nature ecology & evolution

Supplementary information

https://doi.org/10.1038/s41559-025-02845-x

Evolutionary characterization of antiviral SAMD9/9L across kingdoms supports ancient convergence and lineage-specific adaptations

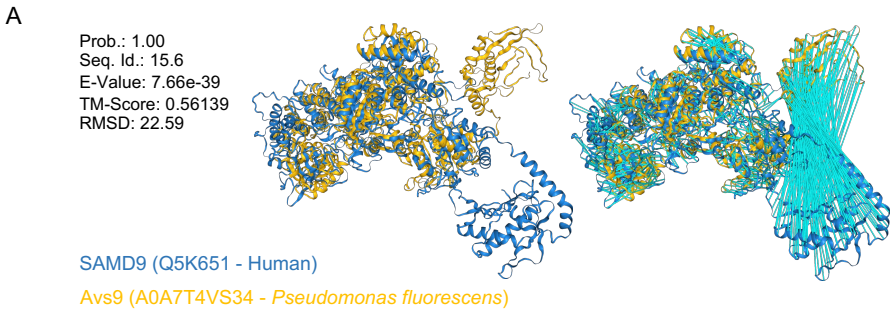
In the format provided by the authors and unedited

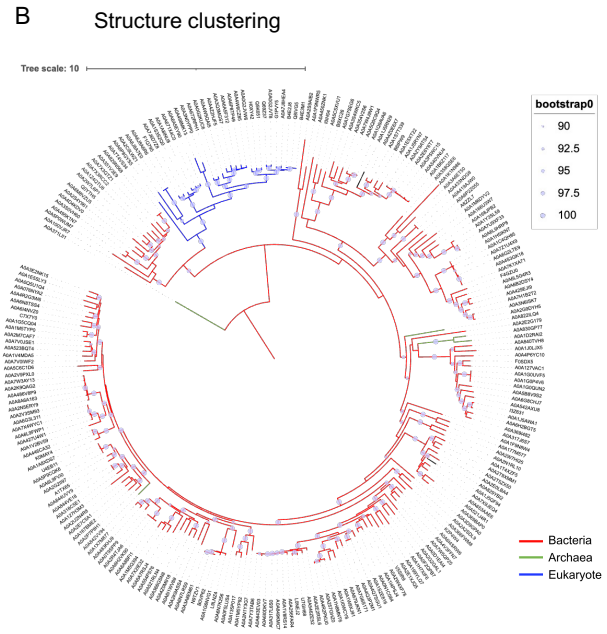
Supplementary Information

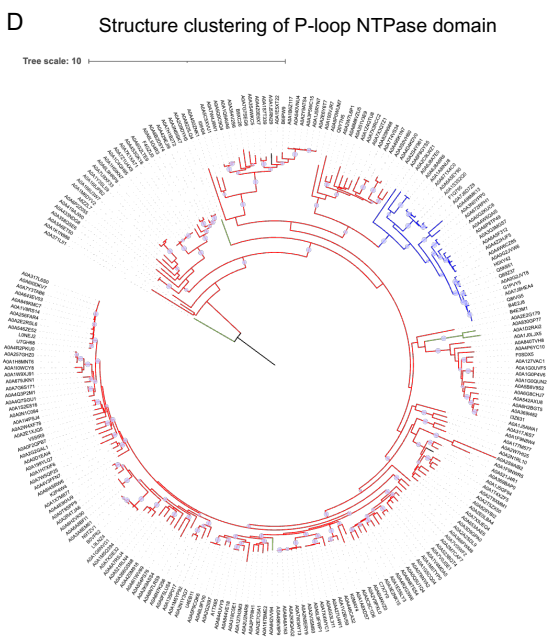
Table of contents

Supplementary Figures

<u>Supplementary Figure 1. Associated to Figure 1. SAMD9/9L structural analogs are extensively present in prokaryotes.</u> A, Left. Protein overlap (FoldSeek) between human SAMD9 and *Pseudomonas fluorescens* A0A7T4VS34 (Avs9). Right. Similar with blue bars linking aligned residues. **B,** Structure clustering of full-length proteins shows widespread and diverse SAMD9/9L analogs in bacteria. **C,** SAMD9/9L domain bounds (coordinates of the protein sequence) used for the biocomputational analyses. **D,** Structure clustering of P-loop NTPase extraction.





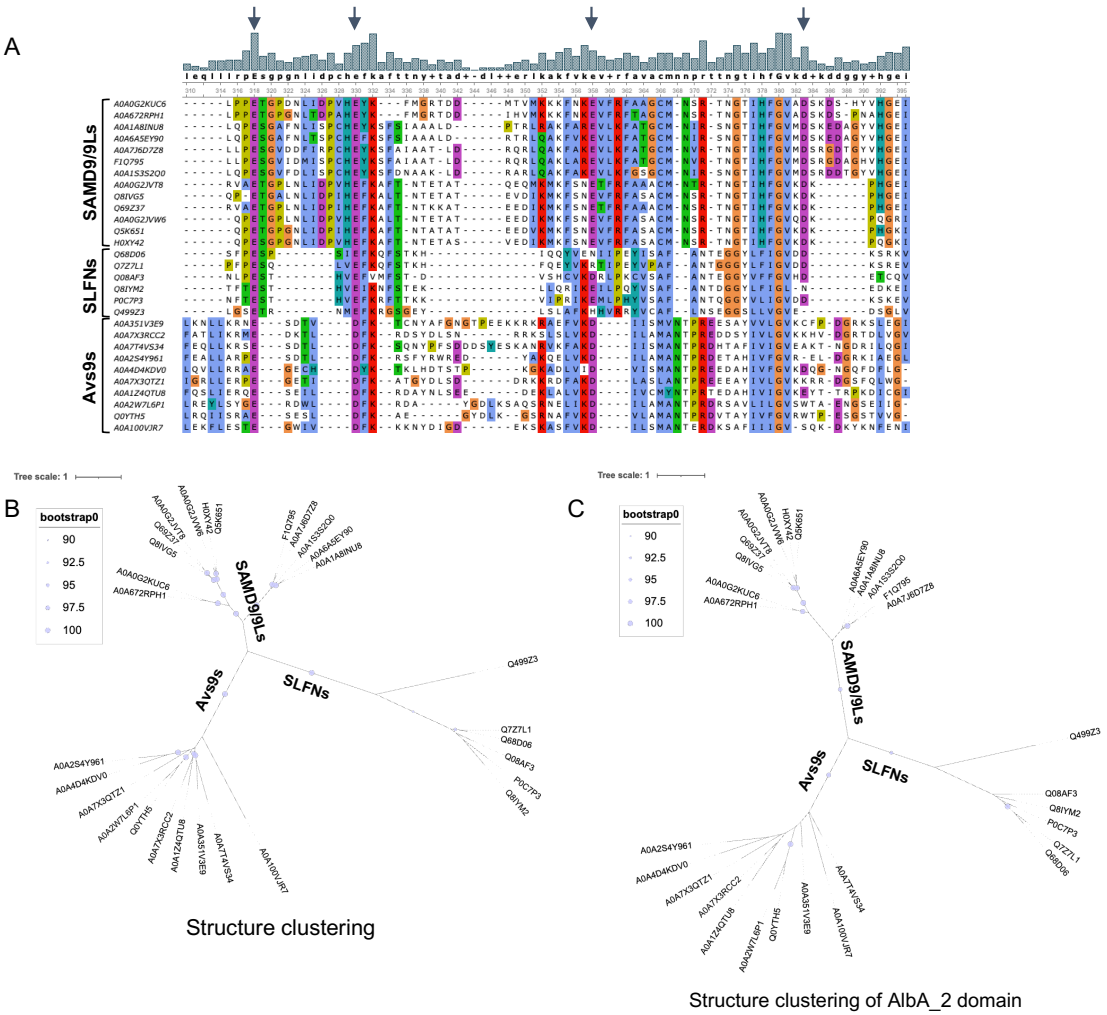


C

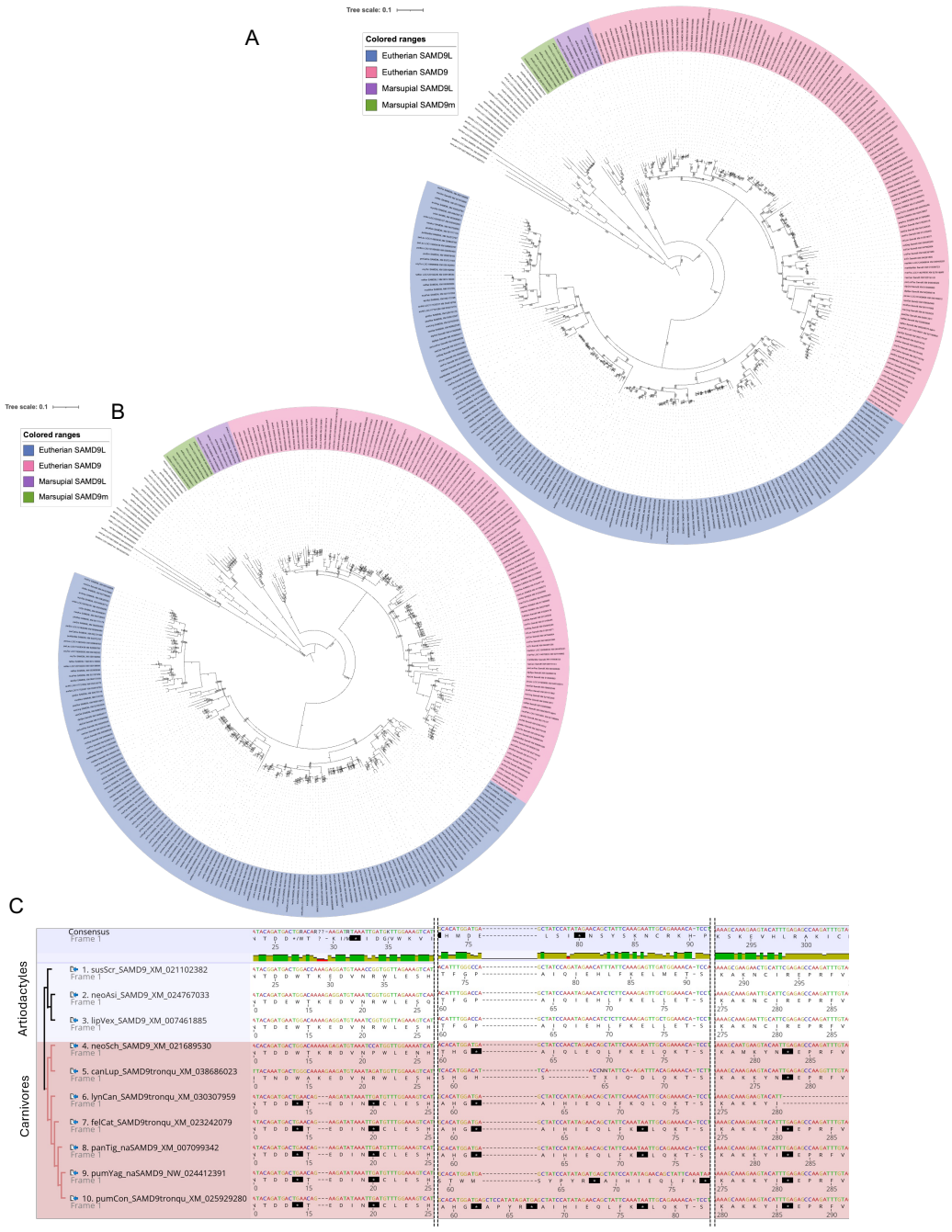
SAMD9		
Domain	Start	End
SAM	0	76
AlbA_2	174	376
SIR2	391	621
P-loop NTPase	626	1036
TPR	1052	1525
OB	1526	1589

SAMD9L		
Domain	Start	End
SAM	0	76
AlbA_2	176	380
SIR2	395	624
P-loop NTPase	629	1040
TPR	1056	1520
OB	1521	1584

Supplementary Figure 2. Associated to Figure 2. AlbA_2 effector domain is shared between human SLFNs, SAMD9/9Ls and bacterial Avs9s, with a conserved catalytic site (A-C), and AlbA_2 effector domain of Avs9 is sufficient for the cell death induction in bacteria (D). A, Representation of the amino acid sequence alignment (from Muscle) showing the shared conserved residues (highlighted by the arrows) between the different protein families – SLFNs, SAMD9/9Ls and Avs9s. The alignment was performed and displayed on UGENE v52.0 with Clustal X colors. **B-C,** Structure clustering of full-length proteins (B) or AlbA_2 domain extraction (C) trees shows three distinct clades for SLFNs, SAMD9/9Ls and Avs9s proteins respectively. **D,** Bacterial drop assay of bacteria transformed with either Avs9$^{1-172}$ or RFP as a control, with induction (500µM IPTG) or without induction (1% glucose). Photos of serial dilutions of bacterial drops.

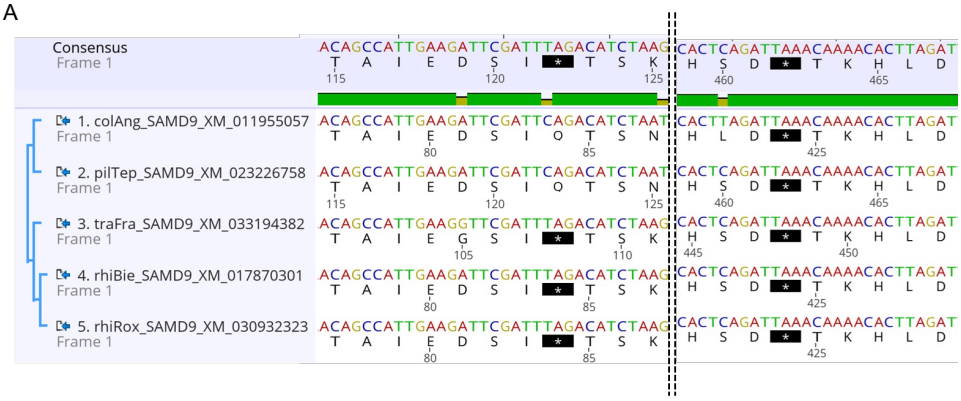


<u>Supplementary Figure 3. Associated to Figure 3. Phylogenetic analyses of SAMD9/9L in mammals associated with Figure 3. A,</u> Maximum likelihood phylogenetic tree of all vertebrate SAMD9s from Supplementary Table 3, determined using IQ-TREE web server (GTR+F+I+G4 identified as the best substitution model by ModelFinder implemented in IQ-TREE). Ultrafast bootstrap support values are shown next to branches. **B,** Maximum likelihood phylogenetic tree determined using PhyML (best model estimated from Smart model selection SMS: GTR+R). aLRT support values are shown next to branches. **C**, Parts of the codon alignment of representative artiodactyl and carnivore SAMD9s, highlighting early STOP (*) codons within carnivore SAMD9s. Amino acid (aa) numbering are shown below sequences. Snapshot from Geneious. Cladogram and sequence names shown on the left (species name abbreviated by three first letters of the genus and three first letters of the species (e.g. felCat, *Felis catus*).

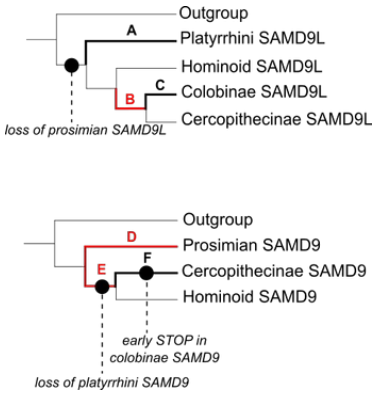


Supplementary Figure 4. Associated to Figure 4. Evidence of early STOP codons in *Colobinae* SAMD9 and of some episodic positive selection during primate SAMD9/9L evolution. A, Snapshot of parts of the codon alignment of *Colobinae* SAMD9 sequences with cladogram and sequence name shown on the left, highlighting early STOP (*) codons. Amino acid (aa) numbering is shown below each aa sequence. View from Geneious. Species name abbreviated by three first letters of the genus and three first letters of the species (e.g. colAng, *Colobus angolensis*). **B,** Simplified primate gene cladograms of *SAMD9L* (top) and *SAMD9* (bottom) showing, in bold, the branches tested for positive selection with aBSREL (HYPHY) and, in red, the branches under significant positive selection (p-value<0.05). The "outgroup" branch corresponds to *Cricetulus griseus* (criGri) and *Tupaia chinensis* (tupChi) for SAMD9 and to criGri only for SAMD9L and is shown here as a rooted tree for representation purposes. **C,** For each branch tested, the table recapitulates the statistical significance of positive selection assessed by LRTs with the Holm-Bonferroni correction, the estimated dN/dS value of the sites in the class under positive selection, and the proportion of sites under positive selection (%).

A



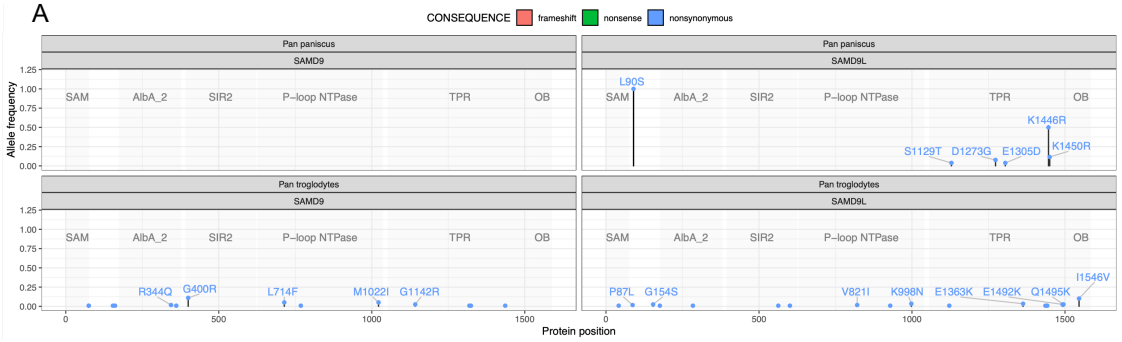
B

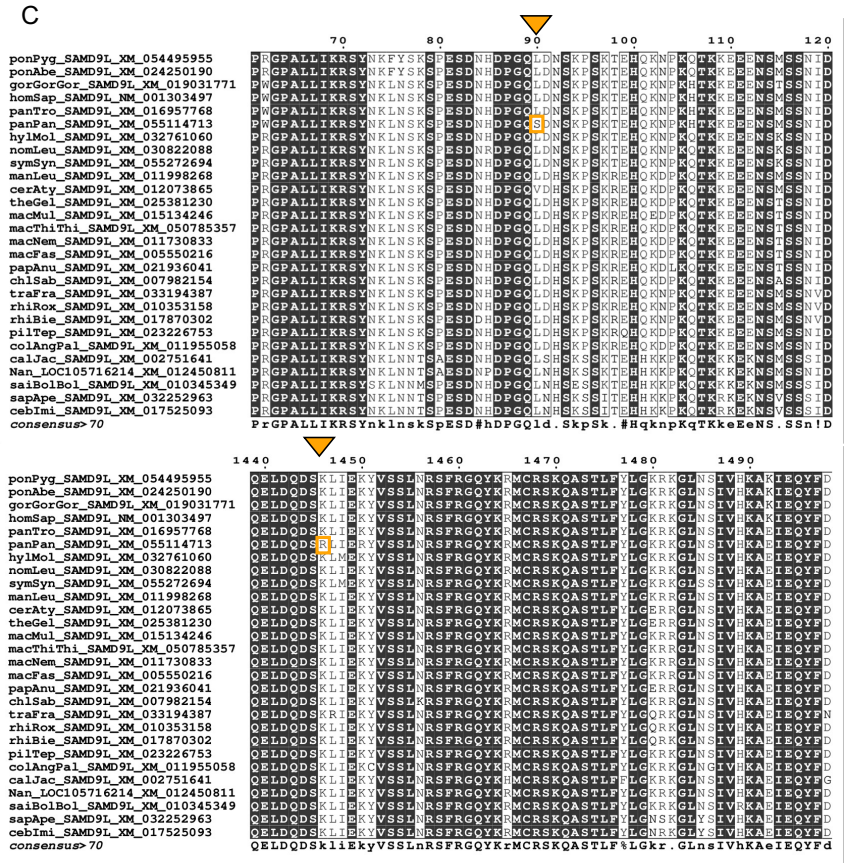


C

Tested branch	p-value	ω class >1	% of sites under positive selection
A	1		
B	0.0059	14	4.7
C	1		
D	0.0453	6.97	5.6
E	0.0001	14.4	2.3
F	1		

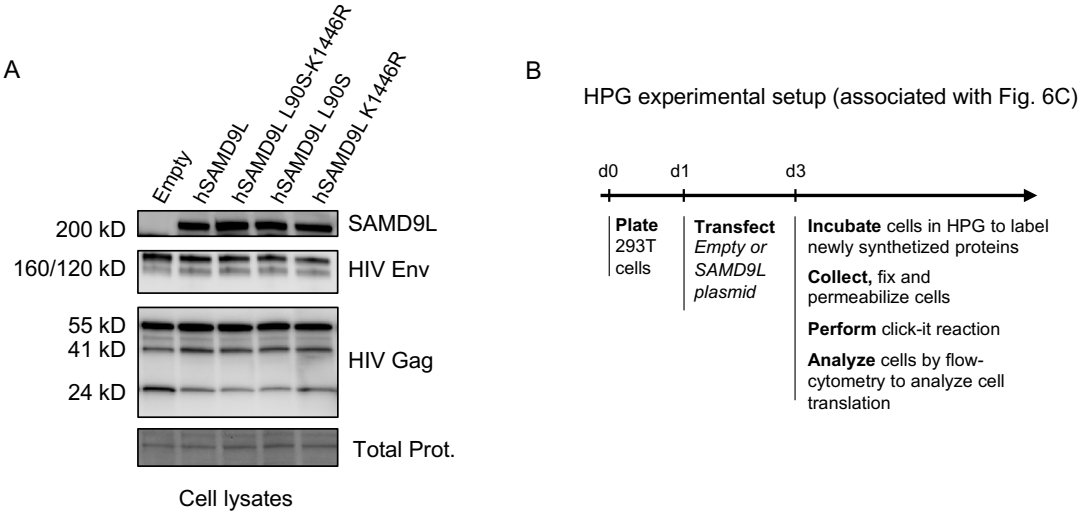
Supplementary Figure 5. Associated to Figure 5. Alleles across SAMD9 and SAMD9L protein sequences in humans, bonobos and chimpanzees. A-B, Polymorphisms frequency in SAMD9 and SAMD9L among the *Pan* (A) and the human (B) populations, with PanTro6 and GRCh38/hg38 as references, respectively. The domains of SAMD9 and SAMD9L are delimited by transparent grey boxes. Labeled coordinates correspond to frequencies > 1%. **C,** Zoom on SAMD9L primate amino acid sequence alignment showing the bonobo-specific variants. Muscle aa alignment representation from ESPript 3.0, https://espript.ibcp.fr (Robert et Gouet 2014). Bonobo-specific variants are at position L90S and K1446R (highlighted in orange).





Supplementary Figure 6. Associated to Figure 6. Effect of bonobo variants on human SAMD9L's activities on cellular and viral protein translation.

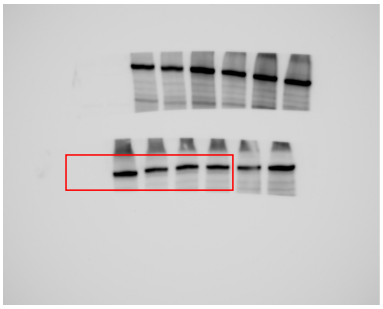
A, Western-blot analysis in the HIV-1 producer cells, following experimental setup of Fig. 5C and associated with experiment shown in Fig. 6A. Shown are the HIV-1 Gag (p55, p41, p24), HIV-1 Env (gp160, gp120), and SAMD9L protein expressions. Loading control is from total protein (BioRad Pre-stained gels). **B,** Experimental setup of the HPG assay to determine impact of SAMD9L bonobo SNPs on cell translation.



Supplementary Figure 7. Uncropped and unprocessed scans of western blots.

Figure 5D

Anti SAMD9L



Total Protein Stainfree revelation

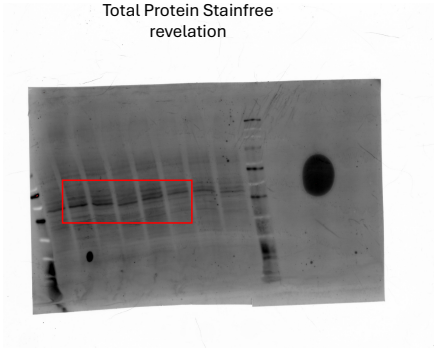
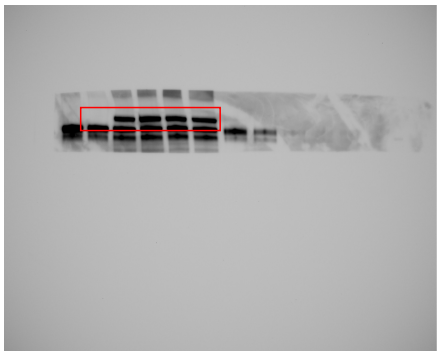


Figure 6B+S6B

Anti SAMD9L
(after Anti Env 16H3, not shown)



Total Protein Stainfree revelation

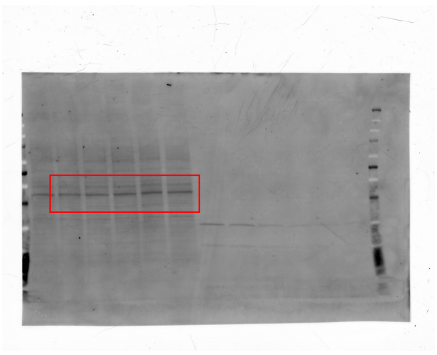
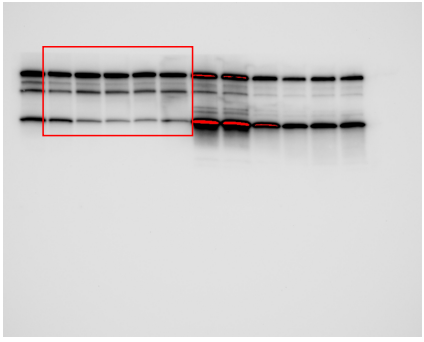


Figure S6B (continued)

Anti Gag p24 183



Anti Env 16H3