



HAL
open science

Recurrent Evolutionary Innovations in Rodent and Primate Schlafen Genes

Joris Mordier, Marine Fraisse, Michel Cohen-Tannoudji, Antoine Molaro

► **To cite this version:**

Joris Mordier, Marine Fraisse, Michel Cohen-Tannoudji, Antoine Molaro. Recurrent Evolutionary Innovations in Rodent and Primate Schlafen Genes. 2024. hal-04418574

HAL Id: hal-04418574

<https://hal.science/hal-04418574v1>

Preprint submitted on 26 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



== REVIEW COMMONS MANUSCRIPT ==

IMPORTANT:

- Manuscripts submitted to Review Commons are peer reviewed in a journal-agnostic way.
- Upon transfer of the peer reviewed preprint to a journal, the referee reports will be available in full to the handling editor.
- The identity of the referees will NOT be communicated to the authors unless the reviewers choose to sign their report.
- The identity of the referee will be confidentially disclosed to any affiliate journals to which the manuscript is transferred.

GUIDELINES:

- For reviewers: <https://www.reviewcommons.org/reviewers>
- For authors: <https://www.reviewcommons.org/authors>

CONTACT:

The Review Commons office can be contacted directly at: office@reviewcommons.org

Recurrent Evolutionary Innovations in Rodent and Primate *Schlafen* Genes

Joris Mordier¹, Marine Fraisse¹, Michel Cohen-Tannoudji², Antoine Molaro^{1#}

¹Institute of Genetics, Reproduction & Development, CNRS UMR6293, INSERM U.1103, Université Clermont Auvergne, 28 Place Henri Dunant, Clermont-Ferrand, France.

²Institut Pasteur, Université Paris Cité, CNRS UMR3738, Early Mammalian Development and Stem Cell Biology, F-75015, Paris, France

Running head: Schlafen Genes in Conflicts

Keywords: Schlafen, duplication, positive selection, sub functionalization, genetic conflicts.

Address correspondence to: antoine.molaro@uca.fr, Antoine Molaro, 28 Place Henri Dunant, iGReD, Clermont-Ferrand, 63000, France.

1 **ABSTRACT**

2 **SCHLAFEN proteins are a large family of RNase-related enzymes carrying essential**
3 **immune and developmental functions. Despite these important roles, *Schlafen* genes**
4 **display varying degrees of evolutionary conservation in mammals. While this appears**
5 **to influence their molecular activities, a detailed understanding of these evolutionary**
6 **innovations is still lacking. Here, we used in depth phylogenomic approaches to**
7 **characterize the evolutionary trajectories and selective forces shaping mammalian**
8 ***Schlafen* genes. We traced lineage-specific *Schlafen* amplifications and found that**
9 **recent duplicates evolved under distinct selective forces, supporting repeated sub-**
10 **functionalization cycles. Codon-level natural selection analyses in primates and**
11 **rodents, identified recurrent positive selection over *Schlafen* protein domains engaged**
12 **in viral interactions. Combining crystal structures with machine learning predictions,**
13 **we discovered a novel class of rapidly evolving residues enriched at the contact**
14 **interface of SCHLAFEN protein dimers. Our results suggest that inter *Schlafen***
15 **compatibilities are under strong selective pressures and are likely to impact their**
16 **molecular functions. We posit that cycles of genetic conflicts with pathogens and**
17 **between paralogs drove *Schlafens*' recurrent evolutionary innovations in mammals.**

18

19

20 **INTRODUCTION**

21

22 *Schlafens* are a class of interferon stimulated genes that participate in immune response to
23 viral infections (Kim and Weitzman 2022). In human, SCHLAFEN (SLFN) proteins have been
24 shown to restrict a wide variety of viruses including HIV (Li et al. 2012; Stabell et al. 2016;
25 Ding et al. 2022); HSV (Kim et al. 2021) and Influenza (Seong et al. 2017). While there is much
26 diversity in their mode of action, most SLFNs also play important core cellular functions most
27 notably by regulating cellular rRNAs and tRNAs metabolism (Pisareva et al. 2015; Murai et al.
28 2018; Yang et al. 2018; Chen and Kuhn 2019; Garvie et al. 2021; Metzner et al. 2022a;
29 Metzner et al. 2022b).

30

31 In rodents and primates, these functions have been tied to unique structural features. All SLFN
32 proteins are structurally related to eukaryotic and viral RNase E and NTPase/Helicase
33 enzymes (Chen and Kuhn 2019; Jo and Pommier 2022). Their N-terminal regions universally
34 encode a "SCHLAFEN core domain" combining a "SCHLAFEN box" and RNase E-related
35 AlbA domain (Pisareva et al. 2015; Yang et al. 2018; Chen and Kuhn 2019; Garvie et al. 2021;

36 Jo and Pommier 2022; Metzner et al. 2022a, Metzner et al. 2022b). This domain promotes
37 SCHLAFEN dimerization creating an RNA binding pocket with endonuclease activity when
38 catalytic residues are present (Garvie et al. 2021; Metzner et al. 2022b). This function is
39 essential to coordinate rRNA and tRNA binding and/or cleavage upon various cellular stresses
40 (Schwarz et al. 1998; Geserick et al. 2004; Zoppoli et al. 2012; Arslan et al. 2017; Fischietti et
41 al. 2018; Yang et al. 2018; Yue et al. 2021; Metzner et al. 2022a).

42

43 The SCHLAFEN core domain is followed by a linker domain containing a “SWADL” motif- and
44 with structural similarities to GTPases (Chen and Kuhn 2019). The C-terminal portion is an
45 NTPase-related Helicase domain, sometimes referred to as P-Loop (Chen and Kuhn 2019; Jo
46 and Pommier 2022). These domains also contribute to dimerization and interaction with other
47 proteins (Chen and Kuhn 2019; Garvie et al. 2021; Kim et al. 2021; Jo and Pommier 2022;
48 Metzner et al. 2022b; Nightingale et al. 2022). For some SLFN proteins, these domains
49 coordinate nucleotide binding - *e.g.*, SLFN12 with GDP and cyclic NTPs (Chen and Kuhn
50 2019; Garvie et al. 2021); SLFN5 with ATP (Metzner et al. 2022a) – albeit with no hydrolysis
51 activity on their own (Garvie et al. 2021). Moreover, the SWADL and Helicase domains have
52 been recently shown to bind ssDNA *in vitro* (Metzner et al. 2022b), and are necessary for viral
53 restriction *in vivo* (Li et al. 2012; Pisareva et al. 2015).

54

55 From an evolutionary stand-point, *Schlafen* genes are found across all vertebrates and likely
56 derive from a single parental RNase E-like ortholog (Chen and Kuhn 2019). All extant
57 vertebrate *Schlafen* genes are found within a single genomic cluster (flanked by *Unc45b* and
58 *Pex12*). They group into four phylogenetically distinct clades: Clade 1 relates to human
59 *SLFN12*; Clade 2 to *SLFN5*; Clade 3 to *SLFN14*; and Clade 4 to *SLFN11* (Bustos et al. 2009).
60 However, the total number of *Schlafen* genes found in this cluster varies greatly between
61 species, as a result of repeated lineage-specific gain and loss events (Bustos et al. 2009). For
62 instance, the *Schlafen* cluster encodes for one coding and two pseudogenized *Schlaferns* in
63 dogs (*Canis familiaris*); six in humans (*Homo sapiens*); and nine coding and one pseudogene
64 in mice (*Mus musculus*) (Bustos et al. 2009; Lilue et al. 2018).

65

66 These expansions and contractions were accompanied by structural changes between and
67 within clades (Bustos et al. 2009; Chen and Kuhn 2019; Jo and Pommier 2022). The most
68 drastic example is the complete loss of the C-terminal helicase domain in the parental gene
69 giving rise to Clade 1 *SLFNs* roughly 100 million years ago (Ma) followed by further shortening
70 of the C-terminal portion during rodent evolution (Bustos et al. 2009). Finally, vertebrate

71 *Schlafen* orthologs have been suggested to display accelerated rates of codon evolution
72 (Bustos et al. 2009; Stabell et al. 2016; Lilue et al. 2018). In the case of primate *SLFN11*,
73 specific codons subject to positive selection were shown to have shaped its antiviral activity
74 most notably against HIV-1 in human cells (Stabell et al. 2016). Nevertheless, for most
75 *Schlaferns*, the link between sequence variation and function remains unknown.

76
77 The leading hypothesis underlying the evolutionary diversification of the *Schlafen* family is
78 their involvement in an ongoing genetic conflict with pathogens (Bustos et al. 2009; Kim and
79 Weitzman 2022). Genetic conflicts arise when co-evolving genetic entities engage in
80 antagonistic interactions (Gardner and Ubeda 2017; McLaughlin and Malik 2017). This
81 promotes rapid adaptation/counter-adaptation cycles, or evolutionary arms races, driving the
82 expansion and positive selection of protein domains engaged in antagonism (Hurst and
83 Werren 2001; McLaughlin and Malik 2017; Daugherty and Zanders 2019; Kuzmin et al. 2022).
84 Considering that many primates SLFN proteins have well described antiviral functions, their
85 rapid diversification could result from an arms race with viral proteins (Kim and Weitzman
86 2022). Supporting this hypothesis, human SLFN5 and SLFN11 are the targets of viral
87 antagonisms in the context of HSV-1 and HCMV restriction respectively (Kim et al. 2021;
88 Nightingale et al. 2022).

89
90 However, *Schlaferns* rapid evolution might also relate to functions beyond antiviral
91 antagonisms. This might be the case in mice where genetic mapping of a strain incompatibility
92 locus was directly linked to the *Schlafen* gene cluster (Cohen-Tannoudji et al. 2000; Bell et al.
93 2006). This suggests that SLFN proteins could play an important function during reproduction
94 and that their rapid evolution contributes to reproductive isolation, and ultimately speciation
95 (Crespi and Nosil 2013). In this case, whether this genetic conflict relates to antiviral
96 antagonism is unknown. More importantly, which *Schlafen* genes and which domains drive
97 this incompatibility remains to be determined.

98
99 Here, we use a detailed phylogenomic approach to determine the evolutionary trajectories and
100 selective forces that shaped mammalian *Schlafen* orthologs and paralogs. We trace the
101 origins of a novel *Schlafen-like* orphan gene in vertebrates and the lineage-specific
102 amplifications patterns of clustered *Schlafen* genes in mammalian genomes. We use in-depth
103 codon-level selection analyses to map recent diversification events in rodent and primate
104 genomes. We show that many recently duplicated *Schlafen* genes evolved under distinct
105 selective forces and identify recurrent positive selection over *SLFN11*, *SLFN5* and *SLFN12*-

106 related genes. Finally, we use crystal structures with machine learning predictions, to identify
107 a novel class of residues subjected to positive selection mapping at the dimer interface
108 between SLFN proteins. We suggest that this diversification shaped both antiviral and
109 developmental functions of *Schlafen* paralogs.

110

111 RESULTS

112

113 Evolutionary trajectory of the *Schlafen* gene cluster and *SLFNL1* in mammals

114 To comprehensively study the origins and selective pressures acting on mammalian *Schlafen*
115 orthologs, we first set out to update the evolutionary history of the *Schlafen* gene cluster using
116 up-to-date vertebrate genomes (Bustos et al. 2009). We used TBLASTN searches with
117 annotated SLFN proteins from Human (*Homo sapiens*), Horse (*Equus caballus*) and Elephant
118 (*Loxodonta africana*) to identify candidate *SLFN* genes across vertebrate and extending to
119 several outgroups, such as Ciona (*Ciona intestinalis*) (Methods). This approach retrieved at
120 least one putative *Schlafen* ortholog in most vertebrate genomes but none with significant
121 homology in species outside of jawed vertebrates - i.e., in Elephant shark (*Callorhinchus milii*)
122 but not in hagfish (*Eptatretus burgeri*) (Figure 1A). In jawless vertebrates, none of the synteny
123 is preserved blurring further homolog identification based on assembled contigs (Figure 1A).

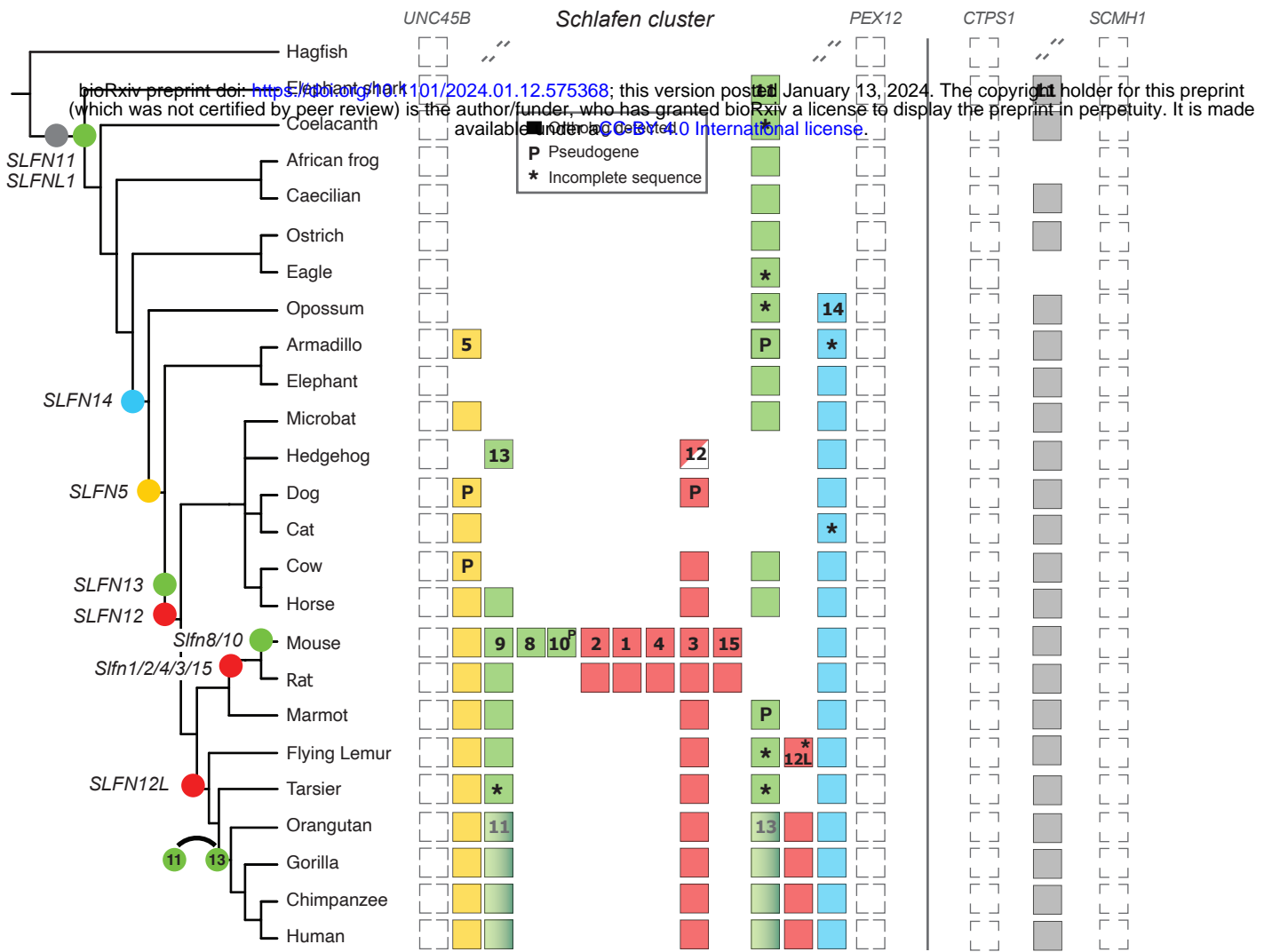
124

125 To identify specific orthologs, we first analyzed coding sequences (CDS) syntenic location on
126 existing genome assemblies or using manual contig annotation (see Methods). We found that
127 most hits mapped back to the expected *Schlafen* gene cluster flanked by the *Unc45b* and
128 *Pex12* genes (Bustos et al. 2009) (Figure 1A). However, one hit consistently mapped outside
129 of the known *Schlafen* cluster between the *Ctps1* and *Scmh1* genes (Figure 1A). This locus
130 codes for a poorly characterized *Schlafen*-related gene found in most vertebrate genomes,
131 named *SLFNL1* in human (RefSeq accession NM_144990). While *SLFNL1* predicted protein
132 shares less than 10% identity with other *SLFNs*, it nonetheless contains a SCHLAFEN AlBA
133 domain with conserved catalytic residues indicative of potential RNAse E functionality
134 (Supplemental Fig. S1). Thus, our syntenic analysis suggests that the last common ancestor
135 to all jawed vertebrates likely encoded two potentially enzymatically active *Schlafen* genes,
136 one within the cluster and one orphan related to *SLFNL1*.

137

138 Having identified the genomic location of all *Schlafen* homologs, we wished to uncover their
139 evolutionary relationships to finalize ortholog assignment based on shared ancestry.
140 Considering the large time span of our analysis – ~500My (Million years) - we used predicted

A



B

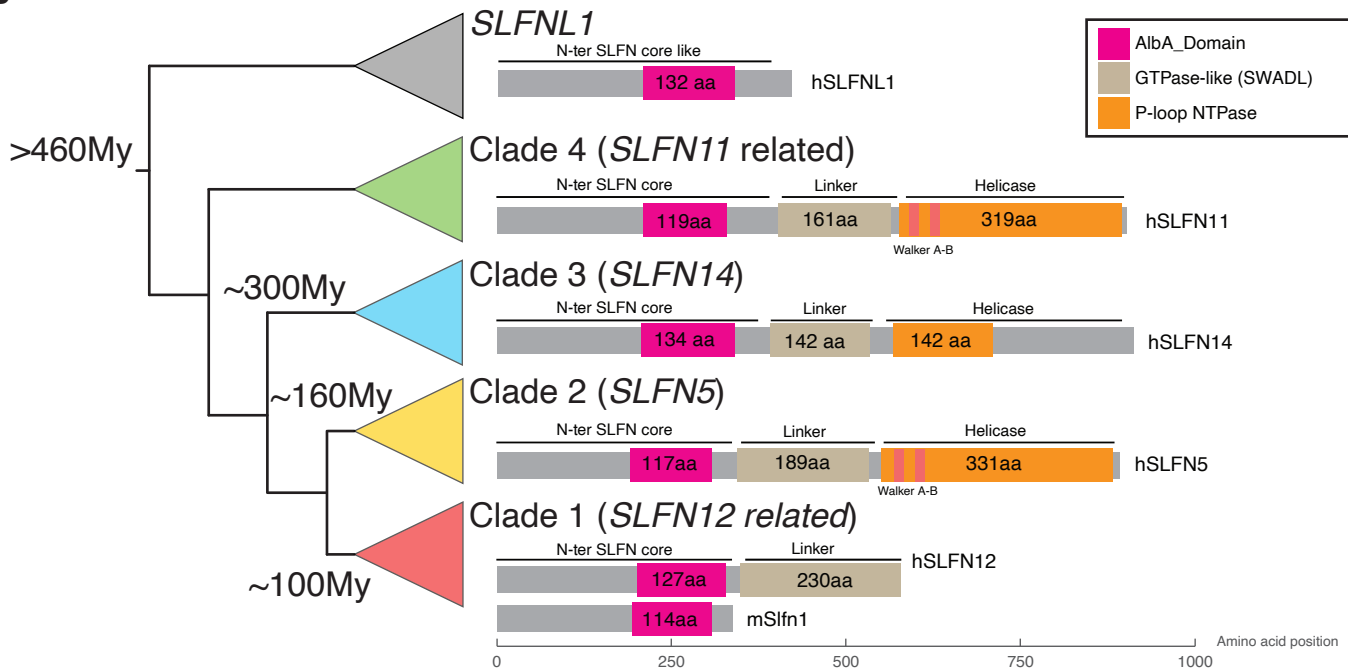


Figure 1. Evolutionary history of *Schlafen* genes in Vertebrates.

(A) Schematic representation of genomic loci encoding clustered *Schlafen* (*SLFN*) and *Schlafen Like 1* (*SLFN1*) genes in selected species of vertebrates. Filled circles denote major gene birth events along the accepted species phylogeny. For each clade, presence of *Schlafen* and *SLFN1* loci are shown with filled boxes with the associated *Schlafen* gene family number shown for the top instance. Empty spaces indicate that no putative ortholog was found in the assembly. Genes flanking each locus are shown with empty boxes and dashed backslashes indicate a break in synteny. “P” denotes pseudogenes and “*” incomplete sequences. A switch in *SLFN13* and *SLFN11* is shown at the base of primates. Incomplete lineage sorting of Hedgehog putative *Sfn12* is shown with a half-filled box.

(B) Inferred evolutionary history of *Schlafen* and *SLFN1* clades based on phylogenetic and syntenic analyses. Clades encoding for *SLFN12* (red), *SLFN5* (yellow), *SLFN14* (blue), *SLFN11* (green) and *SLFN1* (gray) are shown with their estimated date of birth. For each clade, representative secondary structures of human or mouse protein are shown with their associated domains mapped relative to the starting amino-acid (bottom X-axis).

141 SLFN proteins to run maximum likelihood (ML) phylogenetic analyses (see Methods). This
142 identified five distinct *Schlafen* monophyletic clades: the four previously reported *Schlafen*
143 gene clades with the addition of one novel clade related to *SLFNL1* (Figure 1B, Supplemental
144 Fig. S2 & Supplemental Table S1). Increased divergence did not allow proper grouping of non-
145 mammalian vertebrate SLFN11 with their orthologs (Supplemental Fig. S2). Yet, these
146 homologs code for all the key residues and domains characteristic of this family, supporting
147 their shared ancestry (Supplemental Data S1). Finally, within mammals, the hedgehog
148 putative *SLFN12* ortholog grouped outside of both *SLFN5* and *SLFN12* clades (Supplemental
149 Fig. S2). Indeed, while its N-terminal portion is characteristic of other *SLFN12*, it appears to
150 have acquired an extended C-terminal distantly related to *SLFN5* (Supplemental Fig. S3).

151
152 Combining our genomic and phylogenetic analyses, we made several interesting
153 observations. First, the last common ancestor to all jawed-vertebrate likely encoded for one
154 orthologs of *SLFN11* (clade 4) located in the *Schlafen* gene cluster, and one orphan *SLFNL1*
155 (Figure 1A). These two *Schlaferns* remained in most vertebrate genomes, apart from
156 Carnivores that lost *SLFN11*-related genes, and some amphibians and birds that
157 independently lost *SLFNL1* (Figure 1A). Second, the *Schlafen* gene cluster expanded in
158 mammals with the successive birth of *SLFN14* (~300Ma), *SLFN5* (~160Ma) and more recently
159 *Sfn12*-related genes (~100Ma) (Figure 1).

160
161 Following these birth events, *SLFN14* remained intact in all mammalian genomes
162 investigated, while all other clustered *Schlaferns* underwent lineage-specific loss/amplification
163 events (Figure 1A). A few examples include: (1) the duplication of a parental *SLFN11* into
164 *SLFN13/11* paralogs following the split of Atlantogenates (*i.e.*, elephants & armadillos) from
165 the rest of placental mammals; (2) the amplification burst of *Sfn12*-related paralogs following
166 the split between marmots and other rodents (discussed in detail below); or (3) the birth of
167 *SLFN12L* in the last common ancestor of primates and dermoptera (Bustos et al. 2009). In
168 sum, our phylogenomics approach identified a novel orphan *Schlafen*-related gene that likely
169 arose in vertebrates, and confidently assigned orthology between *Schlafen* cluster paralogs
170 allowing for accurate natural selection studies.

171

172

173 Diversifying selection shapes *Schlafen* genes in Primates.

174 As seen for other immune-related genes, SLFN antiviral functions are expected to drive
175 recurrent evolutionary diversification, or positive selection, in response to pathogens

176 (McLaughlin and Malik 2017). While this has been demonstrated for primates' SLFN11 N-
177 terminal region (Stabell et al. 2016), whether similar selective forces also shape other paralogs
178 remains unexplored. To address these shortcomings, we retrieved and annotated *Schlafen*
179 orthologs in simian primates - which includes new world monkeys, old world monkeys and
180 great apes - and used Tarsiers as an outgroup. This group is well suited for natural selection
181 analysis as it spans ~60My without saturation of synonymous mutation rates over 20
182 annotated genomes (Perelman et al. 2011).

183
184 We found that all primates have an intact *Schlafen* cluster encoding orthologs of *SLFN5*,
185 *SLFN11*, *SLFN12*, *SLFN13*, *SLFN12L*, *SLFN14*, in addition to an intact *SLFNL1* (Figure 2,
186 Supplemental Fig. S2 and Supplemental Table S1). However, following the split with tarsiers,
187 an inversion switched *SLFN11* and *SLFN13* locations but left the rest of the cluster unaffected
188 (Figure 1A, and Supplemental Table S1). Closer investigation of the locus could not resolve
189 the genetic nature of this event since *SLFN11* and *SLFN13* surrounding regions as well as the
190 other *Schlafen* loci in the cluster kept their parental orientation. However, primate *Schlafen*
191 orthologies were well supported by maximum likelihood phylogenies (Figure 2 and
192 Supplemental Table S1).

193
194 To determine the strength and direction of natural selection, we measured synonymous (dS)
195 and non-synonymous (dN) codons substitution rates across codon-aligned predicted CDS
196 (see Methods). We first used likelihood approaches to test if modeling dN/dS across the whole
197 gene showed significant deviation from neutral evolution (dN/dS close to 1). While most
198 *Schlafen* genes showed strong support for natural selection (Figure 2 and Supplemental Table
199 S2), we failed to reject neutral evolution for *SLFN13* (mean dN/dS=1.06, see Methods) and
200 *SLFN12L* (dN/dS=0.68). Weak signatures of selection could indicate that these orthologs
201 accumulate mutations without measurable fitness consequences and might be on their way to
202 becoming non-functional in primates.

203
204 Of those *Schlaferns* under selection, *SLFN14* had the lowest overall dN/dS (0.39) likely
205 reflecting predominant purifying selection over core housekeeping functions (Figure 2 and
206 Supplemental Table S2) (Pisareva et al. 2015; Seong et al. 2017). For other paralogs, we
207 turned to codon-level dN/dS likelihood-based modeling to determine which residues and
208 domains might be subject to positive selection (see Methods). Considering the clustering of
209 *Schlafen* genes, they could experience increased recombination rates sometimes leading to
210 pervasive gene conversion events within or between paralogs (Mitchell et al. 2015; Daugherty

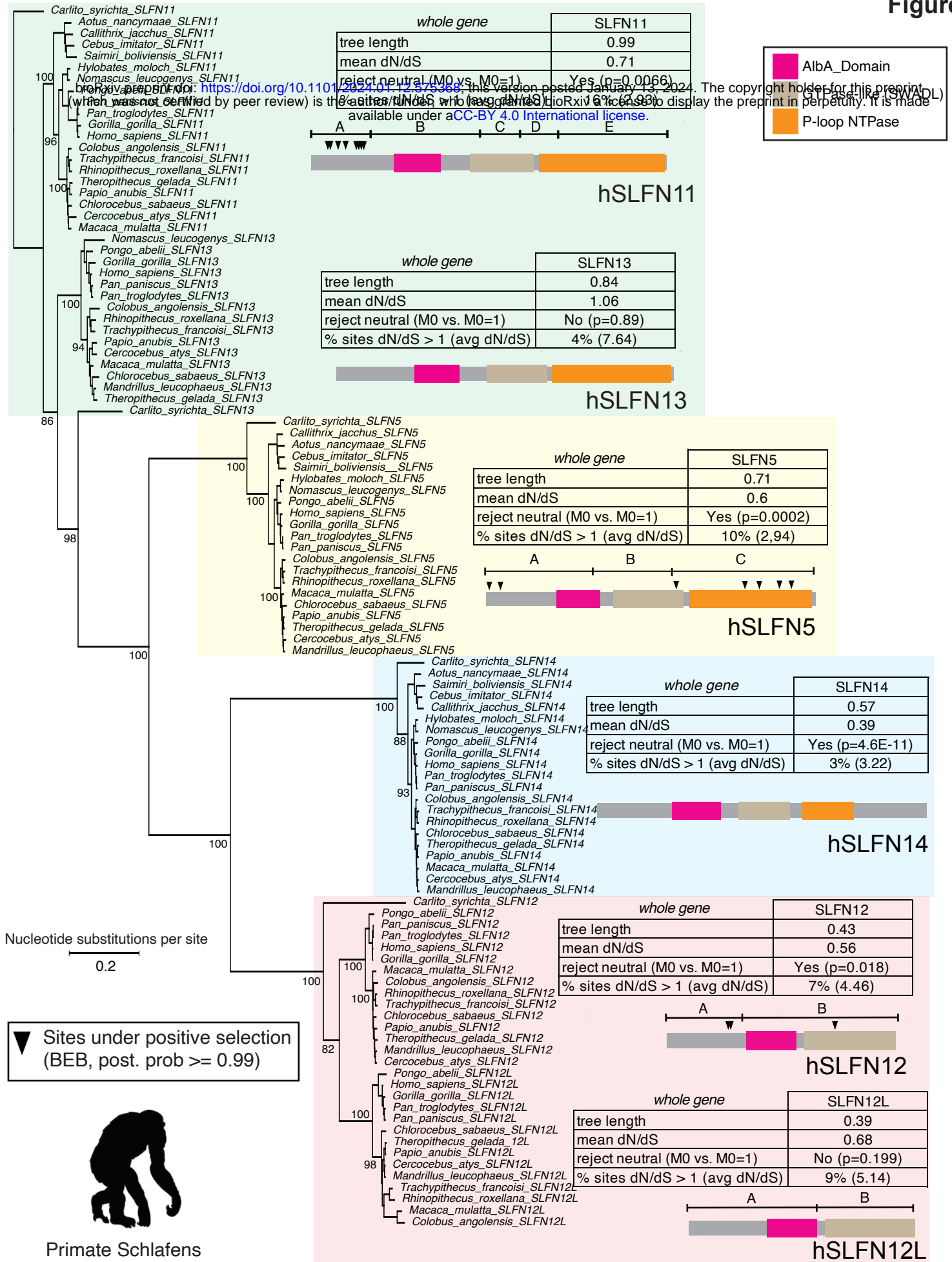


Figure 2. Rapid evolution of *Schlafen* genes in primate genomes.

Maximum likelihood phylogeny of all identified *Schlafen* coding sequences in simian primates. The tree is artificially rooted on Tarsier *SLFN11*. Only major bootstrap support values >50% are shown. For each highlighted clade, a summary table of selection analyses on whole CDS are shown (“whole gene”). Sites evolving under positive selection (PAML, BEB posterior probability >99%) are mapped with arrowheads under each putative recombination segments detected by GARD. Chimpanzee silhouette is from <https://www.phylopic.org>.

211 et al. 2016; Daugherty and Zanders 2019, Molaro et al. 2020). These events can alter local
212 substitution rates over specific CDS portions and thus blur codon-level selection analyses that
213 model homogeneous dN/dS across whole CDS. To account for this potential caveat, we used
214 GARD to predict putative recombination segments (see Methods) (Kosakovsky Pond et al.
215 2006). When detected, we performed separate selection analyses on each predicted segment,
216 otherwise we used whole CDS (Figure 2, Supplemental Table S2).

217
218 This approach identified eight codons under recurrent positive selection clustered over the N-
219 terminal end of SLNF11 (segment A, BEB>0.99, Figure 2, Supplemental Table S2). Of these
220 positively selected codons, seven were directly overlapping with a prior study (M47, Q79,
221 C102, Y133, E137, V140, S142) attesting of the robustness of our analysis (Stabell et al.
222 2016). Using the recently published SLFN11 crystal structure (Metzner et al. 2022b), we found
223 that five of these sites mapped inside the RNase E binding pocket and are thus likely to affect
224 SLFN11 enzymatic function (Supplemental Fig. S4).

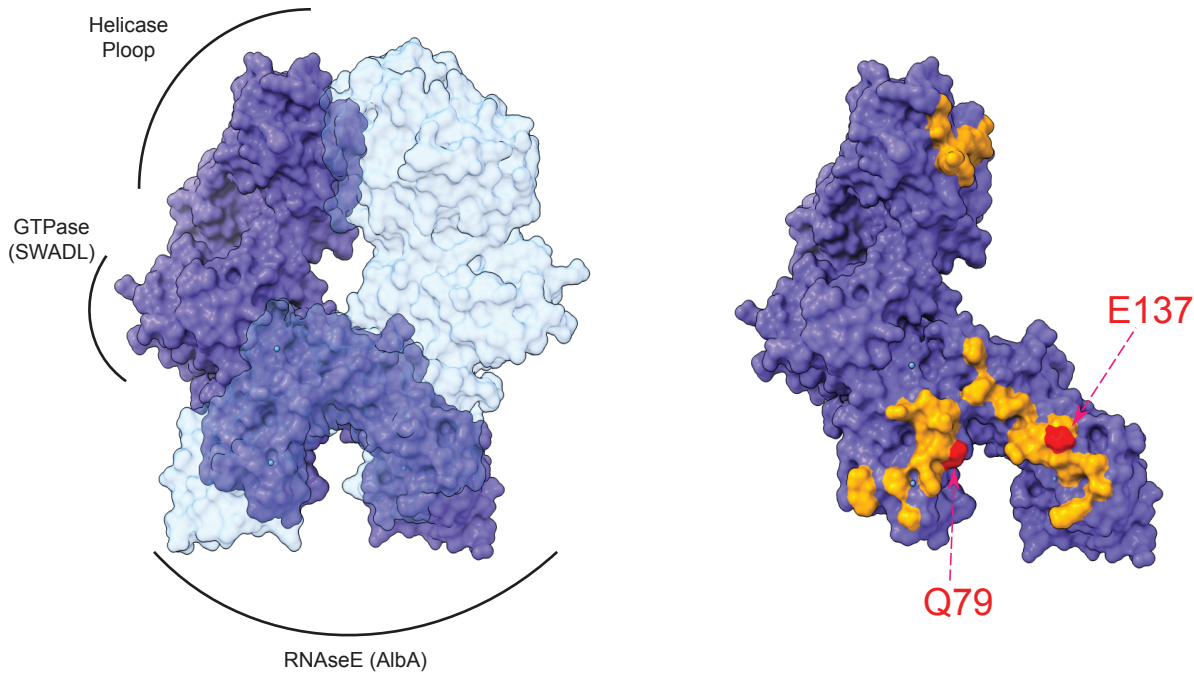
225
226 The remaining three residues mapped toward the surface of the protein. However, unlike what
227 was reported before, we found that two of these residues (Q79 and E137) located at the direct
228 contact interface between SLFN11 molecules in the crystalized homodimer (Figure 3A).
229 Finding 2 out of 8 positively selected residues (1/4) at this interface is highly unlikely since
230 there are only 37/901 (~1/24) total residues within <4Å between the two monomers (chi-square
231 test p-value = 0,011) (Figure 3A). Thus, this indicates the SCHLAFEN contact interface is
232 enriched for diversifying residues. Interestingly, in the case of codon Q79, its side chain in one
233 monomer contacts a region neighboring two other rapidly evolving codon in the second
234 monomer (V140 and S142, Supplemental Fig. 4). This last observation further suggests that
235 intra-SLFN molecular interactions might be subject to ongoing diversification in primates.

236
237 Beside SLFN11, our analysis also uncovered several interesting new cases of positive
238 selection in primates. First, we identified seven positively selected sites in *SLFN5* (Figure 2,
239 Supplemental Table S2). Suggestive of similar selective pressure than those acting on
240 *SLFN11*, two of these rapidly evolving codons mapped to its N-terminal portion (segment A,
241 Figure 2, Supplemental Fig. S5). Since SLFN5 lost its RNase activity, the functional
242 significance of these changes is unknown (Metzner et al. 2022a). Outside of the N-terminus,
243 we found four diversifying residues within the C-terminal Helicase domain (P-loop NTPase,
244 segment C, Figure 2, Supplemental Table S2). SLFN5 helicase is implicated in both viral and
245 retrotransposon activity and is the direct target of the viral antagonist ICP0 encoded by HSV-

A

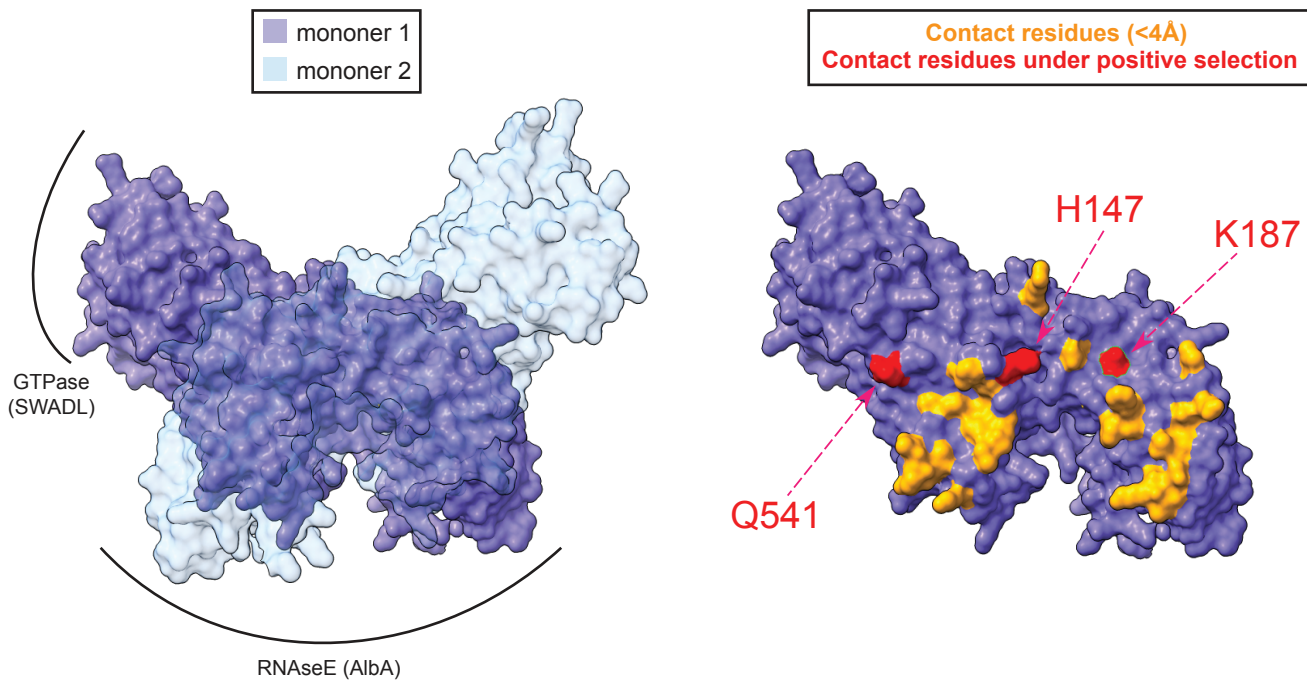
Human SCHLAFEN 11

bioRxiv preprint doi: <https://doi.org/10.1101/2024.01.12.575368>; this version posted January 13, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



B

Predicted Mouse SCHLAFEN 4

**Figure 3. Positive selection over SCHLAFEN homodimer contact residues.**

(A) Surface representation of crystalized Human SLFN11 homodimer (PDB: 7ZEP). Major domain localizations are shown along monomer 1 (purple structure). Monomer 2 is shown in transparency (light blue). Residues within 4Å between monomer 1 and 2 (contact residues) are highlighted in yellow onto monomer 1 (right panel). Contact residues under positive selection are shown in red with their relative position in the human protein.

(B) Same as (A) but for AlphaFold2 predicted mouse SLFN4 homodimer. Positions are relative to the mouse protein.

246 1 (Kim et al. 2021; Ding et al. 2023). Of note, diversification over the helicase domain was not
247 found in other *Schlafens*.

248

249 Second, we identified three positively selected residues over the 5' segment of *SLFN12*
250 (Figure 2, Supplemental Table S2). This contrasts with its duplicate gene *SLFN12L* that
251 displays low overall level of selection (see beginning of section). Mapping these residues over
252 the *SLFN12* structure did not show specific enrichment in the RNase domain or at the interface
253 between monomers, raising questions on the selective forces driving *SLFN12* diversification
254 (Supplemental Fig. S6). Nevertheless, this finding indicates that *SLFN12* and *12L* have been
255 subject to different selective forces following their duplication ~85Ma and likely carry-out
256 independent functions in today genomes.

257

258 In sum, our approach identified recurrent diversifying selection over the primate *Schlafen*
259 genes *SLFN11*, *5* and *12*. While many rapidly evolving residues map to known domains with
260 antiviral functions, we report novel sites that locate at their dimerizing interfaces. This
261 observation suggests that intra- and/or inter-*SLFN* interactions have been under strong
262 selection during primate evolution.

263

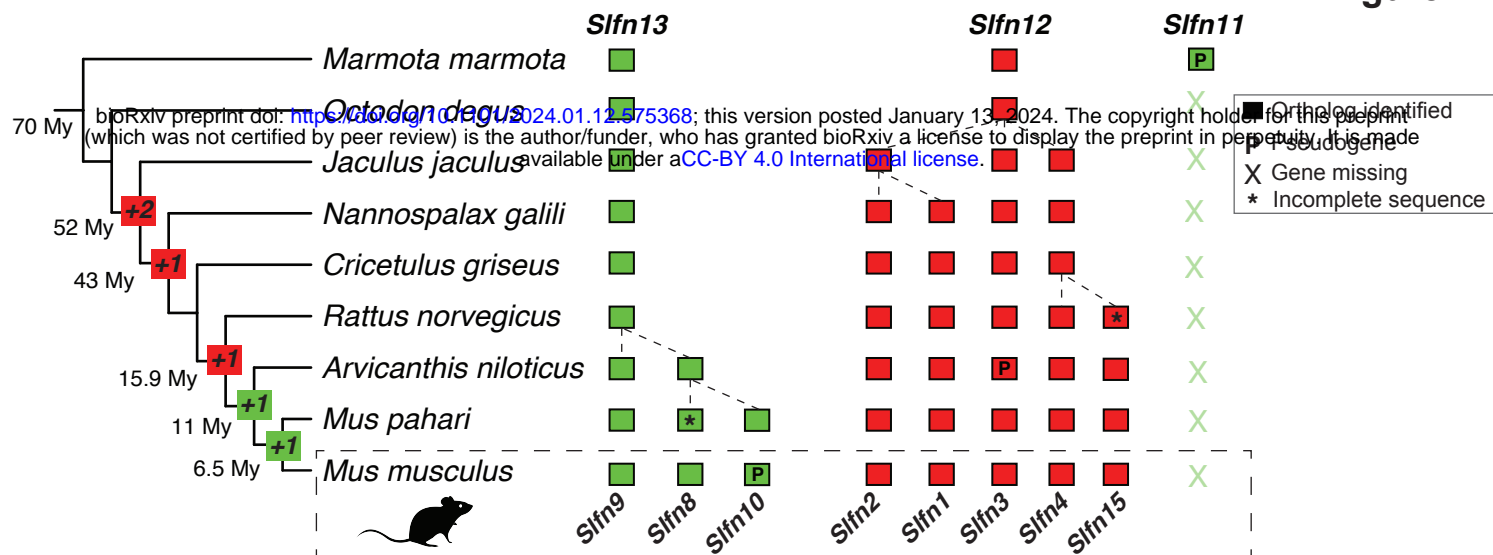
264 Birth of *Slfn12*- and *Slfn13*-related genes in rodents

265 Compared to primates, mouse *Schlafen* genes received less attention. This partially stems
266 from the complex structure of *Schlafen* paralogs within the mouse cluster (Lilue et al. 2018).
267 Indeed, in addition to encoding *Slfn14* and *Slfn5*, and unlike any other mammals, the mouse
268 cluster codes for multiple *Slfn12* and *Slfn13*-related genes with uncharacterized phylogenies
269 (Figure 1A) (Schwarz et al. 1998; Bell et al. 2006; Bustos et al. 2009; Lilue et al. 2018). We
270 used our in-depth phylogenomics approach to map these evolutionary transitions over 22
271 species spanning ~70My of rodent speciation (Kumar et al. 2017; Swanson et al. 2019).

272

273 We found that the amplification of *Slfn12*-related genes occurred in three successive
274 duplication rounds (Figure 4A, Supplemental Table S1). First, an event giving birth to one
275 GTPase-less *Slfn12* and two full-length paralogs (*Slfn3* and *Slfn4*) occurred following the split
276 between degus and other rodents ~50Ma (Figure 4A). A second round of duplication led to
277 the formation of two additional GTPase-less paralogs (*Slfn1* and *Slfn2*) in the last common
278 ancestor of muroids rodents (~43Ma, Figure 4A). One final duplication occurred in the last
279 common ancestor of mice and rats (~16Ma) and gave birth to a third full length gene *Slfn15*
280 (Figure 4A). We note that this gene was previously annotated as *GM11427*. Although

A



B

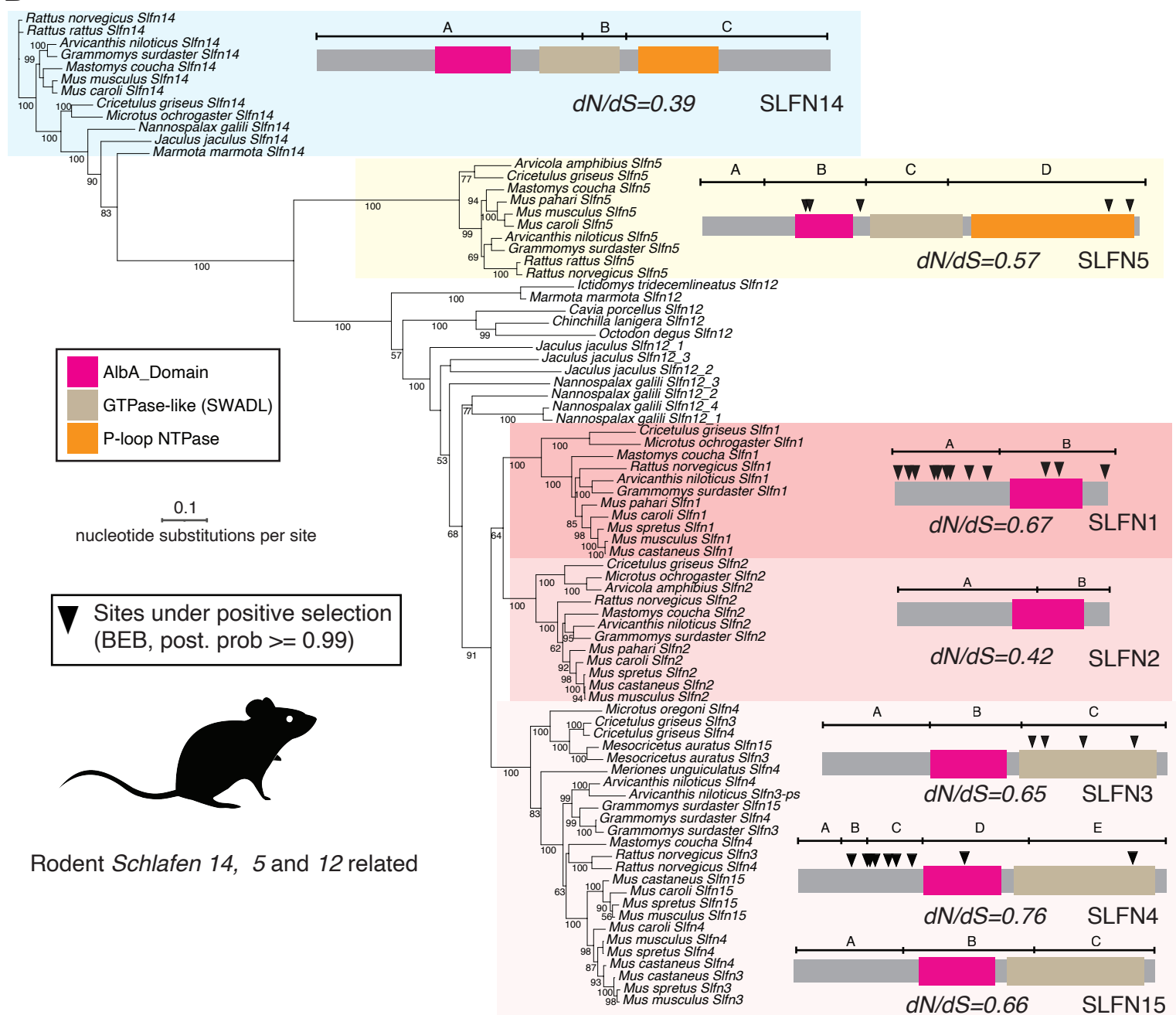


Figure 4. Amplification and rapid evolution of *Schlafen* genes in rodent genomes.

(A) Summary of *Slfn12* and *Slfn13* duplication events along the rodent species tree with estimated divergence times. Identified orthologs are shown with filled boxes. Gene names are shown relative to the Marmot (top) and mouse (bottom) genomes. Pseudogenes are indicated with “P” and incomplete sequences with “*”. *Slfn11* is missing from most genomes following the split with marmots (denoted with “X”).

(B) Maximum likelihood phylogeny of rodent *Slfn14*, *Slfn5* and *Slfn12*-related coding sequences. Bootstrap support values $>50\%$ are indicated. Average dN/dS for each *Slfn* clade is indicated below a schematic protein secondary structure. Codon evolving under positive selection (PAML, BEB posterior probability $>99\%$) are mapped with arrowheads under each putative recombination segments detected by GARD. Mouse silhouette is from <https://www.phylopic.org>.

281 maximum likelihood phylogenies only partially resolved each paralog family, likely due to
282 pervasive gene conversion (see next section), orthology was well supported using syntenic
283 gene structure within contigs or chromosomes (Supplemental Fig. S7).

284

285 *Slfn13* related genes showed a much more recent evolutionary origin, with amplification
286 beginning after the split with rats (~16Ma) and continuing in *Mus* species (Figure 4A). The first
287 duplication gave birth to *Slfn9* and the ancestor of *Slfn8/10*, which subsequently duplicated in
288 the last common ancestor between *Mus pahari* and *Mus musculus* (Figure 4A). This event
289 was quickly followed by a pseudogenization of *Slfn10* in many *Mus musculus domesticus*
290 strains (Supplemental Fig. S8). Phylogenetic analyses, and when possible synteny, supported
291 that *Slfn13* was the parental gene subject to these duplications (Supplemental Fig. S9). This
292 is further supported by the fact that *Slfn11* pseudogenized early during rodent evolution
293 (Figure 4A).

294

295 In sum, we were able to trace the origins of clustered *Schlafen* genes in rodents (leading to
296 ten genes in mouse compared to six in human). We identified a *Slfn12* duplication event
297 accompanied by full domain loss giving rise to GTPase-less clustered *Schlaferns* unique to
298 rodents. These have been retained for at least 50My, indicative of putative functions. We also
299 found an ancient loss of *Slfn11* and, more recently, a burst of *Slfn13* duplications. These likely
300 reflects recent switches in selective pressures since these two genes have been stable during
301 most of mammalian evolution. Next, we explored whether these events were accompanied by
302 signatures of natural selection indicative of recent functional innovations.

303

304 Concerted evolution affecting GTPase-containing *Slfn12* duplicates.

305 To detect putative recombination in rodent *Schlafen* orthologs we used GARD analysis as
306 described in the previous section. This detected several putative recombination segments in
307 all rodent *Schlafen* clades (Supplemental Table S3). We hypothesized this might be due to
308 concerted evolution as many *Schlafen* orthologs failed to segregate according to their
309 orthology in whole CDS phylogenies (Figure 4B).

310

311 Indeed, outside of *Mus* subspecies, *Slfn3*, 4 and 15 paralogs grouped within species and not
312 with their orthologs (Supplemental Fig. S10A). Interestingly, this was not the case for GTPase-
313 less *Slfn1* and *Slfn2* orthologs (Supplemental Fig. S10A). This pattern was also observed in
314 distant species with “unresolved” *Slfn12*-related paralogs (e.g., *Nannospalax galili* or *Jaculus*
315 *jaculus*, Supplemental Fig. S10A). To test whether this was the result of direct gene conversion

316 between paralogs, we compared the genomic identities of the encoding loci (Supplemental
317 Fig.10). This revealed extensive nucleotide identities in the introns and flanking regions of both
318 *Slfn12_4* and *_1* in *Nannosplalax galili* (Supplemental Fig. S10B) and *Slfn3* and *4* in *Cricetulus*
319 *griseus* (Supplemental Fig. S10C). Together these results indicate that GTPase-containing
320 *Slfn12* paralogs have been subject to repeated gene conversion since their birth 50Ma (Figure
321 4A). The selective forces preventing these events from occurring between GTPase-less
322 paralogs, of about the same age, remains unknown.

323

324 Novel signatures of positive selection amongst rodent *Schlafen* genes.

325 For our natural selection analyses we used individual recombination segments of coding CDS
326 (see Methods and previous section). As observed for their primate orthologs, we found that
327 *Slfn14* had the lowest overall dN/dS of all *Schlafen* clades, suggesting purifying selection
328 (Figure 4B, Supplemental Table S3). *Slfn5* orthologs showed pervasive positive selection with
329 several positively selected codons mapping to the helicase domain (Figure 4B, Supplemental
330 Table S3, Supplemental Fig. S11). Thus, *Slfn5* and *Slfn14* are subject to similar selective
331 forces in primates and rodents despite ~87My of divergence. This suggests deeply conserved
332 functions for these two *Schlafen* orthologs including in the context of viral restriction.

333

334 *Slfn12*-related orthologs displayed more complex selective signatures. First, we found that the
335 two GTPase-less paralogs, *Slfn1* and *Slfn2*, evolved under different selective constraints.
336 While *Slfn2* showed signatures of purifying selection (overall dN/dS=0.42, Supplemental Table
337 S3), *Slfn1* orthologs displayed evidence of diversification (dN/dS=0.67, Supplemental Table
338 S3). Codon-level selection analyses revealed that many sites accumulated over *Slfn1* N-
339 terminus with two overlapping the RNase E/Alba domain (Figure 4B, Supplemental Table S3).
340 We currently lack a crystal structure of this protein, and its sequence divergence prevented us
341 from mapping these sites on human SLFN12. However, we used AlphaFold2 to predict SLFN1
342 monomers and dimers (see Methods). While none of the predicted dimers passed our
343 significance threshold, predicted monomers had the expected SLFN structural features (see
344 Methods and Supplemental Fig. S12). When we mapped rapidly evolving sites over this
345 prediction, we found that half were located on the side of the monomer engaged in
346 dimerization in other SCHLAFEN orthologs (L66, S67, S85, P87, V109, N143, Supplemental
347 Fig. S12). This echoes our findings for primate SLFN proteins.

348

349 Second, of all rodent GTPase encoding *Slfn12* paralogs, *Slfn3* showed the strongest signature
350 of selection including positive selection over codons belonging to the GTPase domain (Figure

351 4B, Supplemental Table S3, Supplemental Fig. S13). This contrasts with *Slfn4* orthologs
352 which, albeit weaker evidence for overall selection, accumulated rapidly evolving codons at
353 their N-terminus (Supplemental Table S3, Supplemental Fig. S13). AlphaFold2 predictions of
354 SLFN3 and SLFN4 homodimers matched known SLFN structural features (Supplemental Fig.
355 S13 and Supplemental Fig. S14) (Garvie et al. 2021; Metzner et al. 2022b). Again here, rapidly
356 evolving SLFN4 N-terminal associated sites were enriched at the dimer interface (Figure 3B).
357 Indeed, 3 out of 9 positively selected residues (1/3) locate at this interface, compared to 29
358 out of 602 (~1/21) total residues within 4\AA between the two monomers (chi-square test p-
359 value = 0,0013) (Figure 3A) Finally, we failed to detect evidence for positive selection over
360 *Slfn15* orthologs suggesting they might be evolving primarily under purifying selection, albeit
361 with an average dN/dS=0.6 (Supplemental Table S3). The distinct domain distribution of sites
362 under positive selection between *Slfn3*, *Slfn4* and *Slfn1*, suggests that their parental *Slfn12*
363 gene might have carried multiple functions subject to rapid diversification.

364
365 Considering that most *Slfn13* duplicates arose in *Mus* sub-species, we extended our selection
366 analyses to *Schlafen* orthologs across *Mus musculus musculus* and *domesticus* strains (see
367 Methods). We used whole coding sequences to run McDonald-Kreitman (MK) testing to
368 compare the rates of synonymous to non-synonymous substitutions within *Mus musculus*
369 strains (polymorphisms) to those fixed between species using *Mus Spretus* or *Mus Caroli* as
370 outgroups. This identified *Slfn8* and *Slfn9*, but not *Slfn10*, as evolving under positive selection
371 (Supplemental Table S4). Furthermore, we found whole gene diversifying selection amongst
372 rodent *Slfn13* orthologs (Supplemental Table S4). This suggests that recent *Slfn13* duplicates
373 might share similar selective pressures with their parental gene. Unfortunately, PAML
374 analyses lacked the phylogenetic depth to significantly identify the residues most likely subject
375 to this diversification on rodent *Slfn13* segments or full-length sequences in mouse strains
376 (Supplemental Table S4). Nevertheless, this result is in line with prior reports in mice (Lilue et
377 al. 2018), and indicates that these recent duplicates have continued to diversify their parental
378 gene function.

379
380 In sum, we detected novel events of diversifying selection over recent rodent-specific *Slfn12*
381 duplicates indicating that the birth of *Slfn1/2* and *Slfn3/4* paralogs was accompanied by
382 repeated sub-functionalization. As reported in primates, many rapidly evolving sites
383 accumulated at predicted SLFN interaction surfaces suggesting that SCHLAFEN dimerization
384 is under strong selection in rodents and further supporting its functional importance across
385 mammals.

386

387 DISCUSSION

388

389 Genetic conflicts are key drivers of genome evolution. Most notably, they accelerate genetic
390 innovations over short evolutionary timescales in the form of positive selection and gene family
391 expansion (McLaughlin and Malik 2017). Here, we detected unprecedented signatures of
392 evolutionary diversification of clustered *Schlafen* genes in primate and rodent genomes.
393 *Schlafen* genes engage in at least two genetic conflicts: host/virus restriction in primates; and
394 possibly hybrid incompatibilities in rodents (Cohen-Tannoudji et al. 2000; Bell et al. 2006;
395 Bustos et al. 2009; Lilue et al. 2018; Kim and Weitzman 2022). In the following section, we
396 discuss our findings in the context of these functions.

397

398 Building from pioneering genomics studies, our depth phylogenomic analyses dated the birth
399 of clustered *Schlaferns* to the last common ancestor to all jawed vertebrates, roughly 450Ma
400 (Hedges et al. 2015). We found that a *Slfn11*-like gene was the parental gene of the cluster.
401 SLFN11 is able to restrict retroviruses in several mammalian species suggesting deep antiviral
402 functions (Lin et al. 2016; Stabell et al. 2016; Kim and Weitzman 2022). It is thus tempting to
403 speculate that selective pressures from emerging viruses contributed to *Schlafen* cluster
404 diversification early during vertebrate radiation. This diversification intensified in mammals
405 with the successive birth of *Slfn14*, *Slfn5*, *Slfn13* and finally *Slfn12*. Some of these families,
406 e.g., *Slfn13* and *Slfn12* (Basson et al. 2018; Yang et al. 2018; Garvie et al. 2021; Yue et al.
407 2021), function beyond pathogen restriction indicative of continued sub-functionalization. Our
408 finding of diverging, and sometimes opposite, selective forces over recent amplifications -
409 *SLFN12/12L* and *SLFN11/13* in primates, *Slfn12* duplicates in rodents - further supports this
410 conclusion.

411

412 Our analysis uncovered that one non-clustered *Schlafen* gene, *SLFNL1*, is also ancestral in
413 vertebrate genomes. SLFNL1 has yet to be functionally characterized and shares only that
414 RNase E domain with other SLFNs such as SLFN11. Recently, structurally similar SLFN-like
415 proteins have been shown to help piRNA processing in *C.elegans* (Podvalnaya et al. 2023).
416 Thus, it is possible that orphan genes such as *SLFNL1* might carry out ancestral functions in
417 RNA biogenesis, perhaps related to retrotransposon control (Bartonicek et al. 2022; Ding et
418 al. 2023; Podvalnaya et al. 2023). If so, clustered SLFN antiviral functions might have arose
419 following the addition of regulatory domains such as GTPase and Helicase domains to the
420 "SCHLAFEN core".

421

422 One of the major findings of our study is the pervasive signature of positive selection over both
423 rodent and primate *SLFNs*. Supporting an ongoing genetic conflict with viruses, we found that
424 rapidly evolving codons in primate *SLFN11* and *SLFN5* mapped to known protein domains
425 engaged in viral restriction or subjected to viral antagonism (Stabell et al. 2016; Kim et al.
426 2021). For *SLFN11*, these fall within its RNase pocket. While it is currently unknown whether
427 the *SLFN11* viral antagonist, HCMV protein RL1, specifically targets this domain. Yet, this
428 domain is essential for *SLFN11* function during HIV and HCMV restriction (Stabell et al. 2016;
429 Nightingale et al. 2022). Since we found similar signatures in rodent *Slfn1* and *Slfn4*, it would
430 be interesting to investigate whether RNase activity is also engaged in viral restriction in these
431 species.

432

433 *SLFN5* positive selection mapped to the Helicase domain targeted by HSV-1 protein ICP0
434 supporting a putative antagonistic co-evolution between *SLFN5* and HSV-1 (Kim et al. 2021).
435 Again here, a similar signature is found in rodent *Slfn5* highlighting shared evolutionary
436 constraints. However, this domain is also shown to bind single stranded DNA *in vitro* and is
437 required for retrotransposon restriction in human cells (Metzner et al. 2022b; Ding et al. 2023).
438 Thus, we cannot exclude that additional molecular antagonisms might also drive *SLFN5* rapid
439 evolution. Nevertheless, the diverse immune functions of *Schlafens* family members beyond
440 those reported for *Slfn11* and *Slfn5* unanimously points to host/pathogens arms races being
441 important drivers of their rapid evolution. Thus, future functional investigations of the positively
442 selected sites reported here - most notably over *Slfn12*-related genes - should uncover novel
443 essential interactions shaping mammalian immune responses.

444

445 In addition to expected host/pathogen interactions, we uncovered a novel class of rapidly
446 evolving residues located at the interface between two *SLFN* proteins. So far, we are unaware
447 of studies that directly investigated the consequences of altering *SLFN* protein contact
448 interfaces *in vivo* (Jo and Pommier 2022; Kim and Weitzman 2022). Yet, biochemical work on
449 *SLFN11* and *SLFN12* demonstrated the importance of homodimerization for RNA binding and
450 cleavage likely underlying core *SCHLAFEN* functions (Garvie et al. 2021; Metzner et al.
451 2022b). Thus, our results raise several interesting hypotheses regarding the functional
452 consequences of contact site rapid evolution.

453

454 First, we posit that contact site diversification shapes inter- or intra-*SLFN* compatibilities in the
455 context of rapid paralog expansion (Kuzmin et al. 2022). In this case, the dimerizing surface

456 might be under strong selective pressure to avoid “toxic” inter-paralog dimerization, while
457 keeping intra-ortholog interactions untouched. This type of intragenomic conflict might be
458 particularly important in the context of SLFN functions in hybrid incompatibilities in mice
459 (Cohen-Tannoudji et al. 2000; Bell et al. 2006). This is supported by the recurrent shifts in
460 selection regime we report for rodent *Slfn12* or *Slfn13* duplicates, and by the fact that rapidly
461 evolving residues “face” each other in crystalized or predicted homodimers (SLFN11, SLFN4
462 and SLFN1). Further supporting this model is the observation that SLFN surfaces known to
463 engage with other proteins do not display signatures of diversification (Arslan et al. 2017;
464 Basson et al. 2018; Garvie et al. 2021).

465

466 Another, non-mutually exclusive hypothesis would be that other proteins interfere with
467 dimerization via binding to this surface. One likely candidate is the orthopoxvirus protein p26
468 (v-Slfn poxin) that appears to have acquired a SCHLAFEN-like domain from rodents (Gubser
469 et al. 2007; Bustos et al. 2009). While recent work demonstrated that p26 is key to poxvirus
470 virulence, whether it directly interacts with host Schlafens remains unknown (Eaglesham et al.
471 2020; Hernaez et al. 2020). In both scenario, adaption cycles against “toxic” interactions are
472 expected to trigger evolutionary arms races driving the rapid diversification of *Schlafen* genes.

473

474 In conclusion, recurrent evolutionary innovations of the *Schlafen* gene cluster illustrate
475 genetics conflicts’ profound consequences on genome evolution. Schlafen paralogs have
476 been faced with the challenge of balancing rapid antiviral adaptation with their core cellular
477 functions. Understanding how these evolutionary negotiations influenced their molecular
478 functions should guide future investigations in the context of development, immunity and
479 speciation.

480

481 **METHODS**

482

483 *Identification of vertebrate Schlafen homologs*

484 We performed TBLASTN searches (Altschul et al. 1990) on the NCBI nonredundant nucleotide
485 database of 25 vertebrate genomes (see queried genomes) using translated coding
486 sequences as queries from: (1) human *SLFN5* (NM_144975.4), *SLFN11* (NM_152270.4),
487 *SLFN12* (NM_001289009.2), *SLFN14* (NM_001129820.2), *SLFNL1* (XM_011540945.2); (2)
488 horse *SLFN5* (XM_005597565), *SLFN12* (XM_023653702.1), *SLFN11* (XM_005597562.3),
489 *SLFN14* (XM_023653004.1); and (3) Elephant *SLFN11* (XM_010594583.2), *SLFN14*
490 (XM_023553377.1). We also used Elephant shark and Coelacanth *SLFN11*

491 (XM_007906163.1; XM_014497880.1) or *SLFN1* (XM_007894898.1; XM_014486747.1) as
492 queries to investigate jawless vertebrates and bilaterians. For each search, we collected the
493 longest isoform of hits with more than 40% identity. To search for putative specie-specific
494 homologs, we re-ran TBLASTN searches using within species homologs as queries. Finally,
495 we performed specific TBLASTN searches on whole genome shotgun databases. A gene was
496 considered missing when no homologous sequence (coding or non-coding) could be identified
497 using any of these three approaches. All identified *Schlafen* sequences are available in
498 Supplemental Table S1, raw alignment of all identified vertebrate SCHLAFENS protein in
499 Supplemental Data S1.

500

501 *Pseudogenes and incomplete sequences*

502 *Schlafen* orthologs were considered pseudogenes if we detected stop codons substitutions in
503 their open reading frame or frame shifting mutations disrupting all recognizable homologies in
504 predicted proteins. To detect these events, we used codon alignments with: mouse as
505 reference for rodents; human for primates; and any closely related ortholog for other
506 vertebrates. Partial sequences with missing exons or domains were classified as incomplete.

507

508 *Identification of Primates Schlafens*

509 To identify *SCHLAFEN* homologs in primates, we performed TBLASTN on the NCBI
510 nonredundant nucleotide database in primate genomes (taxid: 9443) using all the translated
511 sequences of human *Schlafen* genes: human *SLFN5* (NM_144975.4), *SLFN11*
512 (NM_152270.4), *SLFN13* (NM_144682.6), *SLFN12* (NM_001289009.2), *SLFN12L*
513 (XM_024450521.1), *SLFN14* (NM_001129820.2), and *SLFN1* (XM_011540945.2). We
514 collected longest isoforms of hits with more than 60 % identity. Raw alignment of all identified
515 primate *SCHLAFEN* coding sequences is available in Supplemental Data S2.

516

517 *Identification of Rodents Schlafens*

518 We used mouse translated sequences to search homologs in rodent species (taxid: 9989) with
519 queries from: *Slfn1* (NM_0114072), *Slfn2* (NM_011408.1), *Slfn3* (NM_001302559.1), *Slfn4*
520 (NM_001302559.1), *Slfn5* (XM_017314615.1), *Slfn8* (NM_1815454.4), *Slfn9*
521 (XM_006533235), *Slfn10-ps* (NR_073523.1), *Slfn14* (NM_001166028.1) and *Slfn15* (refined
522 assembly of XM_487182). For mouse strains, we used TBLASN searches in whole genome
523 assemblies from 13 strains (Ferraj et al. 2023) to identify individual exons and assemble them
524 according to their relative position and orientation in contigs. Raw alignment of all identified

525 rodent *Schlafen* coding sequences is available in Supplemental Data S3 and mouse strains

526 *Slfn13*-related Supplemental Data S4.

527

528 *Queried Genomes*

529 Vertebrates : *Homo sapiens* (UCSC hg38), *Pan troglodytes* (UCSC panTro6), *Gorilla gorilla*
530 (UCSC gorGor6), *Pongo abelii* (UCSC ponAbe3), *Carlito syrichta* (UCSC tarSyr2),
531 *Galeopterus variegatus* (UCSC galVar1), *Marmota marmota* (NCBI GCF_001458135.2),
532 *Rattus norvegicus* (UCSC Rn6), *Mus musculus* (UCSC mm10), *Equus caballus* (UCSC
533 equCab3), *Bos taurus* (UCSC bosTau9), *Felis catus* (UCSC felCat9), *Canis lupus* (canFam4),
534 *Ericaneus europaeus* (UCSC eriEur2), *Myotis lucifugus* (UCSC Myoluc2), *Loxodonta africana*
535 (UCSC loxAfr3), *Dasyurus novemcinctus* (UCSC dasNov3), *Monodelphis domestica* (UCSC
536 MonDom5), *Aquila chrysaetos* (UCSC aquChr2), *Xenopus laevis* (UCSC xenLae2), *Latimeria*
537 *chalumnae* (UCSC LatCha1), *Callorhynchus milii* (UCSC calMil1), *Petromyzon marinus*
538 (UCSC petMar2), *Petromyzon marinus* (UCSC petMar3), *Eptatretus burgeri* (GenBank
539 GCA_024346535.1), *Ciona intestinalis* (UCSC ci3).

540 Additional Primates : *Pan paniscus* (NCBI GCF_029289425.1), *Nomascus leucogenys* (NCBI
541 GCF_006542625.1), *Hylobates moloch* (NCBI GCF_009828535.3), *Colobus angolensis*
542 (NCBI GCF_000951035.1), *Trachypithecus francoisi* (NCBI GCF_009764315.1),
543 *Rhinopithecus roxellana* (NCBI GCF_007565055.1), *Chlorocebus sabaeus* (NCBI
544 GCF_01525025.1), *Theropithecus gelada* (NCBI GCF_003255815.1), *Papio anubis* (NCBI
545 GCF_000264685.3), *Macaca mulatta* (GCF_003339765.1), *Macaca fascicularis* (NCBI
546 GCF_012559485), *Macaca nemestrina* (NCBI GCF_000956065.1), *Cerbocebus atys* (NCBI
547 GCF_000955945.1), *Mandrillus leucophaeus* (NCBI GCF_000951042.1), *Callithrix jacchus*
548 (NCBI GCF_011100555.1), *Cebus imitator* (NCBI GCF_001604975.1), *Saimiri boliviensis*
549 (NCBI GCF_016699345.2), *Aotus nancymaae* (UCSC GCF_000952055.2), *Ptilocolobus*
550 *tephrosceles* (NCBI GCF_002776525.5).

551 Additional Rodents: *Mus spretus* (GenBank GCA_921997135.2), *Mus castaneus* (GenBank
552 GCA_921999005.2), *Mus caroli* (NCBI GCF_900094665.2), *Mus pahari* (NCBI
553 GCF_900095145.1), *Rattus rattus* (NCBI GCF_011064425.1), *Mastomys coucha* (NCBI
554 GCF_008632895.1), *Grammomys surdaster* (NCBI GCF_004785775.1), *Arvicanthis niloticus*
555 (NCBI GCF_011762505.1), *Meriones unguiculatus* (NCBI GCF_002204375.1), *Mesocricetus*
556 *auratus* (NCBI GCF_017639785.1), *Arvicola amphibious* (NCBI GCF_903992535.2), *Microtus*
557 *ochrogaster* (NCBI GCF_000317375.1), *Microtus oregoni* (NCBI GCF_018167655.1),
558 *Nannospalax galili* (NCBI GCF_000622305.1), *Jaculus jaculus* (NCBI GCF_020740685.1),
559 *Octodon degus* (NCBI GCF_000260255.1), *Chinchilla lanigera* (NCBI GCF_020740685.1),

560 *Cavia porcellus* (NCBI GCF_000151735.1), *Ictidomys tridecemlineatus* (NCBI
561 GCF_016881025.1).

562 *Mus musculus* strains (Ferraj et al. 2023): C57/BL6 (GCA_029237415.1), 129S1
563 (GCA_029237445.1), A/J (GCA_029237185.1), C3H/HeJ (GCA_029237405.1), C3H/HeOuj
564 (GCA_029237465.1), BALB/cJ (GCA_029237485.1), BALB/cByJ (GCA_029237445.1), NOD
565 (GCA_029234005.1), NZO/HILtJ (GCA_029233705.1), PWD/PhJ (GCA_029234005.1),
566 PWK/PhJ (GCA_029233695.1), WSB/EiJ (GCA_029233295.1), CAST/EiJ
567 (GCA_029237265.1).

568

569 *Synten analysis*

570 For those genomes without assembled *Schlafen* loci, we determined the location of human
571 *PEX12* (NM_000286.3) and *UNC45B* (NM_001267052.2) in contigs or chromosomes. These
572 genes are slowly evolving and could be mapped in most species queried. We then annotated
573 each *Schlafen* homologs relative to: (1) their position to the flanking genes; (2) other
574 homologs; (3) strand orientation. We repeated the same syntenic analysis *SLFNL1* homologs
575 using: *CTPS1* (NM_001905.4) and *SCMH1* (NM_001394311.1).

576

577 *Alignment and maximum-likelihood phylogenies*

578 Protein and nucleic acid sequences were aligned using MUSCLE 5.1 (Edgar 2004) or MAFFT
579 1.5 (Kato and Standley 2013). We used SMS smart model selection to predict the best model
580 for each alignment (Lefort et al. 2017). Maximum-likelihood phylogenies were built using
581 PhyML 3.3 with either 100 or 500 bootstrap replicates (Guindon et al. 2010). Phylogenetic tree
582 corresponding to SCHLAFEN amino acids sequences in vertebrates was built with JTT model
583 and tree topology optimizing parameters with 500 bootstraps in PhyML. Nucleotides trees was
584 built with the HKY85 model using 100 bootstraps for primate coding sequences and 500
585 bootstraps for rodent coding sequences.

586

587 *Detection of recombination segments*

588 We detected recombination segments in primate and rodent codon alignments using GARD
589 (Kosakovsky Pond et al. 2006) implemented in the HYPHY 2.5.2 package with general
590 discrete model of site to site variation, and three rate classes parameters (Kosakovsky Pond
591 et al. 2020). Recombination segments of <100bp were fused for further analyses.

592

593 *Genomic alignments*

594 To compare the genomic sequences of *Slfn12*-related paralogs in *Nannospalax galili* and
595 *Cricetulus griseus* genomes, we extracted ~16Kb of genomic sequence encompassing the
596 coding loci of *Slfn3* and *Slfn4* or *Slfn12_4* and *Slfn12_1*. We aligned these sequences using
597 the LAGAN global pair-wise alignment program (Brudno et al. 2003) and visualized them using
598 mVISTA (Frazer et al. 2004).

599

600 *Selection analyses*

601 Codon-level dN/dS tests were done using codeml “NSsites”, PAML 4.0 (Yang 2007), with gene
602 trees of degapped Phylip codon alignments. First, we compared NSsites models for evidence
603 of natural selection on full length CDS (“whole gene”, Supplemental Table S2, Supplemental
604 Table S3 and Supplemental Table S4). Statistical significance was measured with χ^2 test of
605 twice the difference in log-likelihood between M0 (calculated dN/dS) or its null M0=1 (dN/dS
606 fixed to 1; genetic drift) with 1 degree of freedom. Genes for which we could reject the null (at
607 $p\text{-val}\leq 0.05$), were tested for positive selection. For this, we used alignments from predicted
608 recombination segments to account for dN/dS variations between gene fragments. In each
609 case, we compared the likelihood of model M8 (allowing for dN/dS>1) with two null models:
610 M7 (not allowing dN/dS>1); and M8a (dN/dS fixed to 1). We used likelihood ratio testing with
611 χ^2 to calculate a p-value for positive selection (1 degree of freedom for M8a, 2 for M7).
612 Segments were considered to be subject to positive selection at $p\text{-val}\leq 0.05$. Sites considered
613 to be under positive selection are those with M8 Bayes Empirical Bayes posterior probability
614 $\geq 99\%$. Results were consistent between codon frequency 2 or 3 and starting omega of 0,4
615 or 1,5. All the PAML results presented in Supplemental Table S2 and Supplemental Table S3
616 are from codon frequency of 2 with starting omega of 0,4. Amino acid coordinates are relative
617 to mouse or human full length CDS orthologs.

618 We also performed FUBAR analyses on degapped alignments from predicted recombination
619 segments using universal code. Sites considered to be under positive selection are those with
620 a posterior probability ≥ 0.95 . (Murrell et al. 2013; Weaver et al. 2018).

621

622 *McDonald-Kreitman testing*

623 McDonald-Kreitman test (<http://mkt.uab.es/mkt/MKT.asp>) on full length degapped CDS
624 alignments from *Slfn8*, *Slfn9* and *Slfn10* using *Mus spretus* as an outgroup for *Slfn9* and *Mus*
625 *caroli* as outgroup for *Slfn8* and *Slfn10*.

626

627 *Species phylogenies*

628 Species divergence time and trees were calculated using timetree.org (Kumar et al. 2017).

629 Species trees were also built using timetree.org specifying a group of species in a list. For
630 more detailed rodent and primate species trees we used previously published phylogenies
631 (Perelman et al. 2011; Stepan and Schenk 2017; Swanson et al. 2019). Mouse strains
632 phylogeny/pedigree was inferred from recent haplotype mapping (Kirby et al. 2010).

633

634 *Protein domain annotations*

635 Domain corresponding to (1) the Schlafen core domain (AlbA and Sln-box); (2) the GTPase-
636 like linker domain (SWADL); and (3) the helicase domain (P-loop NTPase and Walker A and
637 B) were annotated on Human and Mouse proteins according to published structures and
638 predictions (Chen and Kuhn 2019; Garvie et al. 2021; Jo and Pommier 2022; Metzner et al.
639 2022a; Metzner et al. 2022b).

640

641 *Protein structure prediction*

642 Structure predictions of SLFN1, SLFN3 and SLFN4 were calculated with the ColabFold v1.5.2
643 package which uses AlphaFold2 (Jumper et al. 2021; Mirdita et al. 2022). We used
644 SCHLAFEN protein sequences as queries with default parameters, processed with mmseq2
645 alignment with the model type set on auto (Steinegger and Soding 2017). To predict
646 homodimer, we used “:” to specify inter-protein chain breaks for modelling dimers and the
647 model type set on alphafold2_multimer_v3. For each protein, we selected the best ranking
648 model based on the structure alignment confidence and evaluated interface confidence using
649 “predicted aligned error” (PAE) measured in Angströms (Å). We rejected dimer predictions for
650 SLFN1 since all the residues had PAE value > 20Å. Conversely, we retained the prediction of
651 SLFN3 and SLFN4 homodimers because most chain A and B residues were < 5Å.

652

653 *Structure analysis*

654 We used published pdb (protein data bank format) for human SLFN5 (Metzner et al. 2022a),
655 SLFN11 (Metzner et al. 2022b) and SLFN12 (Garvie et al. 2021) or predicted pdb for rodents.
656 All structures were visualized using ChimeraX (Pettersen et al. 2021). For SLFN5, high
657 sequence similarity allowed us to map rodent positively selected sites on the human structure.
658 To identify of protein interfaces, we used the ChimeraX “contact” option with a 4Å threshold
659 between crystalized or predicted dimer chains.

660

661 **DATA ACCESS**

662 All data used and reported in this study can be downloaded freely in public repositories listed
663 in the methods.

664

665 **COMPETING INTEREST STATEMENT**

666 The authors declare no competing interests.

667

668 **ACKNOWLEDGMENTS**

669 We would like to thank Dr. Lucie Etienne, Dr. Deborah Bourc'his and members of the Molaro
670 laboratory for helpful discussions. This work was funded by the Agence Nationale de la
671 Recherche (ANR): ANR-22-CE12-0006-01; the Fondation pour la Recherche Médicale (FRM):
672 AJE201912009932; and the Institute of Genetics, Reproduction and Development (iGReD).

673

674

675 **REFERENCES**

676

677 Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol*
678 *Biol* **215**: 403-410.

679 Arslan AD, Sassano A, Saleiro D, Lisowski P, Kosciuczuk EM, Fischietti M, Eckerdt F, Fish EN,
680 Platanias LC. 2017. Human SLFN5 is a transcriptional co-repressor of STAT1-mediated
681 interferon responses and promotes the malignant phenotype in glioblastoma. *Oncogene* **36**:
682 6006-6019.

683 Bartonicek N, Rouet R, Warren J, Loetsch C, Rodriguez GS, Walters S, Lin F, Zahra D, Blackburn J,
684 Hammond JM et al. 2022. The retroelement Lx9 puts a brake on the immune response to virus
685 infection. *Nature* **608**: 757-765.

686 Basson MD, Wang Q, Chaturvedi LS, More S, Vomhof-DeKrey EE, Al-Marsoummy S, Sun K, Kuhn LA,
687 Kovalenko P, Kiupel M. 2018. Schlafen 12 Interaction with SerpinB12 and Deubiquitylases
688 Drives Human Enterocyte Differentiation. *Cell Physiol Biochem* **48**: 1274-1290.

689 Bell TA, de la Casa-Esperon E, Doherty HE, Ideraabdullah F, Kim K, Wang Y, Lange LA, Wilhemsen
690 K, Lange EM, Sapienza C et al. 2006. The paternal gene of the DDK syndrome maps to the
691 Schlafen gene cluster on mouse chromosome 11. *Genetics* **172**: 411-423.

692 Brudno M, Do CB, Cooper GM, Kim MF, Davydov E, Program NCS, Green ED, Sidow A, Batzoglu S.
693 2003. LAGAN and Multi-LAGAN: efficient tools for large-scale multiple alignment of genomic
694 DNA. *Genome Res* **13**: 721-731.

695 Bustos O, Naik S, Ayers G, Casola C, Perez-Lamigueiro MA, Chippindale PT, Pritham EJ, de la Casa-
696 Esperon E. 2009. Evolution of the Schlafen genes, a gene family associated with embryonic
697 lethality, meiotic drive, immune processes and orthopoxvirus virulence. *Gene* **447**: 1-11.

698 Chen J, Kuhn LA. 2019. Deciphering the three-domain architecture in schlafens and the structures and
699 roles of human schlafen12 and serpinB12 in transcriptional regulation. *J Mol Graph Model* **90**:
700 59-76.

701 Cohen-Tannoudji M, Vandormael-Pournin S, Le Bras S, Coumailleau F, Babinet C, Baldacci P. 2000.
702 A 2-Mb YAC/BAC-based physical map of the ovum mutant (Om) locus region on mouse
703 chromosome 11. *Genomics* **68**: 273-282.

704 Crespi B, Nosil P. 2013. Conflictual speciation: species formation via genomic conflict. *Trends Ecol Evol*
705 **28**: 48-57.

706 Daugherty MD, Schaller AM, Geballe AP, Malik HS. 2016. Evolution-guided functional analyses reveal
707 diverse antiviral specificities encoded by IFIT1 genes in mammals. *Elife* **5**.

708 Daugherty MD, Zanders SE. 2019. Gene conversion generates evolutionary novelty that fuels genetic
709 conflicts. *Curr Opin Genet Dev* **58-59**: 49-54.

710 Ding J, Wang S, Liu Q, Duan Y, Cheng T, Ye Z, Cui Z, Zhang A, Liu Q, Zhang Z et al. 2023. Schlafen-
711 5 inhibits LINE-1 retrotransposition. *iScience* **26**: 107968.

- 712 Ding J, Wang S, Wang Z, Chen S, Zhao J, Solomon M, Liu Z, Guo F, Ma L, Wen J et al. 2022. Schlafen
713 5 suppresses human immunodeficiency virus type 1 transcription by commandeering cellular
714 epigenetic machinery. *Nucleic Acids Res* **50**: 6137-6153.
- 715 Eaglesham JB, McCarty KL, Kranzusch PJ. 2020. Structures of diverse poxins cGAMP nucleases reveal
716 a widespread role for cGAS-STING evasion in host-pathogen conflict. *Elife* **9**.
- 717 Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput.
718 *Nucleic Acids Res* **32**: 1792-1797.
- 719 Ferraj A, Audano PA, Balachandran P, Czechanski A, Flores JI, Radecki AA, Mosur V, Gordon DS,
720 Walawalkar IA, Eichler EE et al. 2023. Resolution of structural variation in diverse mouse
721 genomes reveals chromatin remodeling due to transposable elements. *Cell Genom* **3**: 100291.
- 722 Fischietti M, Arslan AD, Sassano A, Saleiro D, Majchrzak-Kita B, Ebine K, Munshi HG, Fish EN,
723 Plataniias LC. 2018. Slfn2 Regulates Type I Interferon Responses by Modulating the NF-
724 kappaB Pathway. *Mol Cell Biol* **38**.
- 725 Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I. 2004. VISTA: computational tools for
726 comparative genomics. *Nucleic Acids Res* **32**: W273-279.
- 727 Gardner A, Ubeda F. 2017. The meaning of intragenomic conflict. *Nat Ecol Evol* **1**: 1807-1815.
- 728 Garvie CW, Wu X, Papanastasiou M, Lee S, Fuller J, Schnitzler GR, Horner SW, Baker A, Zhang T,
729 Mullahoo JP et al. 2021. Structure of PDE3A-SLFN12 complex reveals requirements for
730 activation of SLFN12 RNase. *Nat Commun* **12**: 4375.
- 731 Geserick P, Kaiser F, Klemm U, Kaufmann SH, Zerrahn J. 2004. Modulation of T cell development and
732 activation by novel members of the Schlafen (slfn) gene family harbouring an RNA helicase-
733 like motif. *Int Immunol* **16**: 1535-1548.
- 734 Gubser C, Goodbody R, Ecker A, Brady G, O'Neill LAJ, Jacobs N, Smith GL. 2007. Camelpox virus
735 encodes a schlafen-like protein that affects orthopoxvirus virulence. *J Gen Virol* **88**: 1667-1676.
- 736 Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and
737 methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML
738 3.0. *Syst Biol* **59**: 307-321.
- 739 Hedges SB, Marin J, Suleski M, Paymer M, Kumar S. 2015. Tree of life reveals clock-like speciation
740 and diversification. *Mol Biol Evol* **32**: 835-845.
- 741 Hernaez B, Alonso G, Georgana I, El-Jesr M, Martin R, Shair KHY, Fischer C, Sauer S, Maluquer de
742 Motes C, Alcami A. 2020. Viral cGAMP nuclease reveals the essential role of DNA sensing in
743 protection against acute lethal virus infection. *Sci Adv* **6**.
- 744 Hurst GD, Werren JH. 2001. The role of selfish genetic elements in eukaryotic evolution. *Nat Rev Genet*
745 **2**: 597-606.
- 746 Jo U, Pommier Y. 2022. Structural, molecular, and functional insights into Schlafen proteins. *Exp Mol*
747 *Med* **54**: 730-738.
- 748 Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R,
749 Zidek A, Potapenko A et al. 2021. Highly accurate protein structure prediction with AlphaFold.
750 *Nature* **596**: 583-589.
- 751 Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements
752 in performance and usability. *Mol Biol Evol* **30**: 772-780.
- 753 Kim ET, Dybas JM, Kulej K, Reyes ED, Price AM, Akhtar LN, Orr A, Garcia BA, Boutell C, Weitzman
754 MD. 2021. Comparative proteomics identifies Schlafen 5 (SLFN5) as a herpes simplex virus
755 restriction factor that suppresses viral transcription. *Nat Microbiol* **6**: 234-245.
- 756 Kim ET, Weitzman MD. 2022. Schlafens Can Put Viruses to Sleep. *Viruses* **14**.
- 757 Kirby A, Kang HM, Wade CM, Cotsapas C, Kostem E, Han B, Furlotte N, Kang EY, Rivas M, Bogue
758 MA et al. 2010. Fine mapping in 94 inbred mouse strains using a high-density haplotype
759 resource. *Genetics* **185**: 1081-1095.
- 760 Kosakovsky Pond SL, Poon AFY, Velazquez R, Weaver S, Hepler NL, Murrell B, Shank SD, Magalis
761 BR, Bouvier D, Nekrutenko A et al. 2020. HyPhy 2.5-A Customizable Platform for Evolutionary
762 Hypothesis Testing Using Phylogenies. *Mol Biol Evol* **37**: 295-299.
- 763 Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH, Frost SD. 2006. GARD: a genetic algorithm
764 for recombination detection. *Bioinformatics* **22**: 3096-3098.
- 765 Kumar S, Stecher G, Suleski M, Hedges SB. 2017. TimeTree: A Resource for Timelines, Timetrees,
766 and Divergence Times. *Mol Biol Evol* **34**: 1812-1819.
- 767 Kuzmin E, Taylor JS, Boone C. 2022. Retention of duplicated genes in evolution. *Trends Genet* **38**: 59-
768 72.
- 769 Lefort V, Longueville JE, Gascuel O. 2017. SMS: Smart Model Selection in PhyML. *Mol Biol Evol* **34**:
770 2422-2424.

- 771 Li M, Kao E, Gao X, Sandig H, Limmer K, Pavon-Eternod M, Jones TE, Landry S, Pan T, Weitzman
772 MD et al. 2012. Codon-usage-based inhibition of HIV protein synthesis by human schlafen 11.
773 *Nature* **491**: 125-128.
- 774 Lilue J, Doran AG, Fiddes IT, Abrudan M, Armstrong J, Bennett R, Chow W, Collins J, Collins S,
775 Czechanski A et al. 2018. Sixteen diverse laboratory mouse reference genomes define strain-
776 specific haplotypes and novel functional loci. *Nat Genet* **50**: 1574-1583.
- 777 Lin YZ, Sun LK, Zhu DT, Hu Z, Wang XF, Du C, Wang YH, Wang XJ, Zhou JH. 2016. Equine schlafen
778 11 restricts the production of equine infectious anemia virus via a codon usage-dependent
779 mechanism. *Virology* **495**: 112-121.
- 780 McLaughlin RN, Jr., Malik HS. 2017. Genetic conflicts: the usual suspects and beyond. *J Exp Biol* **220**:
781 6-17.
- 782 Metzner FJ, Huber E, Hopfner KP, Lammens K. 2022a. Structural and biochemical characterization of
783 human Schlafen 5. *Nucleic Acids Res* **50**: 1147-1161.
- 784 Metzner FJ, Wenzl SJ, Kugler M, Krebs S, Hopfner KP, Lammens K. 2022b. Mechanistic understanding
785 of human SLFN11. *Nat Commun* **13**: 5464.
- 786 Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. 2022. ColabFold: making
787 protein folding accessible to all. *Nat Methods* **19**: 679-682.
- 788 Mitchell PS, Young JM, Emerman M, Malik HS. 2015. Evolutionary Analyses Suggest a Function of
789 MxB Immunity Proteins Beyond Lentivirus Restriction. *PLoS Pathog* **11**: e1005304.
- 790 Molaro A, Malik HS, Bourc'his D. 2020. Dynamic Evolution of De Novo DNA Methyltransferases in
791 Rodent and Primate Genomes. *Mol Biol Evol* **37**: 1882-1892.
- 792 Murai J, Tang SW, Leo E, Baechler SA, Redon CE, Zhang H, Al Abo M, Rajapakse VN, Nakamura E,
793 Jenkins LMM et al. 2018. SLFN11 Blocks Stressed Replication Forks Independently of ATR.
794 *Mol Cell* **69**: 371-384 e376.
- 795 Murrell B, Moola S, Mabona A, Weighill T, Sheward D, Kosakovsky Pond SL, Scheffler K. 2013.
796 FUBAR: a fast, unconstrained bayesian approximation for inferring selection. *Mol Biol Evol* **30**:
797 1196-1205.
- 798 Nightingale K, Potts M, Hunter LM, Fielding CA, Zerbe CM, Fletcher-Etherington A, Nobre L, Wang
799 EY, Strang BL, Houghton JW et al. 2022. Human cytomegalovirus protein RL1 degrades the
800 antiviral factor SLFN11 via recruitment of the CRL4 E3 ubiquitin ligase complex. *Proc Natl Acad
801 Sci U S A* **119**.
- 802 Perelman P, Johnson WE, Roos C, Seuanez HN, Horvath JE, Moreira MA, Kessing B, Pontius J, Roelke
803 M, Rumpler Y et al. 2011. A molecular phylogeny of living primates. *PLoS Genet* **7**: e1001342.
- 804 Pettersen EF, Goddard TD, Huang CC, Meng EC, Couch GS, Croll TI, Morris JH, Ferrin TE. 2021.
805 UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein
806 Sci* **30**: 70-82.
- 807 Pisareva VP, Muslimov IA, Tcherepanov A, Pisarev AV. 2015. Characterization of Novel Ribosome-
808 Associated Endoribonuclease SLFN14 from Rabbit Reticulocytes. *Biochemistry* **54**: 3286-3301.
- 809 Podvalnaya N, Bronkhorst AW, Lichtenberger R, Hellmann S, Nischwitz E, Falk T, Karaulanov E, Butter
810 F, Falk S, Ketting RF. 2023. piRNA processing by a trimeric Schlafen-domain nuclease. *Nature*
811 **622**: 402-409.
- 812 Schwarz DA, Katayama CD, Hedrick SM. 1998. Schlafen, a new family of growth regulatory genes that
813 affect thymocyte development. *Immunity* **9**: 657-668.
- 814 Seong RK, Seo SW, Kim JA, Fletcher SJ, Morgan NV, Kumar M, Choi YK, Shin OS. 2017. Schlafen 14
815 (SLFN14) is a novel antiviral factor involved in the control of viral replication. *Immunobiology*
816 **222**: 979-988.
- 817 Stabell AC, Hawkins J, Li M, Gao X, David M, Press WH, Sawyer SL. 2016. Non-human Primate
818 Schlafen11 Inhibits Production of Both Host and Viral Proteins. *PLoS Pathog* **12**: e1006066.
- 819 Steinegger M, Soding J. 2017. MMseqs2 enables sensitive protein sequence searching for the analysis
820 of massive data sets. *Nat Biotechnol* **35**: 1026-1028.
- 821 Steppan SJ, Schenk JJ. 2017. Muroid rodent phylogenetics: 900-species tree reveals increasing
822 diversification rates. *PLoS One* **12**: e0183070.
- 823 Swanson MT, Oliveros CH, Esselstyn JA. 2019. A phylogenomic rodent tree reveals the repeated
824 evolution of masseter architectures. *Proc Biol Sci* **286**: 20190672.
- 825 Weaver S, Shank SD, Spielman SJ, Li M, Muse SV, Kosakovsky Pond SL. 2018. Datamonkey 2.0: A
826 Modern Web Application for Characterizing Selective and Other Evolutionary Processes. *Mol
827 Biol Evol* **35**: 773-777.
- 828 Yang JY, Deng XY, Li YS, Ma XC, Feng JX, Yu B, Chen Y, Luo YL, Wang X, Chen ML et al. 2018.
829 Structure of Schlafen13 reveals a new class of tRNA/rRNA- targeting RNase engaged in
830 translational control. *Nat Commun* **9**: 1165.

- 831 Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* **24**: 1586-1591.
- 832 Yue T, Zhan X, Zhang D, Jain R, Wang KW, Choi JH, Misawa T, Su L, Quan J, Hildebrand S et al.
- 833 2021. SLFN2 protection of tRNAs from stress-induced cleavage is essential for T cell-mediated
- 834 immunity. *Science* **372**.
- 835 Zoppoli G, Regairaz M, Leo E, Reinhold WC, Varma S, Ballestrero A, Doroshov JH, Pommier Y. 2012.
- 836 Putative DNA/RNA helicase Schlafen-11 (SLFN11) sensitizes cancer cells to DNA-damaging
- 837 agents. *Proc Natl Acad Sci U S A* **109**: 15030-15035.
- 838