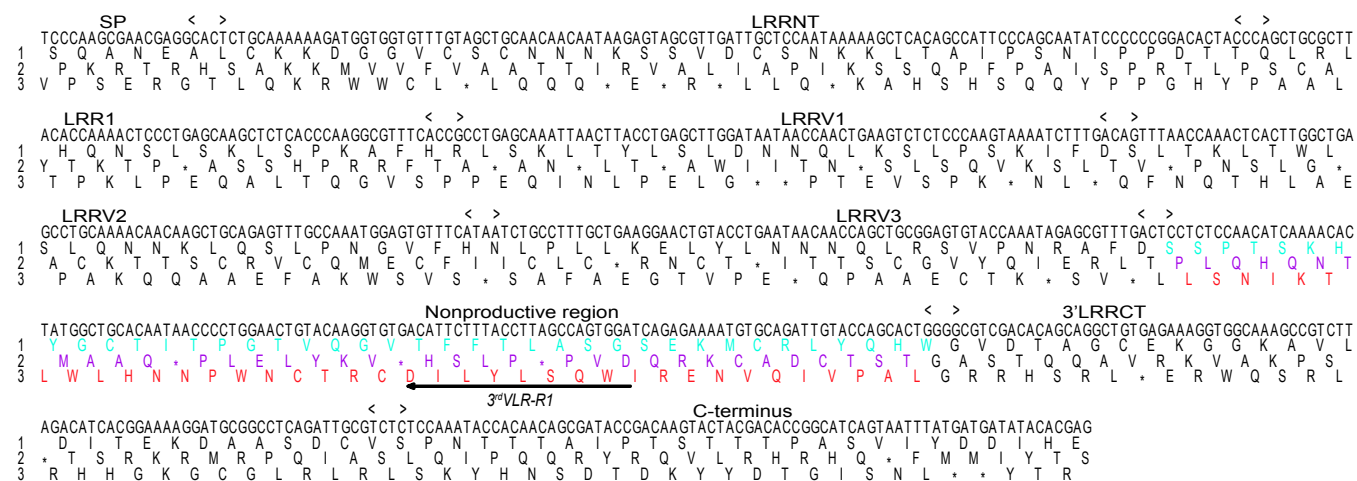


Fig. S1. Hypothetical translation of the hagfish EST clone (GenBank accession no. BAI66885) in three open reading frames (ORF). ORF1 of this clone encoded a nonproductive VLR-like protein lacking a connecting peptide (CP) and intact leucine-rich repeat C-terminal (LRRCT) motif. ORF1 (cyan) and ORF2 (purple) in the nonproductive region do not encode a characteristic CP-5'LRRCT motif, whereas the corresponding region in ORF3 (red) codes for a unique CP-5'LRRCT module. Arrowed line indicates the reverse primer region used for 5'-RACE.

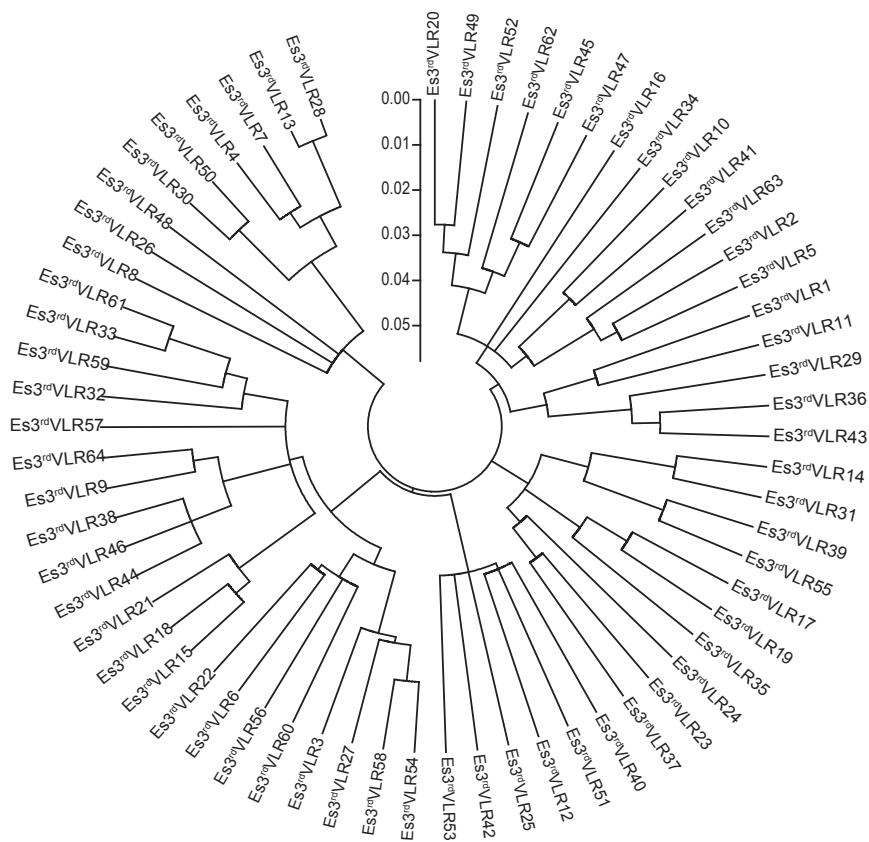


Fig. S2. Diversity of hagfish 3rdVLRs. This neighbor-joining tree was constructed using the diversity region from leucine-rich repeat N-terminal (LRRNT) to LRRCT of 64 mature sequences (GenBank accession nos. KF314046–KF314109).

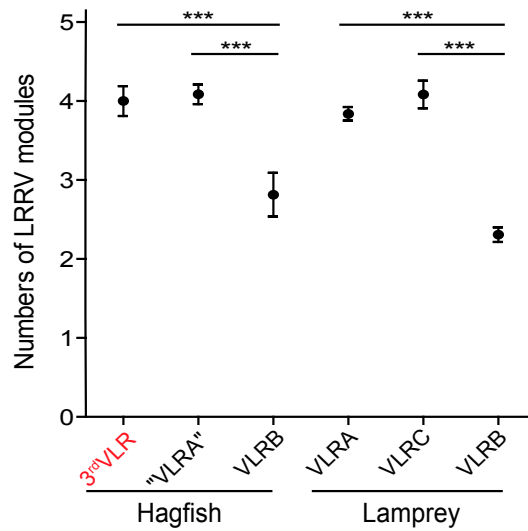


Fig. S3. Numbers of variable leucine-rich repeat (LRRV) modules in the different VLRs. Black circles represent average number of LRRV modules in the indicated VLRs. Bars indicate 95% confidence interval. cDNA clones analyzed: hagfish 3rdVLR ($n = 64$), hagfish "VLRA" ($n = 139$) (1), hagfish VLRB ($n = 70$) (1), lamprey VLRA ($n = 192$) (2), lamprey VLRC ($n = 60$) (3) and lamprey VLRB ($n = 517$) (4).

1. Pancer Z, et al. (2005) Variable lymphocyte receptors in hagfish. *Proc Natl Acad Sci USA* 102(26):9224–9229.
2. Rogozin IB, et al. (2007) Evolution and diversification of lamprey antigen receptors: Evidence for involvement of an AID-APOBEC family cytosine deaminase. *Nat Immunol* 8(6):647–656.
3. Das S, et al. (2013) Organization of lamprey variable lymphocyte receptor C locus and repertoire development. *Proc Natl Acad Sci USA* 110(15):6043–6048.
4. Alder MN, et al. (2005) Diversity and function of adaptive immune receptors in a jawless vertebrate. *Science* 310(5756):1970–1973.

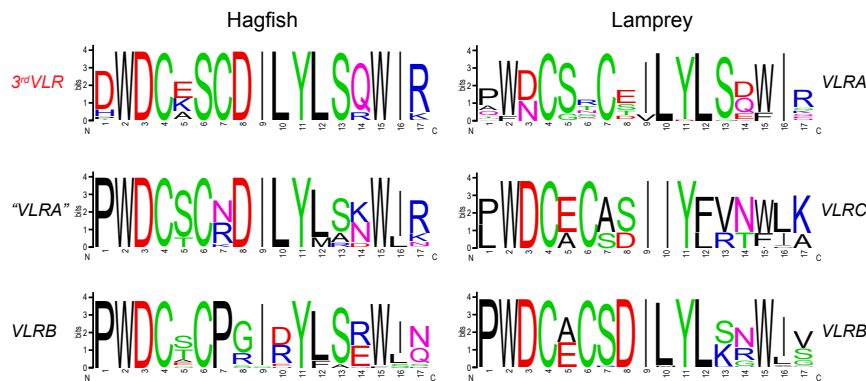


Fig. S4. Sequence signatures for the LRRCT portion of all known VLRs. Consensus sequences were generated by using the 17 amino acid residues from the N-terminal of LRRCT.

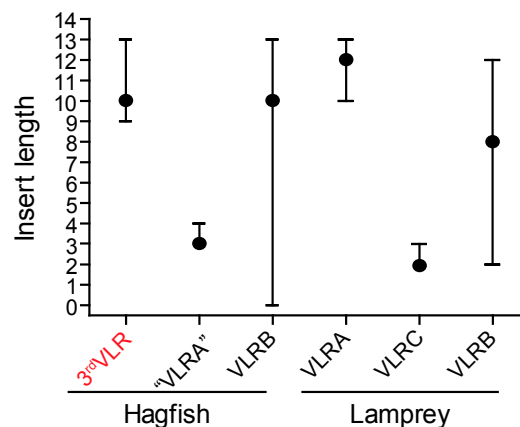


Fig. S5. Variable lengths of the LRRCT insert regions. Black circles represent median length of the indicated VLRs. Lines indicate range of insert length.

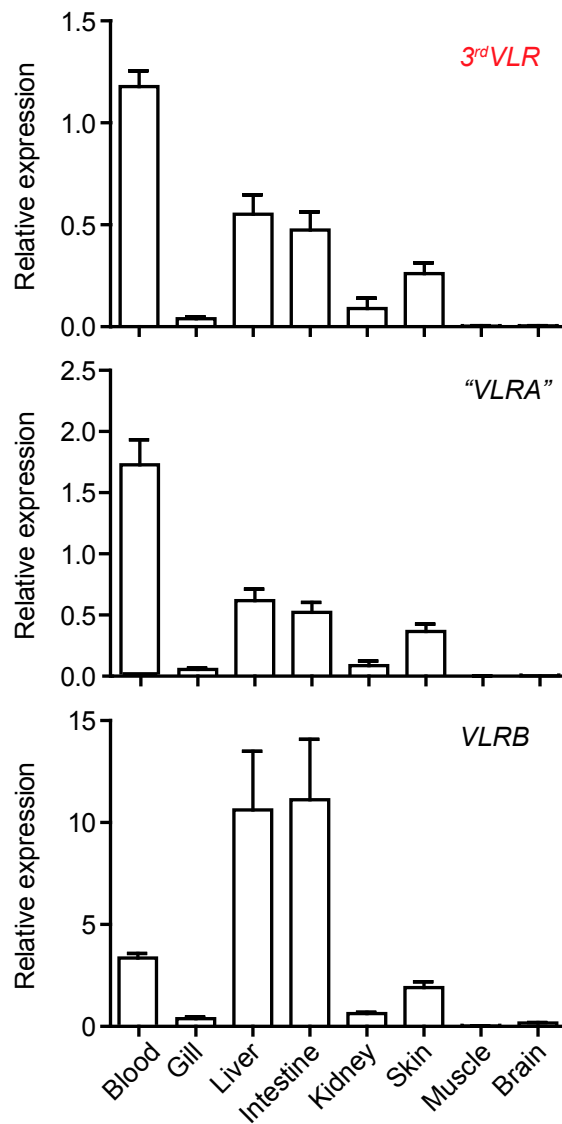


Fig. S6. Analysis of 3^{rd}VLR , "VLRA", and VLRB expression by cells in different tissues. Transcripts were analyzed by real-time RT-PCR with GAPDH as control. Bars indicate SE of mean (SEM); $n = 3$.

Table S1. PCR primers

Primer name	Primer sequence (5'→3')
$3^{\text{rd}}\text{VLR-R1}$	TCCACTGGCTAAGGTAAGAATGTC
$3^{\text{rd}}\text{VLR-SF}$	ATGGAACGATGGTTCAGCAG
$3^{\text{rd}}\text{VLRstop-R}$	TTATGGTTTTCGGAACAACACAC
QEs"VLRA" F	GATACCAAGAGCCAGACCATC
QEs"VLRA" R	GTGGATGACGTACAAAATGGC
QEsVLRBF	TACTTTCCCTTCGTACATCTTCCTC
QEsVLRBR	AGGCGTGACATACGAGGTAA
QEs 3^{rd}VLR F	GAGCTTTTCTCAAAGGCATCG
QEs 3^{rd}VLR R	GGAATCTCCGTAAAAATGTGAATC
QEsGAPDHF	GATCACAGTATTCCAGGAGCG
QEsGAPDHR	AGAGATGACTACCCGTTTTTC