nature ecology & evolution

Article

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

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

Received: 15 September 2024

Accepted: 28 July 2025

Published online: 11 November 2025

Check for updates

Alexandre Legrand[1], Rémi Demeure[1,4], Amandine Chantharath[1,4], Carine Rey[1], Julie Baltenneck[1], Cameron L. M. Gilchrist[2], Joana L. Rocha[3], Clara Loyer[1], Léa Picard[1], Andrea Cimarelli[1], Martin Steinegger[2], Francois Rousset[1], Peter H. Sudmant[3] & Lucie Etienne[1]

Human *SAMD9* and *SAMD9L* are duplicated genes that encode innate immune proteins restricting poxviruses and lentiviruses, such as human immunodeficiency virus (HIV). Mutations in these genes are implicated in genetic diseases and cancer. Here we combined structural similarity searches, phylogenetics and population genomics with experimental assays of SAMD9/9L functions to resolve the evolutionary and functional dynamics of these immune proteins, spanning from prokaryotes to primates. We discovered structural analogues of SAMD9/9L in the antibacteriophage defence system, Avs, resulting from convergent evolution. Further, the predicted nuclease site was conserved in bacterial analogues and was essential for cell death, suggesting a fundamental role in defence across different life kingdoms. Despite this shared immunity, we identified genomic signatures of evolutionary arms races in mammals, with remarkable gene copy number variations. We further unveiled that the absence of *SAMD9* in bonobos corresponds to a recent gene loss still segregating in the population. Finally, we found that chimpanzee and bonobo SAMD9Ls have enhanced anti-HIV-1 functions compared with the human orthologue. These SAMD9/9L adaptations probably resulted from strong viral selective pressures, including by lentiviruses, and could contribute to lentiviral resistance in bonobos. Evolutionary characterization of Avs9 and SAMD9/9L provides a deeper understanding of how the immune system adapts to fight viruses over billions of years of evolution.

The continuous arms races between viruses and their hosts have driven the evolution of several immune defences across all life forms[1,2]. In humans, an emerging actor of cell-autonomous antiviral immunity is the gene family composed of the paralogues sterile alpha motif (SAM) domain-containing proteins 9 and 9L (*SAMD9*/*9L*). These interferon-stimulated genes are in tandem on human chromosome 7 and encode large, multidomain proteins with antiviral properties against poxviruses and lentiviruses[3–5]. SAMD9 and SAMD9L inhibit cellular and viral protein translation[6–8], acting through an essential ribonuclease site in the AlbA_2 domain[5,9]. They are also modulators of endosomal trafficking[10,11] and SAMD9 was recently identified as a nucleic acid sensor[12]. Deleterious germline mutations in human SAMD9 or SAMD9L dysregulate their activity leading to severe life-threatening genetic syndromes, such as MIRAGE (myelodysplasia, infection, restriction of growth, adrenal hypoplasia, genital phenotypes and enteropathy)[11], SAMD9L-associated autoinflammatory

A full list of affiliations appears at the end of the paper. e-mail: lucie.etienne@ens-lyon.fr

disease[13], ataxia-pancytopaenia[14] and normophosphataemic familial tumoural calcinosis[15]. Although the *SAMD9/9L* gene family extends beyond humans, its evolutionary history and functional diversification remain largely unexplored.

Recently, some human immune genes were shown to have a deep evolutionary origin in bacterial defences against phages[16]. For example, major eukaryotic antiviral immune sensors or effectors, such as cGAS or Viperin, have originated from prokaryotic antiviral systems[16–19]. Ancestral immunity therefore broadly defines shared immune defences between prokaryotes and eukaryotes through the presence of conserved immune modules (domains or proteins)[18], that results from horizontal gene transfer, convergent evolution or vertical inheritance.

SAMD9 and SAMD9L are large multidomain proteins, which are presumed members of the STAND (signal transduction ATPases with numerous domains) superfamily[20]. SAMD9/9Ls are composed of (from N to C terminus): a SAM, a Schlafen (SLFN)-like AlbA_2 domain with nuclease site, a SIR2, a P-loop NTPase, tetratricopeptide repeats (TPRs) and an oligonucleotide/oligosaccharide-binding OB-fold. All are predicted domains, except the AlbA_2 for which the structure was recently solved[21]. Computational analyses suggested the presence of homologues of SAMD9/9L domains in other animals and bacteria[20]. However, the ancient evolutionary and functional history of SAMD9/9Ls and its potential link to immunity outside mammals remain unknown.

In mammals, the gene family consists of two paralogues, *SAMD9* and *SAMD9L*, which originally duplicated in placentals and have evolved under positive selection[22], suggesting past molecular arms races with pathogens. Interestingly, these mammalian paralogues exhibit both functional redundancy and divergence in their antiviral functions. Both are restriction factors against poxviruses, but with species-specific variations in their susceptibility to poxviral countermeasures[4]. In humans, they have different functions in human immunodeficiency virus (HIV) and lentiviral infections: SAMD9L is antiviral, while SAMD9 has no, or a proviral, effect[5].

Lentiviruses naturally infect most African non-human primate species (simian immunodeficiency viruses (SIVs)), with the notable exception of some species, such as bonobos[23–26]. SIVcpz from chimpanzees is at the origin of the HIV-1 group M, responsible for the AIDS pandemic[27]. Lentiviruses have co-evolved with primates for millions of years and have been selective drivers of adaptation in primate antiviral defence factors, such as APOBEC3G, Tetherin/BST-2 or TRIM5 (refs. 28–30). These species- and lineage-specific adaptations have further shaped host molecular barriers to cross-species transmissions[31,32].

In this study, we combined AI-predicted structural similarity searches, phylogenetics and population genomics with experimental assays of SAMD9/9L functions in prokaryotes and great apes to resolve the evolutionary and functional dynamics of this antiviral system across scales. We notably found that SAMD9/9Ls and some prokaryotic antiviral STAND (Avs) originated from convergent evolution and depend on AlbA_2 domain for their activity, and that SAMD9/9Ls have evolved under recurrent diversifying evolution in mammals by genomic structural variation, notably gene losses and adaptive polymorphisms to lentiviruses in great apes.

Results

SAMD9/9L bacterial structural analogues with conserved multidomain architecture and predicted antiphage activity

Benefiting from recent advances in structural similarity methods, we investigated the evolutionary history of *SAMD9/9L*s across the tree of life. We first used Foldseek[33], which allows fast protein structure alignment and search, with the AlphaFold-predicted SAMD9/9L structures and amino acid sequences as inputs (Methods). We identified 238 hits that share high structural similarity, mainly belonging to bacteria and metazoa (30% template modelling (TM) score and 80% coverage cut-off; Supplementary Table 2). Analysis of bacterial hits with DefenseFinder[34], which performs hidden Markov model (HMM) searches against a database of prokaryotic antiviral systems, showed that some of these hits belong to the Avs family of antiphage immune proteins[35–38]. Avs proteins are nucleotide-binding oligomerization domain-like receptor (NLR)-like proteins that sense viral phage infections by their C-terminal sensor domains and perform specific antiviral functions through their N-terminal effector domains. The effector domain may have diverse specific activities, such as ATP degradation via a PNP domain or NAD^+ depletion via a SIR2 domain[17,37,39,40]. Here we found that SAMD9/9L shares strong structural similarity with Avs5 from *Desulfobacula* sp. and from a newly identified protein, which we named 'Avs6', from *Labedella endophytica* (Uniprot IDs A0A1F9N8W4 and A0A3S0VH60, respectively) (Fig. 1a). However, Avs5 and Avs6 lack the N-terminal AlbA_2 domain from SAMD9/9L, with Avs6 instead encoding a PNP domain predicted to degrade ATP molecules (Fig. 1a). Remarkably, our analyses also identified other bacterial proteins sharing strong structural similarity with human SAMD9/9L on up to 88% of the protein length (for example, 1,407 of 1,589 amino acids total in SAMD9 and 1,400 of 1,584 in SAMD9L for A0A7T4VS34 from *Pseudomonas fluorescens*). This similarity covers all its functional domains, except the first N-terminal SAM, which seems exclusive to eukaryotes[41]. Despite the low amino acid sequence identity (around 15%), the structural similarity is highly significant (TM score up to 0.56; Supplementary Fig. 1a)[42]. We therefore propose 'Avs9' as a name for bacterial Avs proteins harbouring the same domain composition as SAMD9/9L at the exception of the SAM domain, referring to the latter gene family name (Fig. 1a). Furthermore, our analysis identified hundreds of uncharacterized proteins across various domains of life: metazoa, bacteria and archaea. It is however notable that, although STAND proteins are also present in Plantae and fungi[43,44], we did not retrieve any significant hits in these kingdoms of life, neither in algae nor in protozoa. Further searches with various inputs, including Avs9, and more relaxed parameters did not recover any.

To investigate the evolutionary history of the structural analogues, we performed maximum likelihood phylogenetic analyses, using IQ-TREE, of both the sequence and the structural alignments of these proteins (from Muscle and FoldMason, respectively) (Supplementary Fig. 1b). We found that SAMD9/9Ls and Avs9s were in two distinct clades falling within the same lineage. However, because the proteins bear different domain compositions, we next performed analyses for specific domains, individually (Supplementary Fig. 1c). Starting with the central P-loop NTPase domain, which is a domain common to all hits, we similarly found that Avs9 and SAMD9/9L clades fell into the same lineage in the phylogeny (Supplementary Fig. 1d). Next, we only kept proteins encoding for an AlbA_2 effector domain, resulting in 23 hits (the small number of hits are due to the AlphaFold Database clustered at 50% identity): 13 eukaryotic SAMD9/9Ls and 10 bacterial Avs9s. The resulting tree topology also showed two clades corresponding to SAMD9/9Ls and Avs9s (Fig. 1b). We further observed scattered absence of SAM, OB-fold or SIR2 domains (Fig. 1b). While this suggests that these proteins have a propensity for domain modularity and that these domain losses could be tolerated, the functional impact of such modulation would require further investigations. Overall, the extensive structural similarity of SAMD9/9L- and Avs9-like proteins observed across such a vast evolutionary range is remarkable. It suggests a strong selective advantage driving the structural and functional conservation of this putative immune antiviral system from bacteria to mammals.

Eukaryotic SAMD9/9Ls and bacterial Avs9s result from convergent evolution

To determine whether eukaryotic SAMD9/9Ls result from the ancestral acquisition of a full-length bacterial analogue (such as Avs9) or emerged through convergent evolution, we performed new structural similarity searches for AlbA_2-containing proteins and studied the phylogenetic

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

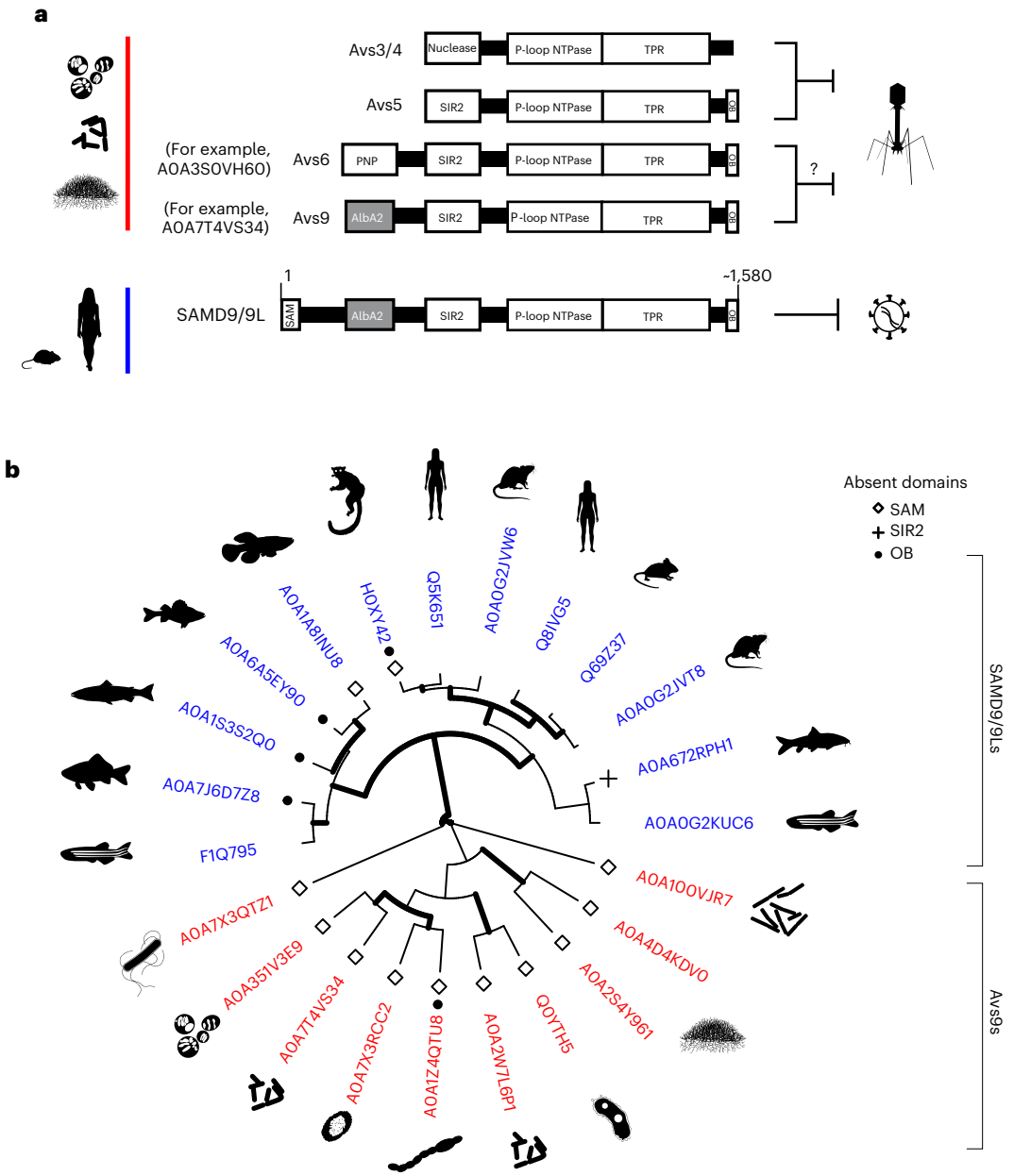


Fig. 1 | Structural homology analyses show strong similarity between *SAMD9/9L* gene family and Avs antiphage proteins. a, Structural homology searches of human SAMD9/9L in prokaryotes identify known and predicted Avs antiviral systems. Linear representation of the multidomain organization of proteins with strong SAMD9/9L structural similarity showing a common conserved organization, on more than 1,300 amino acids for Avs9. Avs3, Avs4 and Avs5 were previously identified[35,37], while Avs6 and Avs9 are proposed names of newly identified Avs members (in parentheses, representative members with protein UniProt ID). **b**, Circular ultrametric tree representing structural clustering of SAMD9/9Ls and Avs from FoldMason MSTA showing widespread and diverse SAMD9/9L structural analogues in bacteria (red) or metazoan (blue) SAMD9/9Ls. Tree rooting is only for representation purposes. Statistical supports are from 1,000 bootstrap replicates (values above 90 are represented by thick lines). Protein with an absent domain have either a square, a cross or a filled circle at the tip of the corresponding branch, for the absence of SAM, SIR2 or OB, respectively. Credit: Species silhouettes from phylopic (https://www.phylopic.org) under a Creative Commons license: *Homo sapiens sapiens*, Edwin Price (CC0 1.0); *Rattus tunneyi*, Carly Monks (CC0 1.0); *Methylococcus capsulatus* and *Acidimicrobium ferrooxidans*, Matthew Crook (CC BY-SA 3.0); *Streptomyces*, Guillaume Dera (CC0 1.0); Enterobacteria phage T2, T. Michael Keesey (CC BY-SA 3.0); *Euoticus pallidus*, Joseph Wolf and T. Michael Keesey (CC0 1.0); *Homo sapiens sapiens*, by Carlo De Rito (CC0 1.0); *Danio rerio*, Ian Quigley (CC BY 3.0); *Carassius auratus*, Andrew Hoadley (CC0 1.0); *Luciobarbus graellsii*, Carlo Cano-Barbacil (CC0 1.0); *Salmo salar*, Domino Joyce (CC0 1.0); *Perca flavescens*, NOAA Great Lakes Environmental Research Laboratory (illustration) and Timothy J. Bartley (silhouette) (CC BY-SA 3.0); *Nothobranchius* sp., Ryan Cupo (CC0 1.0); *Rattus norvegicus*, Mozillian (CC0 1.0); *Mus musculus*, Jiro Wada (CC0 1.0); *Acidobacterium capsulatum*, Poribacteria morphotype 4, *Chlorobium* and *Bacillus subtilis*, Matthew Crook (CC BY-SA 3.0).

distribution of this effector domain across the tree of life (Fig. 2a,b). First, we found that AlbA_2 was not only present in Avs9-like prokaryotic proteins, but was widely represented in bacteria. Concerning eukaryotes, well-known AlbA_2-containing proteins included members of SAMD9/9L and SLFN antiviral immune factors, which, in our analysis, fell in distinct clades (except for SLFNL1, which is distantly related to other SLFNs[45] and whose evolutionary history was uncertain), suggesting different evolutionary origins (Fig. 2b and Supplementary

Fig. 2b,c). As for the ten identified Avs9s, they all clustered within a single clade that was distant from eukaryotic SAMD9/9Ls (Fig. 2b and Supplementary Fig. 2b,c), showing that full-length Avs9- and SAMD9/9L-like proteins did not originate from a single common ancestor. Overall, the presence of evolutionarily distant AlbA_2 domains together with structural similarity and a similar domain organization strongly suggest that bacterial Avs9s and eukaryotic SAMD9/9Ls evolved similar domain architectures through convergent evolution.

Antiphage defence systems tend to physically cluster in defence islands of bacterial genomes[46]. To investigate the possible immune function of bacterial AlbA_2 domains, we used DefenseFinder to evaluate their propensity to be encoded in genomic neighbourhoods of known defence systems. We found a significant association of bacterial AlbA_2 domains with defence systems in all tested clades, including Avs9 (Fig. 2b), strongly suggesting that their primary function is immune defence.

AlbA_2 domain is a shared effector determinant of the activity of SAMD9/9Ls and Avs9s

To investigate the function of bacterial Avs9 proteins, we first looked for the conservation of the AlbA_2 catalytic site. This active ribonuclease site is composed of three to four negatively charged residues, which are necessary for SAMD9/9L and SLFN11/12/13/14 antiviral activities[5,9,47–52]. Here we found similar residues and structure in Avs9 AlbA_2 (E14, D19, D45, E65; Fig. 2c and Supplementary Fig. 2a), suggesting the presence of similar catalytically active site and function.

Second, to investigate Avs9 activity in vivo, we selected and synthesized an Avs9 from *P. fluorescens* (A0A7T4VS34), and we cloned its full-length gene, or the AlbA_2 domain alone, under an inducible promoter into *Escherichia coli*. Expression of full-length Avs9 (Fig. 2d), or the AlbA_2 domain alone (Supplementary Fig. 2d), led to major cell death upon mild induction at 25 °C. Interestingly, the *P. fluorescens* Avs9 cell death induction was abolished by introducing a single-residue mutation in the AlbA_2 predicted catalytic site (D45N; Fig. 2d and Supplementary Fig. 2d). Of note, an equivalent mutation in human SAMD9/9L and SLFN AlbA_2 also abolishes their activity. Therefore, our results strongly suggest a cell-killing defence function of bacterial Avs9, dependent on AlbA_2 and its predicted nuclease active site.

Ancient duplication followed by frequent CNVs of the *SAMD9/9L* gene family in mammals

Although *SAMD9/9L*s are present in several vertebrate hosts (Fig. 1b), the *SAMD9/9L* duplication at the origin of the human paralogous *SAMD9* and *SAMD9L* genes seems more recent. This duplication was previously described in placental eutherians, after the divergence of marsupials from placentals[22]. To address the evolutionary history of the duplicated *SAMD9/9L*, we first performed genomic analyses of the *SAMD9/9L* gene family locus in various vertebrate species focusing on mammals: 38 ungulates, 22 chiropterans, 45 carnivores, 34 primates, 41 rodents and 27 additional species from other mammalian and non-mammalian vertebrate orders (Fig. 3a). We found that all these analysed vertebrate species, except monotremes ($n = 2$), exhibited at least one gene from the *SAMD9/9L* gene family (Fig. 3a).

Furthermore, thanks to genomic sequence advances and contrary to previous findings[22], our observations indicated that marsupials also possess two copies and therefore the gene duplication of *SAMD9/9L*s was not restricted to eutherians (Fig. 3a,b). To elucidate the origin of *SAMD9/9L* in mammals, we performed phylogenetic analyses from 301 homologue sequences of 189 mammalian species spanning 160 million years of divergence (using the National Center for Biotechnology Information (NCBI) blastn implemented in DGINN[53]) and from 18 non-mammalian vertebrate species (12 aves and 6 amphibians) as outgroups (Fig. 3b with selected species and IQ-TREE tree, Supplementary Fig. 3a,b for the complete IQ-TREE and PhyML trees using sequences from Supplementary Table 3; of note, PhyML trees were similar to IQ-TREE trees at key branches). Despite the presence of two copies in marsupial genomes (Fig. 3a), one of them, named *SAMD9m*, did not group with the eutherian *SAMD9* or *SAMD9L* clade in the homologous gene tree, but branched outside (Fig. 3b, statistical significance assessed from 1,000 bootstrap replicates). Therefore, it is likely that the marsupial *SAMD9m* resulted from an ancestral independent duplication predating the divergence of marsupials and eutherians (Fig. 3b). Following this divergence, one copy was potentially lost in eutherians, followed by a subsequent duplication event (Fig. 3c). Alternatively, other hypotheses may involve a single duplication event followed by gene conversion within the two eutherian copies, or loss of a copy through incomplete lineage sorting.

In placentals, following the duplication event that gave rise to *SAMD9* and *SAMD9L* orthologues, we identified at least ten independent losses of one of the two paralogues throughout mammalian evolution (five losses in the main represented lineages of Fig. 3a, as well as additional losses within orders: Fig. 3d and Supplementary Fig. 3c). Notably, artiodactyls experienced the loss of *SAMD9L*, while carnivores lost *SAMD9*. Furthermore, although the synteny and the copy numbers remained largely conserved within the divergence of the last two groups, radiations of primates, bats and rodents, exhibited many *SAMD9/9L* gene losses and some duplication events (Fig. 3d and Supplementary Fig. 3c; Data availability).

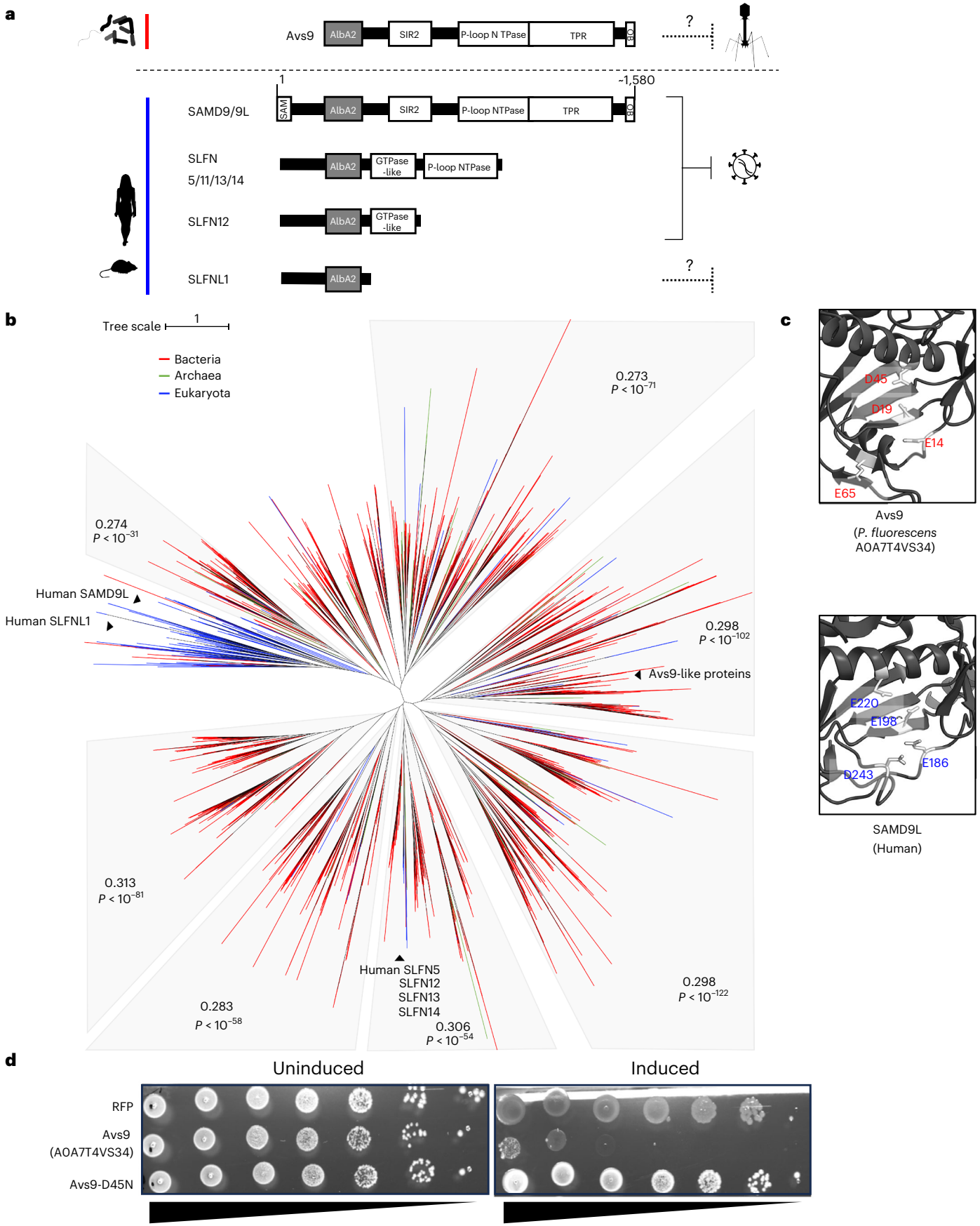
Ancient and recent unfixed gene losses during primate *SAMD9/9L* evolution, through different genetic and genomic mechanisms

Genomic structural changes, such as gene duplication and loss (copy number variations (CNVs)), alongside mutation and recombination events, are an important source of genetic diversity upon which natural selection can act, and thus have strong adaptive potential during virus–host arms races[54,55]. Such genomic adaptations during mammalian evolution have been rampant in response to past lentiviral or poxviral epidemics[30,56]. Because of (1) the very high rate of CNVs in primate *SAMD9/9L*s, (2) the co-evolution of primates with lentiviruses for millions of years[32] and (3) the potential lentivirus-driven adaptation of SAMD9/9L (ref. 5), we performed in-depth phylogenomic, genetic and positive selection analyses in primates.

We identified at least four independent gene loss events during primate evolution (Fig. 4a). *SAMD9L* was lost in prosimians, while bonobos and the common ancestors of Platyrrhini and Colobinae

Fig. 2 | Avs9 is part of prokaryotic defence systems and induces cell death in bacteria, through its SAMD9/9L/SLFN-analogous active site in the AlbA_2 domain. a, Linear representation of the multidomain organization of key proteins bearing an AlbA_2 domain: prokaryotic Avs9 and proteins from the SAMD9/9Ls and SLFNs. **b**, Unrooted phylogenetic tree of an MSA of AlbA_2 domains detected in proteins from kingdoms of life, visualized on iTOL. The scale bar represents the number of substitutions per site. Shown for each clade is the defence score, that is the fraction of bacterial AlbA_2 domains encoded in the vicinity of known defence system, as well as a *P* value representing the significance of this association (Supplementary Table 1; Methods). Statistics were performed using the two-sided binomial tests with FDR correction. **c**, Predicted structures of AlbA_2 domains of *P. fluorescens* A0A7T4VS34 (Avs9) and human SAMD9L, showing conserved SLFN-like nuclease catalytic site. Residues forming the catalytic sites are in clear grey with their coordinates in red (Bacteria) or blue (Eukaryota). **d**, Tenfold serial dilutions of *E. coli* cells transformed with plasmids encoding Avs9, Avs9-D45N or RFP as a control, with induction (100 μM IPTG) or without induction (1% glucose). Shown are photographs of bacterial drops. Credit: Species silhouettes are from phylopic (https://www.phylopic.org) under a Creative Commons license: *Homo sapiens sapiens*, Edwin Price (CC0 1.0); *Rattus tunneyi*, Carly Monks (CC0 1.0); *Acidimicrobium ferrooxidans*, Matthew Crook (CC BY-SA 3.0).

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



experienced the independent loss of *SAMD9*. Genomic analyses showed that *SAMD9* loss in Platyrrhini resulted from the complete loss of the *SAMD9* genomic locus in the common ancestor. In contrast, *SAMD9* loss in Colobinae occurred through a different genetic mechanism by single nucleotide changes introducing several premature stop codons in the coding sequence (Fig. 4a and Supplementary Fig. 4a).

We next specifically tested whether episodes of positive selection occurred in *SAMD9* or *SAMD9L* in some primate lineages that experienced paralogue loss. We performed targeted branch-specific analyses, using the adaptive branch-site random effects likelihood (aBSREL), testing specifically lineages associated with gene loss or paralogue retention ('tested branch', also known as 'foreground' branches) in the *SAMD9* or the *SAMD9L* gene trees (others branches were set as 'background branches')[57] (Supplementary Fig. 4b,c). Our analysis suggested that several primate lineages that lost either *SAMD9* or *SAMD9L* might have been the targets of episodic positive selection (Supplementary Fig. 4b,c). These losses could result from adaptation to ancient pathogen challenges, cost of maintaining the two copies or relaxation of functional constraints that occurred on specific branches.

Bonobos and chimpanzees are human's closest living relatives with a genetic divergence of approximatively 1.3% with humans and only of 0.4% amongst themselves[58]. Despite this strong proximity, bonobos possess a unique genetic distinction, with a 41.46 kilobase (kb) deletion in the *SAMD9/9L* locus. In fact, they stand out as the sole hominid with a single copy of the *SAMD9/9L* gene family, retaining only *SAMD9L* (ref. 59). To characterize the recent loss of *SAMD9*, we investigated the prevalence of the 41.46 kb deletion in the *SAMD9/9L* locus in bonobos at a population level (*Pan paniscus*; $n = 13$) in a joint analysis that additionally included all currently recognized chimpanzee subspecies or populations (*Pan troglodytes* spp.; $n = 59$) (Fig. 4b). We found that, among 13 individuals, 10 bonobos exhibited the same deletion with the complete absence of the *SAMD9* gene, indicating common homozygous genomic deletion (Fig. 4c). However, the remaining three bonobos presented this deletion on a single chromosome, representing a heterozygous genomic absence of *SAMD9* (Fig. 4c). Importantly, the analysed bonobo individuals are not related[60], suggesting that, despite the small sample size, the heterogeneous genomic makeup in SAMD9/9L may be representative of the population. This was in sharp contrast with chimpanzees, which had $SAMD9^{+/+}$ present in all 59 individuals, a pattern shared with humans (Fig. 4c). This suggests a recent *SAMD9* genomic loss event specific to the bonobo lineage still segregating in the population, potentially impacting the immunity of bonobos.

Chimpanzee and bonobo (*Pan*) SAMD9Ls have an increased anti-HIV-1 activity compared with human SAMD9L

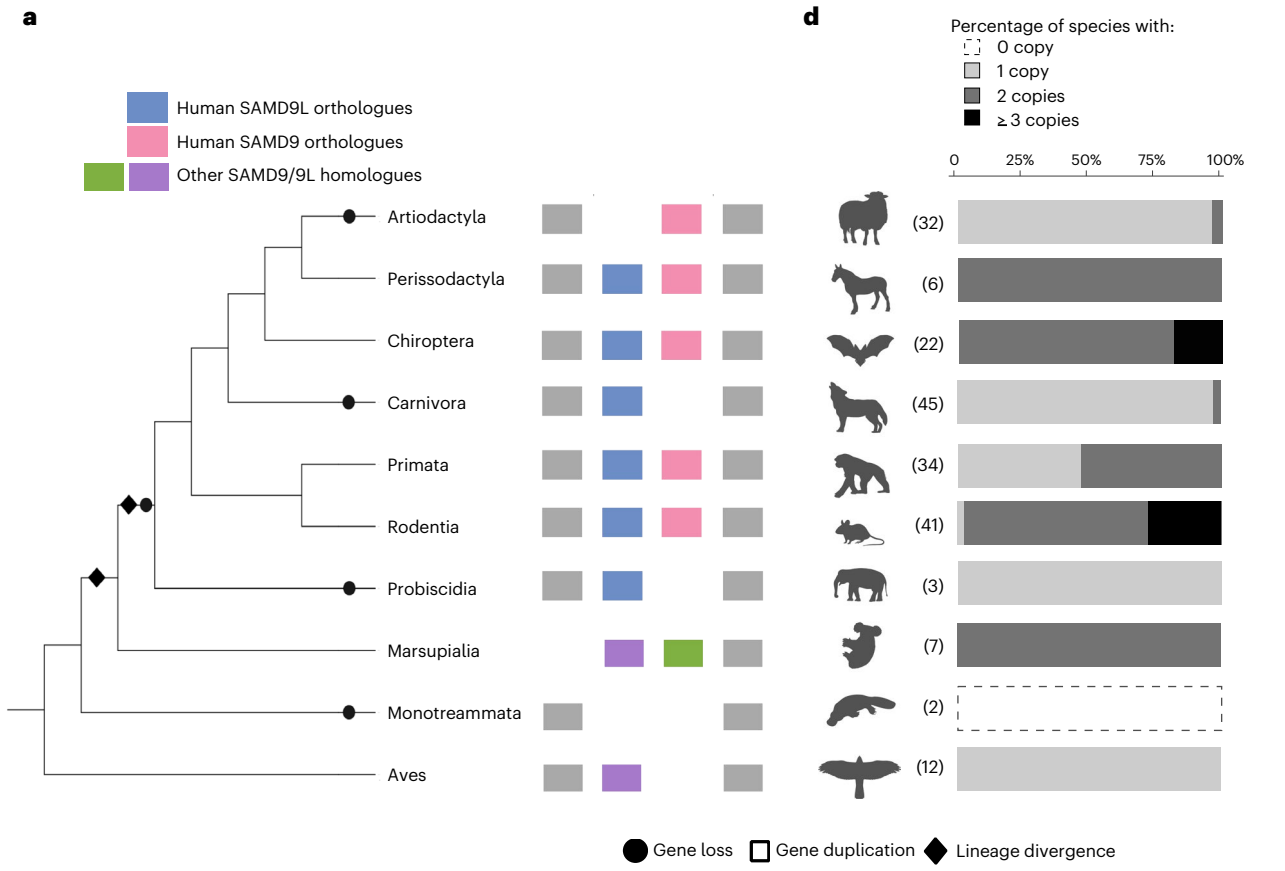
Beyond the genomic loss that occurred in the *SAMD9/9L* locus during hominid evolution, we investigated the evolutionary divergence of *SAMD9* and *SAMD9L* at the genetic level. We therefore analysed the non-synonymous single nucleotide polymorphisms (SNPs) within the coding sequences of both genes for the 72 *Pan* individuals, using panTro6 genome as reference, as well as for more than 4,000 humans (Fig. 5a and Supplementary Fig. 5a,b). For *SAMD9*, we found very few non-synonymous SNPs amongst *Pan*. This amino acid conservation was particularly exemplified in the 12 Eastern chimpanzees (*Pan troglodytes verus*) and in the 3 bonobos encoding a single *SAMD9* copy, as they encoded an identical protein sequence (Fig. 5a). In *SAMD9L*, we found a more widespread distribution of missense SNPs in chimpanzees and bonobos (Fig. 5a). Remarkably, we found two specific variants in *SAMD9L* that stood out because of their high frequency in the bonobo population. All bonobos ($n = 13$) encoded for a homozygous serine (S) at position 90 and 9 out of 13 bonobos encoded an arginine (R) at position 1,446 (either homozygous or heterozygous) (Fig. 5a and Supplementary Fig. 5a). Of note, there were no *SAMD9L* SNPs differentiating $SAMD9^{+/-}$ and $SAMD9^{-/-}$ bonobos. Intriguingly, chimpanzees and all other primates analysed to date, including the 4,099 human genomes, encoded a leucine L90 and a lysine K1446 (Fig. 5a and Supplementary Fig 5a,c), suggesting that these variants are specific to bonobos amongst primates. Furthermore, compared with bonobos, Eastern and Central chimpanzees showed a more widespread distribution of SNPs within *SAMD9L*, and no missense polymorphisms appeared at high frequency in the chimpanzees (Fig. 5a and Supplementary Fig. 5a).

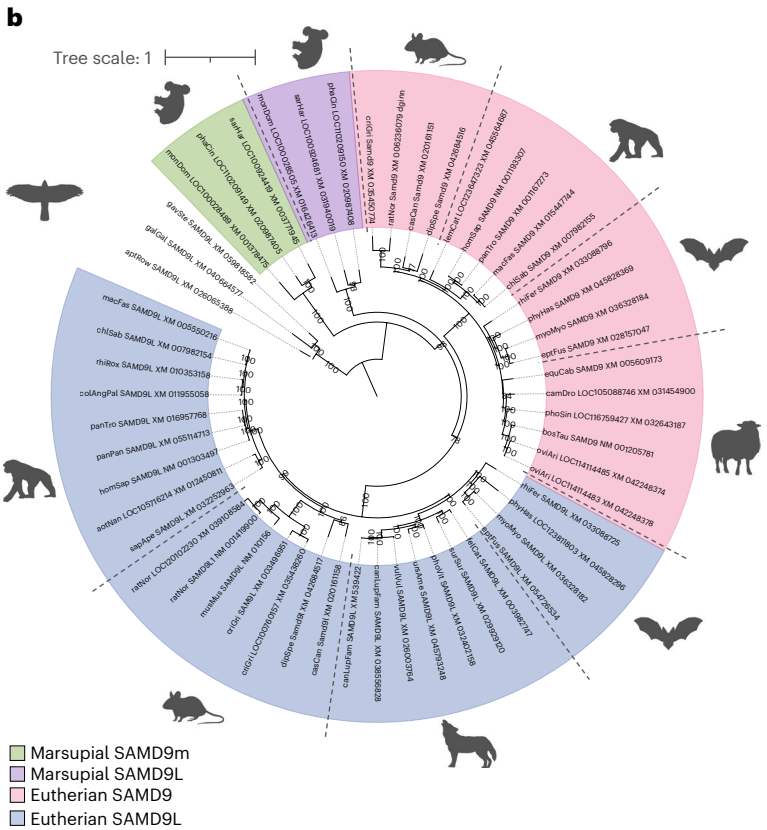
Human SAMD9L inhibits cellular protein synthesis and restricts lentiviral HIV-1 infection[5,8,61,62]. This is particularly interesting in the context of natural lentiviral infections in hominids, where humans and two chimpanzee subspecies are infected by HIV-1 and SIVcpz, respectively. Yet, two chimpanzee subspecies (Eastern and Nigerian-Cameroon chimpanzees) and bonobos have no evidence of modern natural lentiviral infections[25–27,63] (Fig. 4b). Furthermore, pandemic HIV-1 in humans originated from cross-species transmission of SIVcpzPtt from Central chimpanzees (*Pan troglodytes troglodytes*)[27,64].

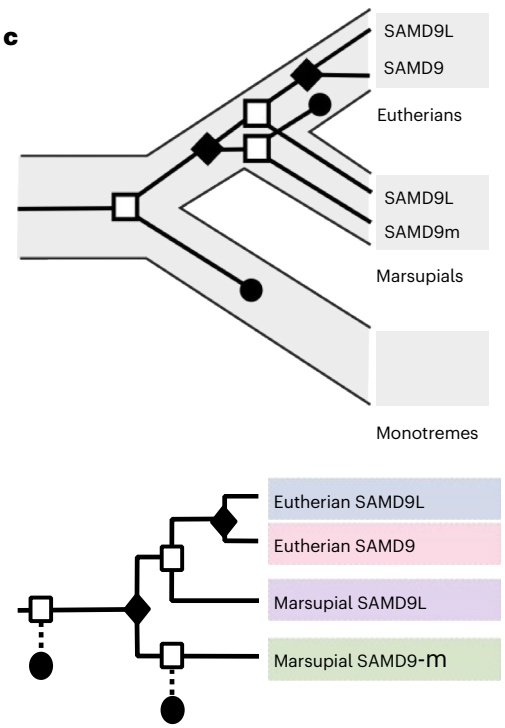
We therefore determined the functional consequences on lentiviral infections of the genotypic differences between human, chimpanzee and bonobo SAMD9Ls. We cloned the native chimpanzee SAMD9L and the two major variants of bonobo SAMD9L (bonobo R1446 and bonobo K1446) in an expression plasmid to compare them with human SAMD9L (Fig. 5b,c). Of note, the chimpanzee SAMD9L has eight amino acid changes compared with the human one (Fig. 5b). We investigated their functions in the context of HIV-1 replication, as in ref. 5, as well as SIVcpzPtt replication (SIVcpz EK505, a kind gift from B. Hahn[27,65]). Briefly, we cotransfected 293 T cells with the mCherry-SAMD9L plasmids along with an infectious molecular clone (IMC) encoding a transmitted/founder HIV-1 natural strain (pWITO, as in ref. 5) or an SIVcpzPtt strain (EK505). Two days later, we measured viral and cell protein expression in the producer cells (Fig. 5c–e and Supplementary Fig. 5d) and the infectious virus yield by Tzm-bl reporter assay (Fig. 5c–e). Remarkably, although all ectopic SAMD9Ls were expressed at similar levels, we found that chimpanzee and bonobo SAMD9Ls had a significant

Fig. 3 | Ancient duplication followed by frequent CNVs of the *SAMD9/9L* gene family in mammals. a, Representation of the *SAMD9*/9L gene family locus in each mammalian order and other vertebrates. Order cladogram is presented on the left with diamonds and rounds on the branches representing events of gene gain and loss, respectively, in the major lineages only. Coloured rectangles represent the gene members of the *SAMD9*/9L gene family with orthologues to human *SAMD9L* and *SAMD9* in blue and pink, respectively. Grey rectangles represent adjacent syntenic genes. **b**, Maximum likelihood phylogenetic tree (IQ-TREE, GTR + F + I + G4 substitution model) generated with selected *SAMD9/9L* homologues from eutherians and marsupials, as well as aves and amphibians, used as outgroups. Bootstraps are from 1,000 replicates. The complete tree is shown in Supplementary Fig. 3a. The scale bar represents the number of substitutions per site. **c**, Schematic diagrams of the origin of *SAMD9/9L* duplication in mammals. Top, grey tree in the background represents the species evolution of the three mammalian groups. Black tree inside the grey one represents the evolution of the *SAMD9/9L* gene tree. Bottom, alternative representation with the gene tree cladogram. Legend is embedded for **a** and **c**. **d**, For each order, CNVs in the *SAMD9/9L* gene family are indicated by histograms. Alignments and trees are given in Data availability. Bar colours (grey scale) represent the number of *SAMD9/9L* copies (legend embedded). Bar lengths indicate the proportion of genomes with the indicated number of copies for each order, with the total number of analysed genomes presented in parentheses for each order (aligned to **a**). Credit: Species silhouettes are from phylopic (https://www.phylopic.org) under a Creative Commons license: *Ovis aries*, Andrés Delgado (CC0 1.0); *Pliohippus*, Zimices (CC BY-SA 3.0); *Plecotus austriacus*, Andy Wilson (CC0 1.0); *Canis lupus monstrabilis*, Tracy A. Heath (CC0 1.0); *Gorilla gorilla*, Andy Wilson (CC0 1.0); *Apodemus sylvaticus*, Anthony Caravaggi (CC0 1.0); *Elephas maximus*, Andy Wilson (CC0 1.0); *Phascolarctos cinereus*, Gavin Prideaux (CC0 1.0); *Ornithorhynchus anatinus*, Sarah Werning (CC0 1.0); *Accipiter gentilis*, Andy Wilson (CC0 1.0).

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







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

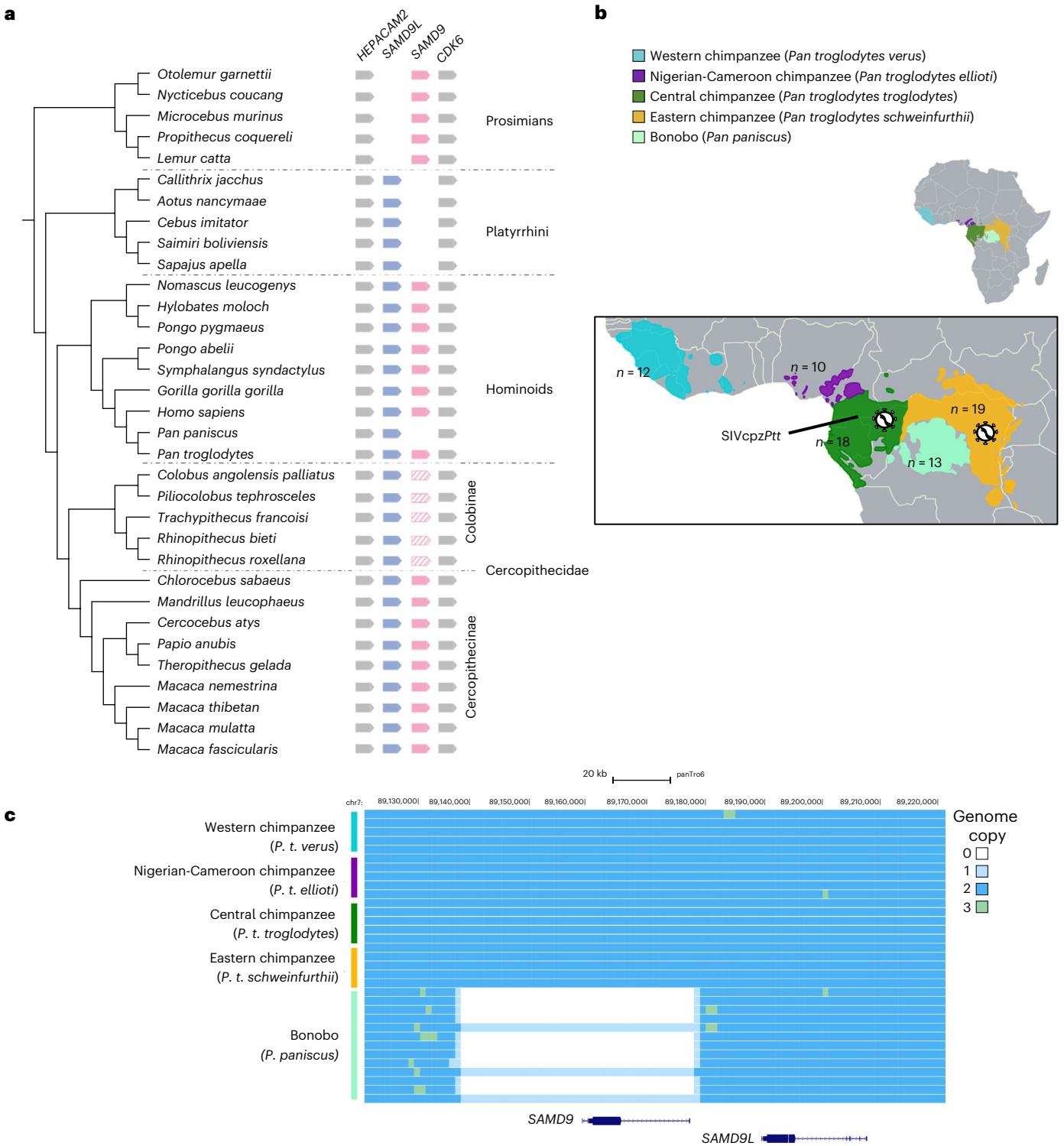


Fig. 4 | Ancient and recent unfixed gene losses in primates, through different genetic and genomic mechanisms. a, Representation of the *SAMD9/9L* genomic locus from primate genomes. Species cladogram is presented on the left. *SAMD9L* and *SAMD9* are in blue and pink, respectively. Adjacent syntenic genes are in grey. Genes containing early stop codons are hatched. **b**, Africa map with inset representing the current geographic ranges of *Pan* populations and their status regarding natural SIV infections. Numbers of individuals studied for their whole-genome sequences are indicated for each (sub)species. Populations naturally infected by SIVcpz are highlighted by a virus symbol. **c**, *SAMD9/9L* genomic locus alignment among *Pan* individuals showing recent unfixed loss of *SAMD9* in the bonobo population.

increase in anti-HIV-1 pWITO activity compared with human SAMD9L (Fig. 5d, $P < 0.005$), suggesting some species-specificity.

By testing effects on SIVcpzPtt EK505, we first showed that human SAMD9L was also restrictive against this SIVcpzPtt strain. Yet, chimpanzee and bonobo SAMD9Ls did not present an increased anti-SIVcpz activity compared with human SAMD9L, suggesting possible lentiviral-strain specificity. Therefore, chimpanzee and bonobo SAMD9Ls seem to have an increased anti-HIV-1 activity,

as compared with human SAMD9L (Fig. 5e). It is further possible that SIVcpz, naturally circulating in chimpanzee populations, adapted to the *Pan* SAMD9L increased antiviral function.

Bonobo-specific polymorphisms confer an increased anti-HIV-1 activity to human SAMD9L without compromising cellular protein synthesis

We wondered whether minimal changes in human SAMD9L informed from some of these natural *Pan* variants could impact human SAMD9L functions. We specifically tested the bonobo SNPs, which could constitute species-specific adaptations with functional implications. We therefore cloned L90S and/or K1446R variants in the context of the human SAMD9L plasmid and investigated their effects on two key functions of human SAMD9L: antilentiviral function as well as cellular protein synthesis shutdown (Fig. 6).

First, we investigated the effect of these bonobo-specific variants on human SAMD9L function in the context of lentiviral replication, as in Fig. 5. Interestingly, we found that ectopic SAMD9L-L90S/K1446R and the single variants were expressed at a similar level to wild-type SAMD9L, but had a significant twofold increase in anti-HIV-1 pWITO activity (Fig. 6a). The double-mutant SAMD9L-L90S/K1446R appeared the most restrictive, while the single mutants had intermediate effects, suggesting additive functions (Fig. 6a). The increased effects of the human SAMD9L mutant with L90S and/or K1446R seemed independent of HIV-1 protein translation shutdown (Supplementary Fig. 6a).

Second, we tested the activity of SAMD9L-L90S/K1446R on SIVcpzPtt EK505 replication and found that, similarly to wild-type *Pan* SAMD9Ls, the specific SNPs in the context of the human SAMD9L did not increase its anti-SIVcpz function (Fig. 6b).

Lastly, we assessed the extent of whole cellular translation in human 293 T cells transfected with high doses of wild-type human mCherry-SAMD9L or mCherry-SAMD9L-L90S/K1446R or the single mutants. We used Click-iT L-homopropargylglycine (HPG) synthesis assays, which measures the incorporation of HPG an analogue of methionine into newly synthesized proteins by flow cytometry Supplementary Fig. 6b). As previously shown[5,8,61,62], we found a dose-dependent shutdown of cellular translation in cells ectopically expressing human SAMD9L (mCherry$^+$) normalized to the control mCherry$^-$ cells (Fig. 6c and Supplementary Fig. 6b). Importantly, we found that hSAMD9L-L90S/K1446R and the single variants showed no differences in the cellular translation repression compared with the human wild-type SAMD9L (Fig. 6c). This suggests that the non-synonymous bonobo variants in SAM and TPR domains do not change the activity of human SAMD9L on cellular protein synthesis.

Overall, bonobo-specific polymorphisms specifically enhance human SAMD9L antiviral function against HIV-1, without affecting the translation shutdown function.

Discussion

This study unveils key aspects of the functional evolution of *SAMD9/9L*s at different timescales, highlighting multidomain and functional convergence between metazoan *SAMD9/9L*s and prokaryotic Avs9s, as well as recurrent genetic and genomic adaptations in mammals from ancient to very recent times. On one hand, we identified SAMD9/9L structural analogues in bacterial defence systems that induce cell death with similar AlbA_2 effector determinants, suggesting a remarkable shared immune factor across billions of years. On the other hand, our analyses of mammalian *SAMD9/9L* revealed dynamic and episodic adaptations, notably in primates, probably in response to epidemics, including from lentiviruses. By an 'immuno-evo' framework, we bring key insights to understand the duality between the maintenance of a key antiviral shared immunity and its constant adaptations to pathogens.

SAMD9/9L as shared immunity in prokaryotes and metazoans

The *SAMD9/9L* gene family may be part of the 'shared immunity', in which antiviral mechanisms are similar between life kingdoms, either resulting from (1) horizontal gene transfer of bacteria, (2) from vertical inheritance originating from LUCA (last universal common ancestor) or (3) from convergent evolution[16,18,66]. The list of shared immune defence systems is rapidly expanding with currently about a dozen identified antiviral systems, including cGAS, Viperin and TLRs[18]. Here we revealed striking structural similarity between human antiviral SAMD9/9L and prokaryotic Avs proteins[37], specifically with a new Avs protein family (Avs9). Notably, they both share the key nuclease domain, AlbA_2, which, in human SAMD9/9L, is responsible for transfer RNA^{Phe} cleavage and viral and cellular translational inhibition[5,9]. In Avs9 from *P. fluorescens*, we uncovered cell-killing activity, which also depended on AlbA_2 and its predicted nuclease site. Future biochemical and functional investigations will determine if Avs9 has bone fide nuclease activity—and which substrate—as well as identify if Avs9 has specific antiphage functions. Furthermore, similar to the Avs system, in which infected bacteria use an altruistic self-killing mechanism for the benefit of the colony[37], tight regulation of SAMD9/9L is probably essential and is certainly possible thanks to the intermediate and C-terminal domains[8].

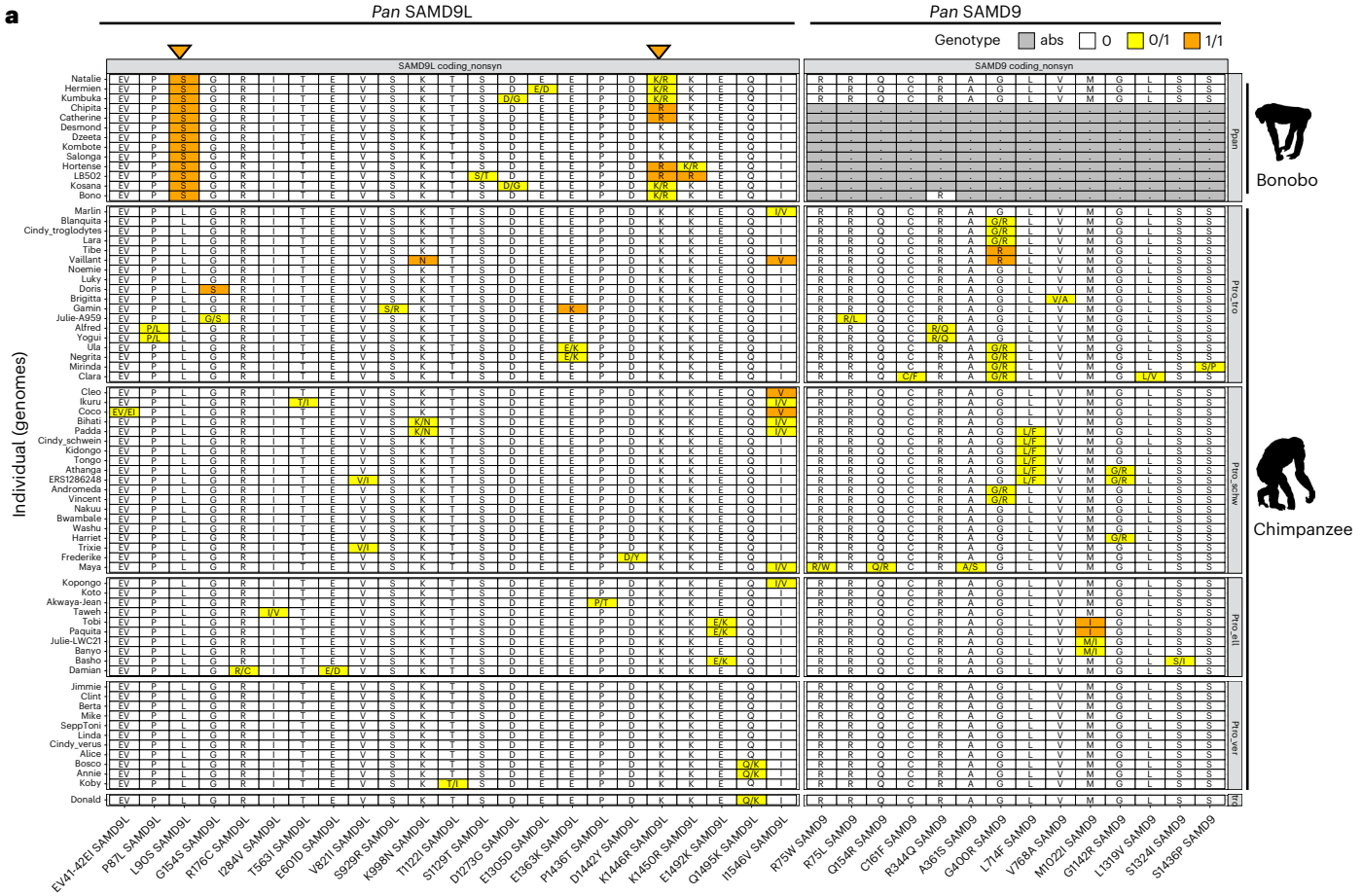
Uncovering phylogenetic relationships of prokaryotic Avs9 and metazoan SAMD9/9L protein families strongly suggest that they result from evolutionary convergence, possibly driven by analogous viral selective pressure. Moreover, domain shuffling is an important driver of protein evolution and can give rise to similar advantageous multidomain combinations. This work exemplifies a surprising evolutionary and functional parallel through the independent emergence of analogous antiviral proteins across vast evolutionary timescales.

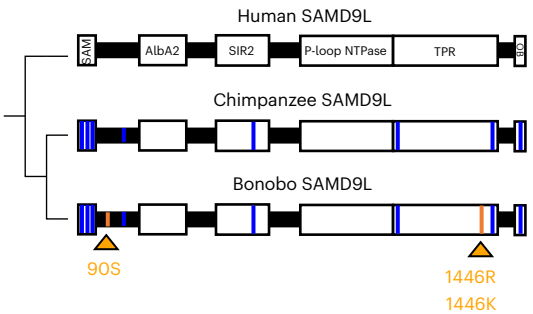
Extensive, and probably adaptative, *SAMD9/9L* CNVs

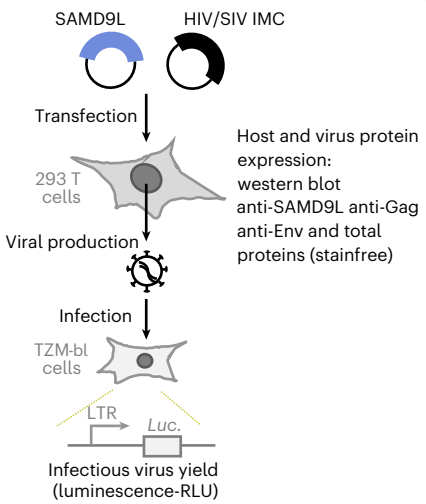
Despite its simultaneous presence in evolutionarily distant organisms, we globally found a patchy distribution of *SAMD9/9L* homologues and structural analogues across domains of life. For example, plants, fungi, algae and protozoa do not seem to harbour complete *SAMD9/9L* homologues. Furthermore, we found a very rapid evolution at the genomic and genetic levels during mammalian evolution in *SAMD9/9L*. Therefore, this ancient conservation is concomitant with rapid evolution, almost certainly as the result of virus–host arms races. Gene loss and duplication may, for example, provide an advantage, similar to observations in other innate immune genes, such as the APOBEC3 family[31,67–69]

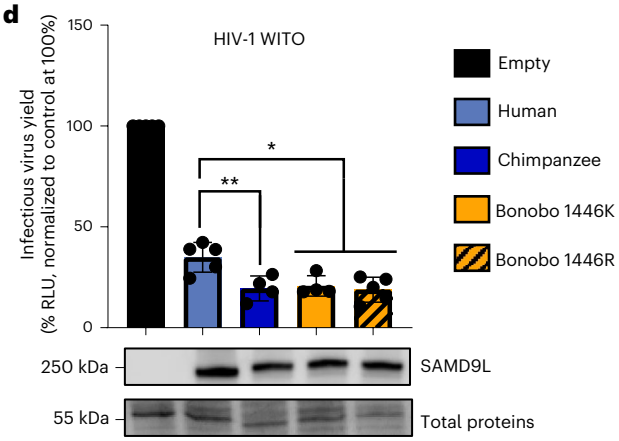
Fig. 5 | Chimpanzee and bonobo (*Pan*) SAMD9L major variants have an increased restrictive activity against HIV-1 pWITO, but not SIVcpz EK505, compared with human SAMD9L. a, Polymorphisms impacting SAMD9 and SAMD9L amino acid coding sequences among the *Pan* populations, with chimpanzee PanTro6 as reference. Highly frequent bonobo-specific SNPs in SAMD9L are highlighted by orange triangles at the top. See Supplementary Fig. 5 for details, human polymorphisms and comparative analyses with other primate species. **b**, Location of the chimpanzee- and bonobo-specific variants as compared with human SAMD9L on the two-dimensional predicted protein domain structure ($n = 8$ for chimpanzees versus humans and $n = 9$ or 10 for bonobos versus humans). **c**, Experimental setup to investigate chimpanzee and bonobo SAMD9L restriction on replication of full-length HIV-1 and SIVcpz IMCs. **d**, Relative infectious virus yields of HIV-1 pWITO in indicated SAMD9L conditions, normalized to the empty control (experimental setup in **c**). The two most frequent variants at position 1,446 in the bonobo populations were functionally tested. Results are from five independent biological replicates. Data are presented as mean ± s.d. Statistics were performed using the two-sided ratio paired *t*-test versus the human SAMD9L condition (**$P < 0.005$; *$P < 0.05$). Of note, all SAMD9Ls significantly restricted HIV-1 pWITO as compared with the empty control ($P < 0.005$). Below, western blot analysis in the producer cells showing similar protein expression of SAMD9Ls. Loading control is from total proteins (prestained gels). **e**, Similar experiment to **c** and **d**, with SIVcpz EK505 strain. Statistics were performed using the two-sided ratio paired *t*-test versus the human SAMD9L condition (NS, not significant) and versus the empty control (**$P < 0.005$). Luc, Luciferase; RLU, Relative light units.

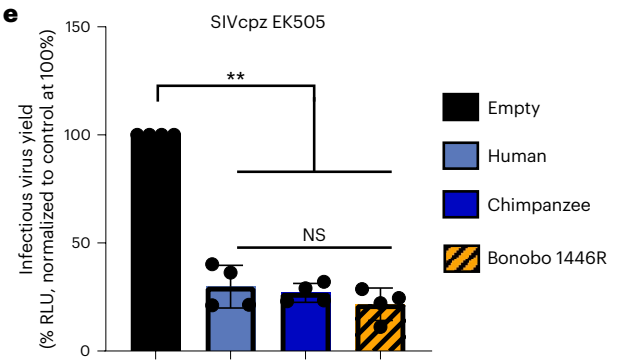
https://doi.org/10.1038/s41559-025-02845-x











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

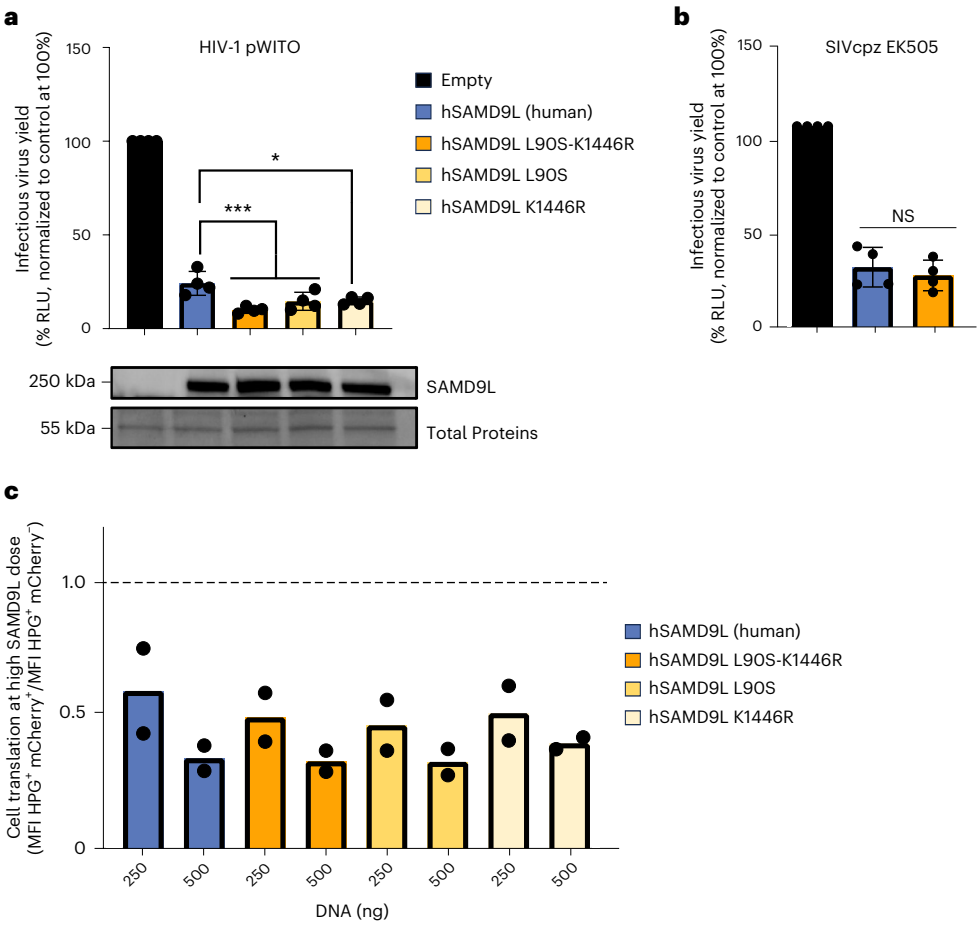


Fig. 6 | Bonobo-specific polymorphisms enhance human SAMD9L anti-HIV-1 activity, without affecting its translation shutdown effect. a,**b**, Relative infectious virus yields of HIV-1 pWITO (**a**) and SIVcpz EK505 (**b**) in indicated SAMD9L conditions (hSAMD9L (human SAMD9L)), normalized to the empty control, with four independent biological replicates. Data are presented as mean ± s.d. Statistics were performed using the two-sided ratio paired *t*-test versus the human SAMD9L condition (***$P < 0.0005$; *$P < 0.05$). Below, western blot analysis in the producer cells showing similar protein expression of SAMD9Ls. Loading control is from total proteins (prestained gels). In **b**, empty and hSAMD9L wild-type conditions are identical to Fig. 5e. **c**, Bonobo-specific SNPs do not modify human SAMD9L restriction of cellular translation. Quantification of protein synthesis assay was performed in two independent biological replicates in the context of ectopic expression of SAMD9Ls. Two doses of input DNA plasmids per condition were tested. HPG MFI ratio was calculated within each experimental condition using the MFI of the mCherry$^+$ cells (expressing mCherry-SAMD9L) normalized to the MFI of the mCherry$^-$ cells (not expressing SAMD9L). MFI, median fluorescence intensity.

and many other factors[1]. It is noteworthy that most genomic variations were gene losses, rather than extensive duplications, at least in most analysed mammalian species. These may be the result of evolutionary trade-offs, where the benefits of losing a gene (potentially escaping viral antagonism or hijacking) outweighed the costs.

At the mechanistic level, while most *SAMD9/9L* losses occurred by genomic loss of a chromosomal region, Colobinae primates seem to have lost *SAMD9* by early stop codons and pseudogenization. However, it cannot be excluded that, in some species, this may lead to the expression of a truncated SAMD9 retaining the AlbA_2 effector function but lacking the crucial regulatory intermediate and C-terminal domains, similar to some SAMD9L autoinflammatory gain-of-function variants, for example[8,61]. Further study on its selective pressure, and on messenger RNA transcript and protein expression in natural tissues from Colobinae species under immune stimulation would resolve this question.

The specific case of *SAMD9* unfixed loss in bonobos and adaptive chimpanzee and bonobo *SAMD9L*s: modern implications, functions and potential past lentiviral drivers

Bonobos harbour a recent unfixed loss of *SAMD9*, which occurred through a large chromosomal deletion. Despite bonobos, chimpanzees and humans being closely related, the variability observed at this locus is intriguing. Two bonobo-specific missense polymorphisms in SAMD9L that confer an increased antiviral activity against HIV-1 are located in the SAM and TPR domains, which could be involved in protein–protein or protein–RNA interactions[37,70–72]. One possibility is that the variants may modulate SAMD9L sensing, especially for the TPR domains, which have been reported to act as (viral) sensors in Avs and IFIT proteins (interferon-induced protein with TPRs)[37,73]. Further, although most human deleterious gain-of-function mutations in SAMD9/9L associated diseases are in the P-loop NTPase domain, some are described in the TPR or SAM domains[74]. However, we did not observe a gain-of-function phenotype on global cellular translation for the bonobo-specific variants. This therefore suggests that the variants may not destabilize the inactive closed form of the protein, nor change SAMD9L basal activities. Instead, it might modify its specificity and sensitivity in viral sensing, potentially adapting its interface with viruses. Otherwise, it may impact other potential functions of SAMD9L, for example in endosomal trafficking, increasing specific anti-HIV functions[5,10,11]. It would also be interesting to determine if, and how, chimpanzee and bonobo SAMD9Ls restrict other infections from poxviruses or other RNA viruses[3,12], and whether those may have driven some of the adaptations.

Our data show that chimpanzee and bonobo SAMD9Ls have an increased anti-HIV-1 phenotype compared with human. The additional loss of the pro-HIV-1 SAMD9 (ref. 5) may have been particularly advantageous (overall increased fitness) during past lentiviral infections in bonobo ancestors (that is, increased antiviral SAMD9L and loss of prolentiviral SAMD9). The presence of 3 of 13 unrelated bonobos with 1 *SAMD9*/*SAMD9L* allele and 1 *SAMD9L*-only allele suggests that *SAMD9* loss is unfixed, and probably recent. If selection favours individuals without *SAMD9*, the gene could eventually be completely lost in the bonobo population. The modern genomic makeup of the *SAMD9*/*9L* locus in chimpanzees, and even more so in bonobos, therefore suggests adaptation to lentiviral-like epidemics that occurred in *Pan*, as well as since the bonobo–chimpanzee divergence. In the future, performing this evo-functional study with a larger bonobo sample size, genomic phased data and long-read sequencing would enable robust determination of the exact selective pressures shaping antiviral defence mechanisms in this species. It would also help in determining the haplotype structures on which the *SAMD9* gene was lost and the *SAMD9L* substitutions occurred, potentially providing further insight into the epistatic interactions between these genetic changes.

Unlike some chimpanzees and humans that are infected by SIVcpz and HIVs, respectively, and suffer from AIDS symptoms[75–77], modern bonobos are not known to be naturally infected by any lentiviruses[25,26]. Overall, SAMD9/9L adaptation may nowadays participate, with other factors[78], in bonobo population resistance against lentiviral/SIV infections.

Finally, it is noteworthy that in this study SIVcpzPtt EK505 did not show an increased sensitivity to chimpanzee and bonobo SAMD9Ls, or to *Pan* SNPs in the context of human SAMD9L. This may be the result of virus–host co-evolution[1,30], where SIVcpz has adapted to the natural genetic makeup of its host, particularly of chimpanzee antiviral innate immunity.

Altogether, our findings highlight the strength of evo-immuno approaches in unravelling links between the evolutionary history of innate immunity and contemporary challenges in human health. The identification of SAMD9/9L homologues and structural–functional analogues across diverse taxa, as prokaryotes and primates, shows a shared convergent immunity. Common challenges, such as fighting viral infections, drive both conservation or convergence of key immune systems as well as their rapid evolution through arms races. In this regard, using diverse models (human, diverse eukaryotic cells and bacteria) and natural variants in closely related species for functional studies can bring valuable insights with broader medical applications, such as the incorporation of potentiator mutations in antiviral factors (protein engineering) or the use of bacterial antiviral proteins that could act against human viruses.

Methods

Comparative genomics, phylogenetics and positive selection analyses in mammals

To obtain the coding sequences of the *SAMD9* and *SAMD9L* homologues in bats, rodents, primates, ungulates and carnivores, we used the detection of genetic innovations (DGINN) pipeline[53] with, respectively, *Myotis myotis, Rattus norvegicus, Homo sapiens, Hippopotamus amphibius* and *Phoca vitulina* Refseq *SAMD9* and *SAMD9L*, as queries. Briefly, the coding sequences from each group were automatically retrieved with NCBI blastn[79,80], cleaned and aligned with MAFFT[81]. Homologous sequences from marsupials, aves and amphibians were retrieved using NCBI blastn. Of note, these analyses are based on publicly available genome annotations (not necessarily genes annotated as '*SAMD9/9L*-like' but regions annotated as coding regions), so it is not excluded that some unannotated genes were not analysed. The species and accession numbers are presented in Supplementary Table 2. Nucleotide alignments of *SAMD9* or *SAMD9L* from each mammalian group were manually curated before being used as input in DGINN for automatic codon alignment using the probabilistic alignment kit (PRANK)[82] and phylogenetic tree building using PhyML[83] (with default settings in DGINN).

Furthermore, all codon alignments, as well as outgroup sequences, were realigned in a three-step fashion to obtain a high-quality mammalian-wide codon alignment (1) using Muscle[84], (2) manually curating the sequences, (3) codon aligned with PRANK. A phylogenetic tree was inferred from this alignment using IQ-TREE webserver (GTR + F + I + G4 identified as the best substitution model by ModelFinder implemented in IQ-TREE)[85]. We also performed analyses with PhyML (best model estimated from Smart model selection, SMS: GTR + R). These analyses allowed us to attribute the phylogenetically aware *'SAMD9'* or *'SAMD9L'* nomenclature.

We next tested for positive selection occurring at specific branches using the aBSREL program on the DataMonkey webserver[57]. We used as inputs the codon alignments of (1) primate *SAMD9* with *Cricetulus griseus* (criGri) and *Tupaia chinensis* (tupChi) *SAMD9*s as outgroups and (2) primate *SAMD9L* with *Cricetulus griseus* (criGri) *SAMD9L* as outgroup. Each branch that we tested for evidence of positive selection was defined as 'tested branch' (that is, 'foreground') and the remaining as 'background branches'. Two models were fit to each tested branch: one that allows for episodic diversifying selection (with $\omega > 1$) and one that does not. A likelihood ratio test is then used to compare these models and assess whether the tested branch shows evidence of positive selection. For branches where selection is detected, aBSREL also estimates the proportion of codon sites that are subject to positive selection.

For cases in which we suspected gene losses, we confirmed the absence of coding genes by several methods. We analysed the *SAMD9*/*9L* genomic locus (between *HEPACAM2* and *CDK6* syntenic genes) on NCBI genome data viewer of specific species. We verified that there were no missing data or low sequence quality in this genomic region. Pseudogenes, here identified by several and very early stop codons, were only analysed systematically in primates. For the two monotreme species, in which no *SAMD9*/*9L* homologues could be retrieved by genome-wide blast, the HEPACAM2-CDK6 genomic locus was retrieved. We found no missing data (no 'N') and no homology using blast or alignments (with relaxed parameters) of non-annotated regions with human *SAMD9*/*9L*.

Genome alignment and SNPs analyses in hominids

The genomic sequences of 13 bonobos, 59 chimpanzees and 4,099 humans were retrieved from public online databases: 1000 Genomes Project, Human Genome Diversity Project (HGDP), NCBI bioprojects PRJNA189439, SRP018689 and PRJEB15083 (refs. 60,86) (The 1000 Genomes Project Consortium 2015). DNA sequences from *Pan* individuals were aligned to Clint_PTRv2/panTro6 reference genome using BWA-MEM[87]. Then, variant calling was done using FreeBayes[88] to obtain a vcf file. The *SAMD9* and *SAMD9L* locus regions (chromosome 7: 88599930–89350079 in panTro6) were extracted and parsed using an ad hoc R script using VariantAnnotation, GenomicFeatures, AnnotationHub, org.Pt.eg.db, ggplot2, R packages. This script was used to identify non-synonymous variants among SNPs and to visualize them. The equivalent genomic region in human was retrieved by using the LiftOver tool of the UCSC genome browser (chromosome 7: 92540798–93289621 in the human reference GRCh38/hg38). These coordinates were then used to subset the HGDP + 1000 Genomes Project vcf for this region.

Structure similarity search, structurally aware alignment and phylogenetic analyses across kingdoms of life

Protein structures were obtained from AlphaFold DB[89,90] and RCSB PDB[91]. Foldseek[33] was used for detection of structural similarity. We used a sequential strategy. First, we queried Foldseek v.427df8a with SAMD9 (PDB ID Q5K651) and SAMD9L (Q8IVG5) against the AlphaFold database clustered at 50% sequence identity (AFDB50),

using a 30% TM-score threshold and a maximum *E*-value of 0.0001 (default settings). Then, we constructed a query database consisting of SAMD9 (Q5K651), SAMD9L (Q8IVG5) and bacterial top hits—AVAST type V (A0A1F9N8W4) and Avs9 (A0A100VJR7, A0A2S4Y961 and A0A7T4VS34)—and used it for searches under the same settings. Search hits were subsequently filtered for a minimum 80% query coverage to cover at least 1,000 amino acids of the query structures, resulting in 238 analogue structures. FoldMason v.333d54c (ref. 92) was used to generate a multiple structure alignment (MSTA) of the identified structural analogues and MUSCLE v.5 was used to generate a multiple amino acid sequence alignment (MSA). The domain coordinates used for presence/absence analysis of each domain or for the extraction of a given domain MSTA are presented in Supplementary Fig. 1c and are based on the predicted SAMD9 three-dimensional structure. Additionally, using FoldMason, an MSTA was generated on the human SLFNs (SLFNL1, SLFN5 and SLFN11–14, respectively, corresponding to PDB IDs Q499Z3, Q08AF3, Q7Z7L1, Q8IYM2, Q68D06 and P0C7P3) with the 23 hits over 238 containing an AlbA_2 domain. Phylogenetic trees were constructed from both the MSA and MSTA using IQ-TREE 2.3.0 (ref. 93) with the LG + F + G4 substitution model and 1,000 bootstrap replicates, and visualized using the ggtree R library[94] or interactive tree of life (iTol)[95].

Phylogenetic analysis of prokaryotic and eukaryotic AlbA_2 domains

The HMM profile of the Pfam AlbA_2 (PF04326) domain was retrieved from the Pfam database[96]. This profile was searched against a custom protein database combining: (1) 41,150 complete bacterial genomes downloaded from Refseq in August 2024, filtered for redundancy using the clusthash function of MMseqs2 (v.13.45111) using default parameters[97]; (2) 455 complete archaeal genomes downloaded from Refseq in August 2024; (3) 993 representative eukaryotic genomes from the EukProt database[98], filtered for redundancy using the clusthash function of MMseqs2 (13.45111) using default parameters. The AlbA_2 HMM profile was searched into this combined protein database using hmmsearch (v.3.3.2) with default parameters[99]. Hits with at least 90 covered profile residues were selected and the amino acid sequences of the aligned regions complemented with ten residues on each side were extracted. Sequences were clustered using the easy-cluster function of MMseqs2 (v.13.45111) with parameter --min-seq-id 0.8 (ref. 97). Cluster representatives were aligned with Clustal-Omega[100] with default parameters and the alignment was trimmed using ClipKit[101]. The trimmed alignment was used to compute a tree using IQ-TREE with parameters -m L -bb 10000 -nm 10000, which was then visualized using iTOL.

For each genome, defence systems were detected using DefenseFinder (v.1.3)[34]. For each clade, we calculated a defence score as the fraction of bacterial genes found within ten genes upstream or downstream of a defence protein as detected by DefenseFinder in their genome of origin. As a control, we calculated for each clade the defence score expected by chance as the fraction of all genes (from the same genomes) that are found within ten genes upstream or downstream of a defence protein. This value was used to assess for each clade whether AlbA_2-containing genes colocalize with antiphage systems more frequently than expected by chance, using a binomial test and false discovery rate (FDR) correction (Supplementary Table 1).

Bacterial strains and plasmids

The codon-optimized open reading frame encoding Avs9 from *P. fluorescens* (Uniprot ID A0A7T4VS34) was ordered as a gene fragment from Twist Bioscience, cloned into the pBbA6c vector[102] by T5 exonuclease-dependent assembly cloning[103] and transformed into *E. coli* DH5α λpir. Constructions were sequence-verified by Sanger sequencing (Microsynth). The complete Avs9 gene fragment sequence synthesized and used in this study is provided in Data availability. We further made the Avs9-D45N mutant by site-directed mutagenesis using polymerase chain reaction amplification with Q5 polymerase (New England Biolabs) and KLD cloning (New England Biolabs).

Bacterial drop assays

E. coli DH5α λpir cells carrying pBbA6c, pBbA6c-Avs9 or pBbA6c-Avs^{D45N} were grown for 6 h at 37 °C and 180 rpm in Luria–Bertani (LB) medium supplemented with chloramphenicol at 20 μg ml^{-1} and glucose at 1 g ml^{-1}. After tenfold serial dilutions, 5 μl of each dilution were spotted on LB agar plates supplemented with 20 μg ml^{-1} of chloramphenicol and either 1% glucose or 100–500 μM isopropyl β-D-1-thiogalactopyranoside (IPTG). Plates were incubated at 37 and 25 °C for 24 and 48 h, respectively. Strong toxicity of pBbA6c-Avs9 was observed upon incubation at 25 °C.

Plasmids for expression in human cells

HIV-1 T/F pWITO (human immunodeficiency virus 1 pWITO.c/2474, ARP-11739) encoding a full-length IMC was contributed by J. Kappes and C. Ochsenbauer through the National Institutes of Health (NIH) AIDS repository programme. IMC for SIVcpz EK505 was a gift from B. Hahn[27,65]. The pMT06-Flag-mCherry-SAMD9L plasmid was constructed by cloning the synthetized human *SAMD9L* gene into a pMT06-Flag-mCherry backbone, from the original RRL.sin.cPPT.CMV/Flag-E2-crimson.IRES-puro.WPRE (MT06, a gift from C. Goujon: Addgene plasmid no. 139448; http://n2t.net/addgene:139448; RRID: Addgene_139448)[104]. The pMT06-Flag-mCherry-chimp SAMD9L plasmid (chimpanzee SAMD9L) was synthesized and cloned by Azenta Genewiz. The bonobo SAMD9L plasmids were generated through site-directed mutagenesis of the pMT06-FLAG-mCherry-chimp SAMD9L plasmid. Human SAMD9L-L90S/K1446R, SAMD9L-L90S and SAMD9L-K1446R (double and single mutant) plasmids were generated from pMT06-Flag-mCherry-SAMD9L plasmid, using the QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent) following the manufacturer's instructions. Sequences were confirmed through full-length plasmid and/or Sanger sequencing (Microsynth).

Cell lines and culture

Human embryonic kidney 293 T cell lines (ATCC, catalogue no. CRL-3216) and TZM-bl (NIH AIDS Research and Reference Reagent Program, catalogue no. 8129) were grown in Dulbecco modified Eagle medium containing 10% fetal calf serum (Sigma catalogue no. F7524) and 100 U ml^{-1} of penicillin/streptomycin. TZM-bl cells express the cell-surface proteins CD4, CCR5 and CXCR4 and encode for luciferase and β-galactosidase under the control of the long-terminal repeat (LTR) promoter. They are commonly used for lentiviral titration of culture supernatants (Tzm-bl assays).

Production and quantification of replication-competent lentivirus

A total of 293 T cells were initially seeded in 6-well plates at a density of 0.2 M cells ml^{-1} (400,000 cells total per well). After 24 h, the cells were cotransfected using TransIT-LT1 (Mirus) with a plasmid encoding a fully replication-competent lentivirus (IMCs), alongside either a plasmid encoding SAMD9L or an empty control. The quantity of DNA used was 250 ng for host plasmids and 1,200 ng for virus plasmids. Subsequently, 48 h post-transfection, cells were harvested for western blot. The supernatants were collected and stored at −80 °C for further titration of infectious virus yield via TZM-bl cells. For titration, TZM-bl cells were plated in 96-well plates and exposed to serial dilutions of viral supernatant. Following 48 h of infection, cell lysis was performed using BrightGlow Lysis Reagent (Promega E2620) and relative light units were measured using the Tecan Spark Luminometer. Infectious virus yields under various conditions were consistently expressed as fold-change compared with paired viral infection conditions in the absence of SAMD9L.

Western blot analysis

Cells were harvested and lysed using ice-cold RIPA buffer (composed of 50 mM Tris pH 8, 150 mM NaCl, 2 mM EDTA and 0.5% NP40) supplemented with protease inhibitors (Roche), followed by sonication. Proteins from cell lysates or supernatants were separated by electrophoresis and transferred onto a PVDF membrane via overnight wet transfer at 4 °C. Stain-Free gel (BioRad) was used for loading and protein transfer controls. Following blocking in TBS-T 1X solution (Tris buffer saline, consisting of Tris HCl 50 mM pH 8, NaCl 30 mM and 0.05% Tween 20) with 5% powdered milk, the membranes underwent incubation with primary antibodies for a duration ranging from 1 h to overnight, followed by subsequent 1-h incubation with secondary antibodies. Detection was carried out using SuperSignal West Pico Chemiluminescent Substrate (ThermoFisher Scientific) and imaged using the Chemidoc Imaging System (BioRad). Antibodies used included anti-SAMD9L (Proteintech, 25173-1-AP), anti-Gag (NIH HIV Reagent Program, 183-H12-5C), anti-HIV-1-gp120 (Aalto, D7324; NIH HIV Reagent Program, 16H3), as well as secondary IgG-peroxidase conjugated anti-mouse (Sigma, catalogue no. A9044) and anti-rabbit (Sigma, catalogue no. AP188P). 'Total proteins were used as a loading control with BioRad stain-free gel.

Protein synthesis assay

A total of 293 T cells were seeded at 0.2 M cells ml^{-1} in 12-well plates (200,000 cells total per well). At 24 h after seeding, cells were transfected with 250 or 500 ng of host DNA plasmid, using TransIT-LT1 (Mirus) following the manufacturer's instructions. At 48 h post-transfection, cells were incubated in HPG for 30 min at 37 °C. Medium was discarded and cells were harvested and fixed with PFA 4%. Cells were then washed with PBS BSA 3% and permeabilized in PBS 0.5% Triton X-100 for 15 min. Click-iT Plus Alexa Fluor Picolyl Azide assay was then performed and cells were analysed on MACSQuant VYB Cytometer (Miltenyi Biotec, SFR BioSciences).

Other softwares and statistical analyses

DefenseFinder[34] was used to analyse sequences from Foldseek analyses[33]. Sequencing analyses and representations were conducted using Geneious (Biomatters), ESPript 3.0 https://espript.ibcp.fr (ref. 105) and UGENE v.52.0 (ref. 106). R scripts were used to conduct analyses of genomic data. Graphic representations and statistical analyses were carried out using GraphPad Prism 9 and R scripts. In the figures, data are presented as mean ± s.d. and each point correspond to an independent biological replicate. Statistics were performed using the two-ratio paired *t*-test, except where indicated in the figure legends. *P* values are according to figure legends and exact P values are reported in Supplementary Table 1 and Source data.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All data and reagents are available in this manuscript (including Supplementary Information, Supplementary Tables and Supplementary Data) or accessible upon request to the corresponding author. The alignments and phylogenetic trees are all openly available via figshare (phylogenetics associated: with Fig. 1 and Supplementary Fig. 1 at https://doi.org/10.6084/m9.figshare.29082710.v1 (ref. 107); with Fig. 2 and Supplementary Fig. 2 at https://doi.org/10.6084/m9.figshare.29082710 (ref. 108); with Fig. 3 and Supplementary Fig. 3 at https://doi.org/10.6084/m9.figshare.29042762 (ref. 109); with Fig. 4 and Supplementary Fig. 4 at https://doi.org/10.6084/m9.figshare.29087987 (ref. 110); and with Supplementary Fig. 5c at https://doi.org/10.6084/m9.figshare.29088050 (ref. 111). The Avs9 sequence used for experimental assays in Fig. 2d and Supplementary Fig. 2d is available via figshare at https://doi.org/10.6084/m9.figshare.29082719 (ref. 112). Source data are provided with this paper.

Code availability

The scripts for the analyses of the polymorphic sites are available via GitLab at http://gitbio.ens-lyon.fr/ciri/lp2l/vcf_to_nice_figures.git.

References

1. Tenthorey, J. L., Emerman, M. & Malik, H. S. Evolutionary landscapes of host–virus arms races. *Annu. Rev. Immunol.* **40**, 271–294 (2022).
2. Georjon, H. & Bernheim, A. The highly diverse antiphage defence systems of bacteria. *Nat. Rev. Microbiol.* **21**, 686–700 (2023).
3. Liu, J. & McFadden, G. SAMD9 is an innate antiviral host factor with stress response properties that can be antagonized by poxviruses. *J. Virol.* **89**, 1925–1931 (2014).
4. Meng, X. et al. A paralogous pair of mammalian host restriction factors form a critical host barrier against poxvirus infection. *PLoS Pathog.* **14**, e1006884 (2018).
5. Legrand, A. et al. SAMD9L acts as an antiviral factor against HIV-1 and primate lentiviruses by restricting viral and cellular translation. *PLoS Biol.* **22**, e3002696 (2024).
6. Meng, X. & Xiang, Y. RNA granules associated with SAMD9-mediated poxvirus restriction are similar to antiviral granules in composition but do not require TIA1 for poxvirus restriction. *Virology* **529**, 16–22 (2019).
7. Sivan, G., Glushakow-Smith, S. G., Katsafanas, G. C., Americo, J. L. & Moss, B. Human host range restriction of the vaccinia virus C7/K1 double deletion mutant Is mediated by an atypical mode of translation inhibition. *J. Virol.* **92**, e01329-18 (2018).
8. Russell, A. J. et al. SAMD9L autoinflammatory or ataxia pancytopenia disease mutations activate cell-autonomous translational repression. *Proc. Natl Acad. Sci. USA* **118**, e2110190118 (2021).
9. Zhang, F. et al. Human SAMD9 is a poxvirus-activatable anticodon nuclease inhibiting codon-specific protein synthesis. *Sci. Adv.* **9**, eadh8502 (2023).
10. Nagamachi, A. et al. Haploinsufficiency of SAMD9L, an endosome fusion facilitator, causes myeloid malignancies in mice mimicking human diseases with monosomy 7. *Cancer Cell* **24**, 305–317 (2013).
11. Narumi, S. et al. SAMD9 mutations cause a novel multisystem disorder, MIRAGE syndrome, and are associated with loss of chromosome 7. *Nat. Genet.* **48**, 792–797 (2016).
12. Hou, G. et al. SAMD9 senses cytosolic double-stranded nucleic acids in epithelial and mesenchymal cells to induce antiviral immunity. *Nat. Commun.* **16**, 3756 (2025).
13. de Jesus, A. A. et al. Distinct interferon signatures and cytokine patterns define additional systemic autoinflammatory diseases. *J. Clin. Invest.* **130**, 1669–1682 (2020).
14. Gorcenco, S. et al. Ataxia-pancytopenia syndrome with SAMD9L mutations. *Neurol. Genet.* **3**, e183 (2017).
15. Topaz, O. et al. A deleterious mutation in *SAMD9* causes mormophosphatemic familial tumoral calcinosis. *Am. J. Hum. Genet.* **79**, 759–764 (2006).
16. Wein, T. & Sorek, R. Bacterial origins of human cell-autonomous innate immune mechanisms. *Nat. Rev. Immunol.* **22**, 629–638 (2022).
17. Rousset, F. Innate immunity: the bacterial connection. *Trends Immunol.* **44**, 945–953 (2023).
18. Bernheim, A., Cury, J. & Poirier, E. Z. The immune modules conserved across the tree of life: towards a definition of ancestral immunity. *PLoS Biol.* **22**, e3002717 (2024).
19. Ledvina, H. E. & Whiteley, A. T. Conservation and similarity of bacterial and eukaryotic innate immunity. *Nat. Rev. Microbiol.* **22**, 420–434 (2024).

20. Mekhedov, S. L., Makarova, K. S. & Koonin, E. V. The complex domain architecture of SAMD9 family proteins, predicted STAND-like NTPases, suggests new links to inflammation and apoptosis. *Biol. Direct* **12**, 13 (2017).
21. Peng, S. et al. Structure and function of an effector domain in antiviral factors and tumor suppressors SAMD9 and SAMD9L. *Proc. Natl Acad. Sci. USA* **119**, e2116550119 (2022).
22. Lemos de Matos, A., Liu, J., McFadden, G. & Esteves, P. J. Evolution and divergence of the mammalian SAMD9/SAMD9L gene family. *BMC Evol. Biol.* **13**, 121 (2013).
23. Peeters, M. et al. Risk to human health from a plethora of simian immunodeficiency viruses in primate bushmeat. *Emerg. Infect. Dis.* **8**, 451–457 (2002).
24. Chahroudi, A., Bosinger, S. E., Vanderford, T. H., Paiardini, M. & Silvestri, G. Natural SIV hosts: showing AIDS the door. *Science* **335**, 1188–1193 (2012).
25. Li, Y. et al. Eastern chimpanzees, but not bonobos, represent a simian immunodeficiency virus reservoir. *J. Virol.* **86**, 10776–10791 (2012).
26. Van Dooren, S. et al. Lack of evidence for infection with simian immunodeficiency virus in bonobos. *AIDS Res. Hum. Retrovir.* **18**, 213–216 (2002).
27. Keele, B. F. et al. Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. *Science* **313**, 523–526 (2006).
28. Neil, S. J. D., Zang, T. & Bieniasz, P. D. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature* **451**, 425–430 (2008).
29. Sawyer, S. L., Wu, L. I., Emerman, M. & Malik, H. S. Positive selection of primate TRIM5α identifies a critical species-specific retroviral restriction domain. *Proc. Natl Acad. Sci. USA* **102**, 2832–2837 (2005).
30. Duggal, N. K. & Emerman, M. Evolutionary conflicts between viruses and restriction factors shape immunity. *Nat. Rev. Immunol.* **12**, 687–695 (2012).
31. Etienne, L. et al. The role of the antiviral APOBEC3 gene family in protecting chimpanzees against lentiviruses from monkeys. *PLoS Pathog.* **11**, e1005149 (2015).
32. Compton, A. A., Malik, H. S. & Emerman, M. Host gene evolution traces the evolutionary history of ancient primate lentiviruses. *Philos. Trans. R. Soc. B* **368**, 20120496 (2013).
33. van Kempen, M. et al. Fast and accurate protein structure search with Foldseek. *Nat. Biotechnol.* **42**, 243–246 (2024).
34. Tesson, F. et al. A comprehensive resource for exploring antiphage defense: DefenseFinder Webservice, Wiki and Databases. *Peer Community J.* **4**, e91 (2024).
35. Wang, Y., Tian, Y., Yang, X., Yu, F. & Zheng, J. Filamentation activates bacterial Avs5 antiviral protein. *Nat. Commun.* **16**, 2408 (2025).
36. Béchon, N. et al. Diversification of molecular pattern recognition in bacterial NLR-like proteins. *Nat. Commun.* **15**, 9860 (2024).
37. Gao, L. A. et al. Prokaryotic innate immunity through pattern recognition of conserved viral proteins. *Science* **377**, eabm4096 (2022).
38. Gao, L. et al. Diverse enzymatic activities mediate antiviral immunity in prokaryotes. *Science* **369**, 1077–1084 (2020).
39. Rousset, F. et al. A conserved family of immune effectors cleaves cellular ATP upon viral infection. *Cell* **186**, 3619–3631 (2023).
40. Garb, J. et al. Multiple phage resistance systems inhibit infection via SIR2-dependent NAD+ depletion. *Nat. Microbiol.* **7**, 1849–1856 (2022).
41. Schultz, J., Ponting, C. P., Hofmann, K. & Bork, P. SAM as a protein interaction domain involved in developmental regulation. *Protein Sci. Publ. Protein Soc.* **6**, 249–253 (1997).
42. Xu, J. & Zhang, Y. How significant is a protein structure similarity with TM-score=0.5? *Bioinformatics* **26**, 889–895 (2010).
43. Leipe, D. D., Koonin, E. V. & Aravind, L. STAND, a class of P-loop NTPases including animal and plant regulators of programmed cell death: multiple, complex domain architectures, unusual phyletic patterns, and evolution by horizontal gene transfer. *J. Mol. Biol.* **343**, 1–28 (2004).
44. Loquet, A. & Saupe, S. J. Diversity of amyloid motifs in NLR signaling in fungi. *Biomolecules* **7**, 38 (2017).
45. Mordier, J., Fraisse, M., Cohen-Tannoudji, M. & Molaro, A. Recurrent evolutionary innovations in rodent and primate Schlafen genes. Preprint at *bioRxiv* https://doi.org/10.1101/2024.01.12.575368 (2024).
46. Doron, S. et al. Systematic discovery of anti-phage defense systems in the microbial pan-genome. *Science* **359**, eaar4120 (2018).
47. Valenzuela, C., Saucedo, S. & Llano, M. Schlafen14 impairs HIV-1 expression in a codon usage-dependent manner. *Viruses* **16**, 502 (2024).
48. Kobayashi-Ishihara, M. et al. Schlafen 12 restricts HIV-1 latency reversal by a codon-usage dependent post-transcriptional block in CD4+ T cells. *Commun. Biol.* **6**, 487 (2023).
49. Metzner, F. J., Huber, E., Hopfner, K.-P. & Lammens, K. Structural and biochemical characterization of human Schlafen 5. *Nucleic Acids Res.* **50**, 1147–1161 (2022).
50. Garvie, C. W. et al. Structure of PDE3A-SLFN12 complex reveals requirements for activation of SLFN12 RNase. *Nat. Commun.* **12**, 4375 (2021).
51. Yang, J.-Y. et al. Structure of Schlafen13 reveals a new class of tRNA/rRNA-targeting RNase engaged in translational control. *Nat. Commun.* **9**, 1165 (2018).
52. Li, M. et al. DNA damage-induced cell death relies on SLFN11-dependent cleavage of distinct type II tRNAs. *Nat. Struct. Mol. Biol.* **25**, 1047–1058 (2018).
53. Picard, L. et al. DGINN, an automated and highly-flexible pipeline for the detection of genetic innovations on protein-coding genes. *Nucleic Acids Res.* **48**, e103 (2020).
54. Wellenreuther, M., Mérot, C., Berdan, E. & Bernatchez, L. Going beyond SNPs: the role of structural genomic variants in adaptive evolution and species diversification. *Mol. Ecol.* **28**, 1203–1209 (2019).
55. Radke, D. W. & Lee, C. Adaptive potential of genomic structural variation in human and mammalian evolution. *Brief. Funct. Genomics* **14**, 358–368 (2015).
56. Daugherty, M. D. & Malik, H. S. Rules of engagement: molecular insights from host–virus arms races. *Annu. Rev. Genet.* **46**, 677–700 (2012).
57. Smith, M. D. et al. Less is more: an adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Mol. Biol. Evol.* **32**, 1342–1353 (2015).
58. Prüfer, K. et al. The bonobo genome compared with the chimpanzee and human genomes. *Nature* **486**, 527–531 (2012).
59. Mao, Y. et al. A high-quality bonobo genome refines the analysis of hominid evolution. *Nature* **594**, 77–81 (2021).
60. Prado-Martinez, J. et al. Great ape genetic diversity and population history. *Nature* **499**, 471–475 (2013).
61. Allenspach, E. J. et al. Germline SAMD9L truncation variants trigger global translational repression. *J. Exp. Med.* **218**, e20201195 (2021).
62. Thomas, M. E. et al. Pediatric MDS and bone marrow failure-associated germline mutations in SAMD9 and SAMD9L impair multiple pathways in primary hematopoietic cells. *Leukemia* **35**, 3232–3244 (2021).
63. Locatelli, S. et al. Why are Nigeria-Cameroon chimpanzees (*Pan troglodytes ellioti*) free of SIVcpz infection? *PLoS ONE* **11**, e0160788 (2016).
64. Gao, F. et al. Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes*. *Nature* **397**, 436–441 (1999).

65. Barbian, H. J. et al. Neutralization properties of simian immunodeficiency viruses infecting chimpanzees and gorillas. *mBio* **6**, e00296-15 (2015).
66. Burroughs, A. M. & Aravind, L. Identification of uncharacterized components of prokaryotic immune systems and their diverse eukaryotic reformulations. *J. Bacteriol*. https://doi.org/10.1128/jb.00365-20 (2020).
67. Sawyer, S. L., Emerman, M. & Malik, H. S. Ancient adaptive evolution of the primate antiviral DNA-editing enzyme APOBEC3G. *PLoS Biol*. **2**, e275 (2004).
68. Münk, C., Willemsen, A. & Bravo, I. G. An ancient history of gene duplications, fusions and losses in the evolution of APOBEC3 mutators in mammals. *BMC Evol. Biol*. **12**, 71 (2012).
69. Ito, J., Gifford, R. J. & Sato, K. Retroviruses drive the rapid evolution of mammalian APOBEC3 genes. *Proc. Natl Acad. Sci. USA* **117**, 610–618 (2020).
70. Kim, C. A. & Bowie, J. U. SAM domains: uniform structure, diversity of function. *Trends Biochem. Sci*. **28**, 625–628 (2003).
71. Knight, M. J., Leettola, C., Gingery, M., Li, H. & Bowie, J. U. A human sterile alpha motif domain polymerizome. *Protein Sci*. **20**, 1697–1706 (2011).
72. Johnson, B. et al. Human IFIT3 modulates IFIT1 RNA binding specificity and protein stability. *Immunity* **48**, 487–499 (2018).
73. Fleith, R. C. et al. IFIT3 and IFIT2/3 promote IFIT1-mediated translation inhibition by enhancing binding to non-self RNA. *Nucleic Acids Res*. **46**, 5269–5285 (2018).
74. Sahoo, S. S. et al. Clinical evolution, genetic landscape and trajectories of clonal hematopoiesis in SAMD9/SAMD9L syndromes. *Nat. Med*. **27**, 1806–1817 (2021).
75. Keele, B. F. et al. Increased mortality and AIDS-like immunopathology in wild chimpanzees infected with SIVcpz. *Nature* **460**, 515–519 (2009).
76. Etienne, L. et al. Characterization of a new simian immunodeficiency virus strain in a naturally infected *Pan troglodytes troglodytes* chimpanzee with AIDS related symptoms. *Retrovirology* **8**, 4 (2011).
77. Greenwood, E. J. D. et al. Simian immunodeficiency virus infection of chimpanzees (*Pan troglodytes*) shares features of both pathogenic and non-pathogenic lentiviral infections. *PLoS Pathog*. **11**, e1005146 (2015).
78. Pawar, H., Ostridge, H. J., Schmidt, J. M. & Andrés, A. M. Genetic adaptations to SIV across chimpanzee populations. *PLoS Genet*. **18**, e1010337 (2022).
79. Camacho, C. et al. BLAST+: architecture and applications. *BMC Bioinf*. **10**, 421 (2009).
80. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol*. **215**, 403–410 (1990).
81. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol*. **30**, 772–780 (2013).
82. Löytynoja, A. Phylogeny-aware alignment with PRANK. *Methods Mol. Biol*. **1079**, 155–170 (2014).
83. Guindon, S. et al. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol*. **59**, 307–321 (2010).
84. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. **32**, 1792–1797 (2004).
85. Trifinopoulos, J., Nguyen, L.-T., von Haeseler, A. & Minh, B. Q. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res*. **44**, W232–W235 (2016).
86. Bergström, A. et al. Insights into human genetic variation and population history from 929 diverse genomes. *Science* **367**, eaay5012 (2020).
87. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Preprint at https://doi.org/10.48550/arXiv.1303.3997 (2013).
88. Garrison, E. & Marth, G. Haplotype-based variant detection from short-read sequencing. Preprint at https://doi.org/10.48550/arXiv.1207.3907 (2012).
89. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
90. Varadi, M. et al. AlphaFold protein structure database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res*. **50**, D439–D444 (2022).
91. Berman, H. M. et al. The protein data bank. *Nucleic Acids Res*. **28**, 235–242 (2000).
92. Gilchrist, C. L. M., Mirdita, M. & Steinegger, M. Multiple protein structure alignment at scale with FoldMason. Preprint at *bioRxiv* https://doi.org/10.1101/2024.08.01.606130 (2024).
93. Minh, B. Q. et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol*. **37**, 1530–1534 (2020).
94. Yu, G., Smith, D. K., Zhu, H., Guan, Y. & Lam, T. T.-Y. ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evol*. **8**, 28–36 (2017).
95. Letunic, I. & Bork, P. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Res*. **52**, W78–W82 (2024).
96. Paysan-Lafosse, T. et al. The Pfam protein families database: embracing AI/ML. *Nucleic Acids Res*. **53**, D523–D534 (2025).
97. Steinegger, M. & Söding, J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat. Biotechnol*. **35**, 1026–1028 (2017).
98. Richter, D. J. et al. EukProt: a database of genome-scale predicted proteins across the diversity of eukaryotes. *Peer Community J*. **2**, evae231 (2022).
99. Eddy, S. R. Accelerated profile HMM searches. *PLoS Comput. Biol*. **7**, e1002195 (2011).
100. Sievers, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol*. **7**, 539 (2011).
101. Steenwyk, J. L., Iii, T. J. B., Li, Y., Shen, X.-X. & Rokas, A. ClipKIT: a multiple sequence alignment trimming software for accurate phylogenomic inference. *PLoS Biol*. **18**, e3001007 (2020).
102. Lee, T. S. et al. BglBrick vectors and datasheets: a synthetic biology platform for gene expression. *J. Biol. Eng*. **5**, 12 (2011).
103. Xia, Y. et al. T5 exonuclease-dependent assembly offers a low-cost method for efficient cloning and site-directed mutagenesis. *Nucleic Acids Res*. **47**, e15 (2019).
104. Doyle, T. et al. The interferon-inducible isoform of NCOA7 inhibits endosome-mediated viral entry. *Nat. Microbiol*. **3**, 1369–1376 (2018).
105. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res*. **42**, W320–W324 (2014).
106. Okonechnikov, K., Golosova, O., Fursov, M. & UGENE team Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics* **28**, 1166–1167 (2012).
107. Etienne, L. Alignments and trees associated with Fig. 1 and S1: structural homology analyses show strong similarity between SAMD9 gene family and Avs antiphage proteins. *figshare* https://doi.org/10.6084/m9.figshare.29082653.v2 (2025).
108. Etienne, L. Alignments and trees associated with Fig. 2 and S2: Avs9 is part of prokaryotic defense systems and induces cell death in bacteria, through its SAMD9/SLFN-analogous active site in the AlbA_2 domain. AlbA_2 effector domain is shared between human SLFNs, SAMD9s and bacterial Avs9, with a conserved catalytic site. *figshare* https://doi.org/10.6084/m9.figshare.29082710.v1 (2025).

109. Etienne, L. Alignments and trees associated with Fig. 3 and S3: SAMD9/9L gene tree in mammals (with some other vertebrates as outgroups). *figshare* https://doi.org/10.6084/m9.figshare.29042762.v1 (2025).
110. Etienne, L. Alignments and trees associated with Fig. 4 and S4: primate phylogenetic analyses of SAMD9 and SAMD9L. Codon alignments used as inputs. *figshare* https://doi.org/10.6084/m9.figshare.29087987.v1 (2025).
111. Etienne, L. Alignments associated with S5c: SAMD9L primate amino acid sequence alignment showing the bonobo-specific variants. *figshare* https://doi.org/10.6084/m9.figshare.29088050.v1 (2025).
112. Etienne, L. Avs9 gene fragment sequence synthesized and used in this study, associated with Fig 3 and S3. *figshare* https://doi.org/10.6084/m9.figshare.29082719.v1 (2025).

Acknowledgements

We thank F. Fiorini (MMSB), F. Roucher-Boulez (HCL), A. Bernheim and H. Vaysset (Pasteur Institute) for feedback on the project, LP2L past and present laboratory members for support and discussion, L. Guéguen (LBBE) for providing guidance in the positive selection analyses in DGINN. We thank A. Le Corf (CIRI), C. Goujon (IRIM), L. Guiguettaz and E. Ricci (LBMC), B. Hahn (University of Pennsylvania, Perelman School of Medicine), M. Emerman (Fred Hutch) and the NIH AIDS reagent programme for sharing reagents. We acknowledge the contribution of SFR Biosciences (Université Claude Bernard Lyon 1, CNRS UAR3444, Inserm US8, ENS de Lyon) ANIRA platform, especially V. Barateau for her assistance, as well as D. Decimo of the BSL-3 ENS-Lyon platform. We also thank all the contributors of the publicly available bioinformatic programmes and genomic sequences. This work was funded by grants from: the French Research Agency on HIV and Emerging Infectious Diseases ANRS/MIE (grant nos ECTZ19143 and ECTZ245897) to L.E., the Sidaction (grant no. 23-1-AEQ-13601 to L.E. and 2020—grant no. 12673 and 2023—grant no. 13574 PhD fellowships to A.L.), a 2021 France-Berkeley Fund Award to L.E. and PHS, Agence Nationale de la Recherche under France 2030 bearing the reference ANR-24-RRII-0005 on funds administered by Inserm (EvoCure) to L.E. and F.R., Institute of General Medical Sciences (grant no. R35GM142916) to P.H.S., the Vallee Scholars Award to P.H.S., the Weill Neurohub Award to P.H.S. L.E. and P.H.S. are further supported by CNRS International Research Project IRP RAPIDvBAT. L.E. and A.C. are supported by the CNRS. 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

A.L. and L.E. conceived the project. L.E. supervised the project. A.L., R.D., A. Chantharath, C.L.M.G., M.S., F.R., P.H.S. and L.E. participated in the development or design of the methodology. A.L. and C.L.M.G. performed Foldseek analyses and downstream phylogenetics, with supervision from L.E. and M.S. A.L. and F.R. performed phylogenetic analyses across the tree of life. J.B. and F.R. performed experiments in bacteria. A. Chantharath and L.E. performed phylogenetics, genomics and positive selection analyses on mammals and primates, with help from L.P. who performed preliminary analyses. J.L.R. generated geographic maps and provided support for the population genetic and genomic analyses. P.H.S. performed the genomic locus alignment and supervised A.L. in genomic analyses in *Pan* individuals. C.R. generated the initial script for polymorphism analyses and supported A.L. in biocomputational analyses. C.L. provided technical support for molecular biology assays. A.L., R.D. and L.E. performed lentiviral infection experiments. A.L. performed translational shutoff assay, with help from A. Chantharath. A.L., A. Chantharath, R.D., C.R., C.L.M.G., A. Cimarelli, M.S., F.R., P.H.S., L.E. provided resources for this work. A.L., A. Chantharath and L.E. wrote the original draft of the manuscript, with help from R.D., J.L.R., C.R. and F.R. for the visualization. A.L., A. Chantharath, R.D., C.R., J.B., C.L.M.G., J.L.R., C.L., L.P., A. Cimarelli, M.S., F.R., P.H.S. and L.E. reviewed and edited the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41559-025-02845-x.

Correspondence and requests for materials should be addressed to Lucie Etienne.

Peer review information *Nature Ecology & Evolution* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

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

[1]Centre International de Recherche en Infectiologie (CIRI), Inserm U1111, UCBL1, CNRS UMR5308, ENS de Lyon, Université de Lyon, Lyon, France. [2]School of Biological Sciences, Seoul National University, Seoul, South Korea. [3]Department of Integrative Biology, University of California, Berkeley, CA, USA. [4]These authors contributed equally: Rémi Demeure, Amandine Chantharath. ✉e-mail: lucie.etienne@ens-lyon.fr

nature portfolio

Corresponding author(s): Lucie ETIENNE

Last updated by author(s): May 20, 2025

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	*Provide a description of all commercial, open source and custom code used to collect the data in this study, specifying the version used OR state that no software was used.*
Data analysis	DGINN, NCBI Blastn, IQ-TREE 2.3.0 , aBS-REL from DataMonkey webserver, R script using VariantAnnotation, GenomicFeatures, AnnotationHub, org.Pt.eg.db, ggplot2, R packages, LiftOver tool of the UCSC genome browser , AlphaFold DB , Foldseek , FoldMason v.333d54c , iTol , MMseqs2 (v 13.45111) , hmmsearch (v3.3.2) , DefenseFinder (v 1.3), Geneious (Biomatters), ESPript 3.0 and UGENE v52.0, GraphPad Prism 9, The scripts for the analyses of the polymorphic sites are available at http://gitbio.ens-lyon.fr/ciri/lp2l/vcf_to_nice_figures.git.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Availability of codes, data and reagents. All data and reagents are available in this manuscript (including in Dataset S1-S3) or accessible upon request to the corresponding author. The scripts for the analyses of the polymorphic sites are available at http://gitbio.ens-lyon.fr/ciri/lp2l/vcf_to_nice_figures.git. The alignments and phylogenetic trees are all openly available through FigShare deposits (Phylogenetics associated: with Fig. 1 and S1 https://doi.org/10.6084/m9.figshare.29082653 ; with Fig. 2 and S2 https://doi.org/10.6084/m9.figshare.29082710 ; with Fig. 3 and S3 https://doi.org/10.6084/m9.figshare.29042762 ; with Fig. 4 and S4 https://doi.org/10.6084/m9.figshare.29087987 ; with Fig. S5C https://doi.org/10.6084/m9.figshare.29088050). Avs9 sequence used for experimental assays in Fig. 3D and S3D is available at https://doi.org/10.6084/m9.figshare.29082719.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	na
Reporting on race, ethnicity, or other socially relevant groupings	na
Population characteristics	na
Recruitment	na
Ethics oversight	na

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	na
Data exclusions	If the host and viral dna transfected in the producer cells did not show any protein expression by Western-blot (which occured in one isolated case), we excluded it, because it certainly resulted from transfection problem in a single well.
Replication	Each infection experiment was performed in intra-assay duplicates (technical replicates). Each experiment was itself performed in 3-5 independant replicates (biological replicates), except in Fig. 6c (2 independent replicates). In the manuscipt, data are presented as mean ± SD, and each point correspond to an independent biological replicate. Statistics were performed using the two-sided ratio paired t-test (*, p value < 0.05; **, p value < 0.005; ***, p value <0.0005).
Randomization	na
Blinding	na

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used: anti-SAMD9L (Proteintech, 25173-1-AP), anti-Gag (NIH HIV Reagent Program, 183-H12-5C), anti-HIV-1-gp120 (Aalto, D7324; NIH HIV Reagent Program, 16H3), as well as secondary IgG-Peroxidase conjugated anti-mouse (Sigma, cat. A9044) and anti-rabbit (Sigma, cat. AP188P).

Validation: na

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s): Human embryonic kidney 293T (ATCC, cat. CRL-3216) and TZM-bl (NIH AIDS Research and Reference Reagent Program, Cat. 8129)

Authentication: none

Mycoplasma contamination: Confirm that cells were tested for mycoplasma and were negative.

Commonly misidentified lines (See ICLAC register): na

Plants

Seed stocks: na

Novel plant genotypes: na

Authentication: na

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation: Medium was discarded, and cells were harvested and fixed with PFA 4%. Cells were then washed with PBS BSA 3% and permeabilized in PBS 0,5% Triton X-100 for 15 min. Click-iT® Plus Alexa Fluor® Picolyl Azide assay was then performed and cells were analyzed on MACSQuant® VYB Cytometer (Miltenyi Biotec, SFR BioSciences).

Instrument	MACSQuant® VYB Cytometer (Miltenyi Biotec, SFR BioSciences).
Software	FloJo
Cell population abundance	Available in Dataset_S1-gating-strategy.pdf
Gating strategy	Available in Dataset_S1-gating-strategy.pdf

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.