



HAL
open science

In silico reconstruction of a salmonid alphavirus virion reveals distinctive molecular features implicated in virulence

Stéphane Biacchesi, Calvin Fauvet, Emilie Mérour, Julie Bernard, Annie Lamoureux, Delphine Lallias, Jean K Millet

► **To cite this version:**

Stéphane Biacchesi, Calvin Fauvet, Emilie Mérour, Julie Bernard, Annie Lamoureux, et al.. In silico reconstruction of a salmonid alphavirus virion reveals distinctive molecular features implicated in virulence. *iScience*, 2026, 29 (3), pp.115070. <10.1016/j.isci.2026.115070>. <hal-05553990>

HAL Id: hal-05553990

<https://hal.inrae.fr/hal-05553990v1>

Submitted on 16 Mar 2026

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.

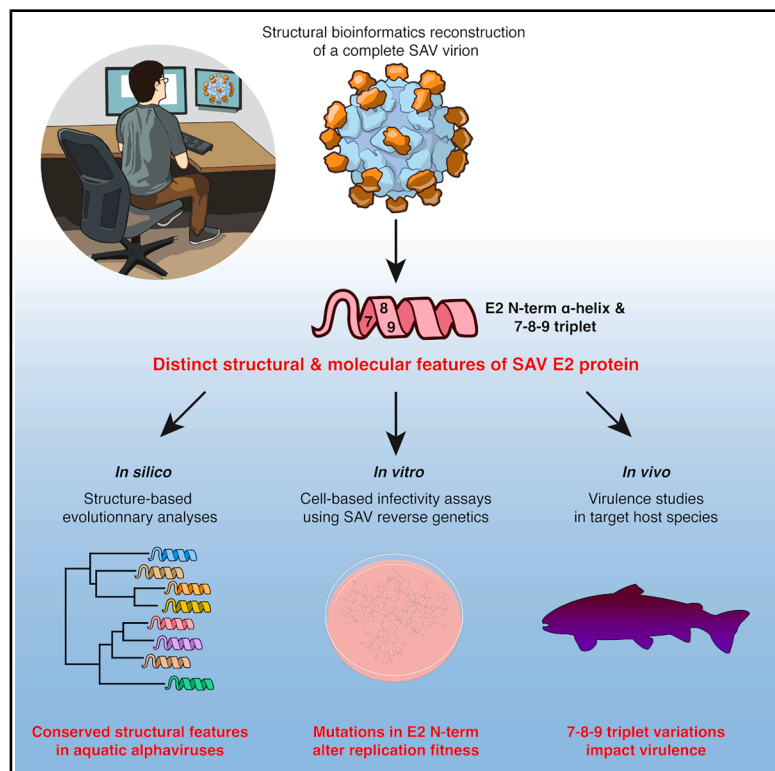
L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

In silico reconstruction of a salmonid alphavirus virion reveals distinctive molecular features implicated in virulence

Graphical abstract



Authors

Stéphane Biacchesi, Calvin Fauvet, Emilie Mérour, Julie Bernard, Annie Lamoureux, Delphine Lallias, Jean K. Millet

Correspondence

jean.millet@inrae.fr

In brief

Virology; structural biology; bioinformatics

Highlights

- Reconstruction of an entire salmonid alphavirus virion by structural bioinformatics
- Discovery of an exposed α -helix and 7-8-9 triplet motif at E2 protein N-terminus
- The E2 α -helix is a conserved feature of aquatic alphaviruses
- Mutations in E2 N-terminus alter viral fitness *in vitro* and virulence *in vivo*



Article

In silico reconstruction of a salmonid alphavirus virion reveals distinctive molecular features implicated in virulence

Stéphane Biacchesi,¹ Calvin Fauvet,¹ Emilie Mérour,^{1,3} Julie Bernard,¹ Annie Lamoureux,¹ Delphine Lallias,² and Jean K. Millet^{1,4,*}

¹Université Paris-Saclay, INRAE, UVSQ, Virologie et Immunologie Moléculaires, Jouy-en-Josas, France

²Université Paris-Saclay, INRAE, AgroParisTech, GABI, Jouy-en-Josas, France

³Present address: Université Paris-Saclay, CNRS, Institut de Chimie des Substances Naturelles, UPR 2301, Gif-sur-Yvette, France

⁴Lead contact

*Correspondence: jean.millet@inrae.fr

<https://doi.org/10.1016/j.isci.2026.115070>

SUMMARY

Salmonid alphavirus (SAV) poses a significant disease threat to aquaculture. Recently, new alphaviruses from several fish species have been discovered. However, little is known about their biology and pathogenicity. Alphaviruses are considered to have originated from a marine environment; therefore, studying fish alphaviruses can inform on the evolutionary history of the genus. Contrary to many terrestrial alphaviruses, there are currently no experimentally determined structures for aquatic alphaviruses, severely limiting their study. In this work, we harness the power of structural bioinformatics and AlphaFold to reconstruct an entire SAV virion, thereby revealing an exposed and distinctive α -helical feature in its E2 envelope protein. Using an integrative approach, we explore the sequence diversity and evolutionary conservation of this predicted feature and investigate the functional consequences of variations on viral fitness and virulence. This study provides a framework paving the way to better understand aquatic alphavirus pathogenicity and host species adaptation.

INTRODUCTION

Epizootics caused by emerging and re-emerging viruses are a serious threat to the development of aquaculture which is recognized as the fastest-growing food sector. Salmonid alphavirus (SAV) is a WOAHL-listed viral pathogen that causes sleeping disease (SD) in rainbow trout (*Oncorhynchus mykiss*) and pancreas disease in Atlantic salmon (*Salmo salar*).¹

Members of the *Alphavirus* genus within the *Togaviridae* family are RNA viruses found in diverse aquatic and terrestrial ecosystems and infect invertebrates, fishes, birds, and mammals. Most terrestrial alphaviruses are arboviruses posing a significant threat for human and animal health. A landmark genome-wide study suggested that alphaviruses originated from a marine environment.² This was confirmed by a large scale metatranscriptomics study, which demonstrated that fish-infecting viral species within distinct genera of RNA viruses, such as alphaviruses and rhabdoviruses, form lineages basal to vector-borne terrestrial representatives.³ Recently discovered alphaviruses include viral species infecting several fish hosts such as striated frogfish (*Antennarius striatus*), crested flounder (*Plagiopsetta* sp.), jawless inshore hagfish (*Eptatretus burgeri*), and Mediterranean comber (*Serranus cabrilla*).^{3,4} Whether these viruses found in wild reservoir species can pose a disease threat to aquaculture species remains to be addressed. Despite these advances in virus discovery, little is known about their biology,

the disease they cause, and their relationship to SAV and terrestrial alphaviruses.

SAV is the best-studied aquatic alphavirus and like other members of its genus it has a positive-sense non-segmented RNA genome (11–12 kb range) that is 5' capped and 3' polyadenylated. Its genome encodes four non-structural proteins (nsP1–4) which are first synthesized as a polyprotein which is then proteolytically processed. Alphavirus nsP proteins mediate viral transcription, replication, and counteract host innate immune responses. A subgenomic RNA encodes the structural polyprotein composed of capsid (Cp), E3, E2, 6k, transframe (TF), and E1 proteins, which are also cleaved into individual proteins by proteolytic enzymes. As their name implies, the structural proteins participate in virion formation and can also suppress or antagonize host innate immunity. There are six major subtypes of SAV, numbered 1 to 6, based on partial sequences of nsP3 and E2. Distinctive SAV features setting them apart from terrestrial alphaviruses include shorter non-coding regions at the 5' and 3' genomic ends, a low temperature of replication (10°C–15°C), and transmission occurring directly through water without the need for a vector, although SAV nucleic acids have been detected in sea lice (*Lepeophtheirus salmonis*).^{5–7} Another unique characteristic of SAV is the longer length of the E1 and E2 envelope proteins, due to the presence of relatively long (~10 aa) inserts. The role of these inserts and more broadly the structure-function relationships of fish alphavirus proteins constitute major



limitations in our understanding of their biology and evolutionary relationship with their terrestrial counterparts.

Much of what is known about alphavirus structure stems from studies on terrestrial representatives such as chikungunya virus (CHIKV)⁸ and Venezuelan equine encephalitis virus (VEEV).⁹ Although enveloped, the surface of alphavirus virions (~70 nm) adopts an ordered $T = 4$ icosahedral symmetry architecture composed of 80 trimeric spikes with each trimer made of heterodimers of E2 and E1, which are both transmembrane proteins and the main determinants of host cell receptor attachment and membrane fusion, respectively.¹⁰ Within the alphavirus virion, the Cp protein encapsidates the RNA genome forming an inner nucleocapsid shell.⁹ To date, there are no available experimentally determined structures for SAV or any other fish alphavirus representatives. The lack of widely accessible structural data limits the depth of structure-function analyses of these viral proteins. Since E1 and E2 proteins are the main drivers of viral entry and host range, the absence of structural information precludes a deeper understanding of the molecular determinants of virulence, host specificity, and interspecies transmission.

From a functional standpoint, our group has previously demonstrated that single amino acid substitutions within the E2 envelope protein profoundly impact virulence in rainbow trout, mirroring similar phenomena observed with terrestrial alphaviruses.¹¹ A comparable attenuation was described for CHIKV strain 181/clone 25, a live-attenuated vaccine candidate derived from a south-east Asian human isolate, where only two E2 mutations in domains responsible for receptor recognition were found to alter virulence.¹² A single substitution (A226V) in CHIKV E1 fusion protein identified in an epidemic strain from La Réunion island was shown to be responsible for a significant increase in infectivity in *Aedes albopictus*, providing a plausible explanation for how this variant caused an epidemic in a region lacking its typical vector, *A. aegypti*.¹³ Moreover, our group has previously demonstrated that single substitutions within an epitope motif of SAV2 E2 were associated with altered virulence *in vivo*.¹⁴ Similar mutational sensitivity (i.e., small amino acid changes in envelope proteins leading to large and measurable phenotypical effects) has also been described for other alphaviruses, including VEEV, Sindbis virus (SINV), and Semliki Forest virus (SFV).^{15–17} The findings that such mutational sensitivity is shared among several terrestrial and aquatic alphaviruses suggest it could be a genus-wide conserved trait. Since SAV occupies a basal position within *Alphavirus* genus phylogeny, its study can not only inform the biology of fish alphaviruses but also the evolutionary relationships and commonalities with terrestrial alphaviruses.

Recently, structural bioinformatics has undergone a revolution with the emergence of artificial intelligence-based algorithms capable of highly accurate protein structure predictions, as exemplified by AlphaFold.^{18,19} The development of these tools is particularly useful for virology as, until recently, a huge proportion of viral protein sequences had no corresponding structural information. Importantly, these predictive programs open up the possibility to conduct structure-guided studies on viral proteins that previously lacked experimentally obtained structures.

Here, we leverage the groundbreaking advances brought by AlphaFold and other structural bioinformatics tools by modeling the structural proteins of SAV and reconstructing the structure of a complete virion. Using an integrative strategy, we uncover a unique α -helical structural feature found at a site near the N-terminus (N-term) of E2 that is shared among aquatic alphaviruses. By combining phylogenetic analysis, reverse genetics and *in vitro* and *in vivo* assays, we delve into the sequence diversity of a broad set of SAV isolates to investigate how variations found at this site can impact replicative fitness and virulence in rainbow trout. This work will allow to anticipate the functional impacts of virulence-associated mutations in a better way, which may inform surveillance strategies for disease outbreak investigation and ultimately offer insights into the evolutionary dynamics of alphaviruses.

RESULTS

Phylogeny of SAV structural proteins

To position SAV and more recently identified fish alphaviruses within the context of the *Alphavirus* genus, the sequences of structural proteins of 17 alphaviruses were aligned to generate a maximum-likelihood (ML) phylogenetic tree of representatives from both aquatic and terrestrial environments (Figure 1A and Table S1). The overall topology recapitulates previously established phylogenetic relationships among the *Alphavirus* genus and highlights the branching of fish-infecting alphaviruses forming a distinct basal clade from that of the other alphaviruses. This analysis is in line with the evidence for the ancient evolutionary history of Alphaviruses and their proposed marine origin.^{2,3} Notably, aquatic alphaviruses that infect mammalian hosts, such as southern elephant seal virus (SESV) and Alaskan harbor porpoise alphavirus (AHPV) are more closely related to terrestrial alphaviruses from the SFV antigenic complex group, with a solidly supported branching (bootstrap support of 97).

Among fish-infecting aquatic alphaviruses, most branch nodes are strongly supported (bootstrap support of 100), with only one node with a relatively lower value (bootstrap support of 63) corresponding to the branching between SAV (ICTV species *Alphavirus salmon*) and Wenling hagfish alphavirus (WHAV), which is most closely related to SAV. SAV infects Atlantic salmon (*Salmo salar*), rainbow trout (*Oncorhynchus mykiss*), Arctic charr (*Salvelinus alpinus*) and common dab (*Limanda limanda*). Among species closely related to SAV, WHAV was isolated from inshore hagfish (*Eptatretus burgeri*), and is followed by Comber alphavirus (CAV) which was isolated from *Serranus cabrilla*. A distinct clade associated with a robustly supported node (bootstrap support of 100) is formed by two very closely related viruses Wenling crested flounder alphavirus (WCFV) and Wenling striated frogfish alphavirus (WSFV), isolated from *Plagiopsetta* sp. and *Antennarius striatus*, respectively.

Several terrestrial alphaviruses, including representatives of the major terrestrial antigenic groupings (VEE, EEE, WEE, and SFV), have been extensively studied structurally. This is attested by the abundance of experimentally determined whole-virion structures available in the Protein DataBank (PDB, Figure 1A, right). Importantly, to date there are no experimentally

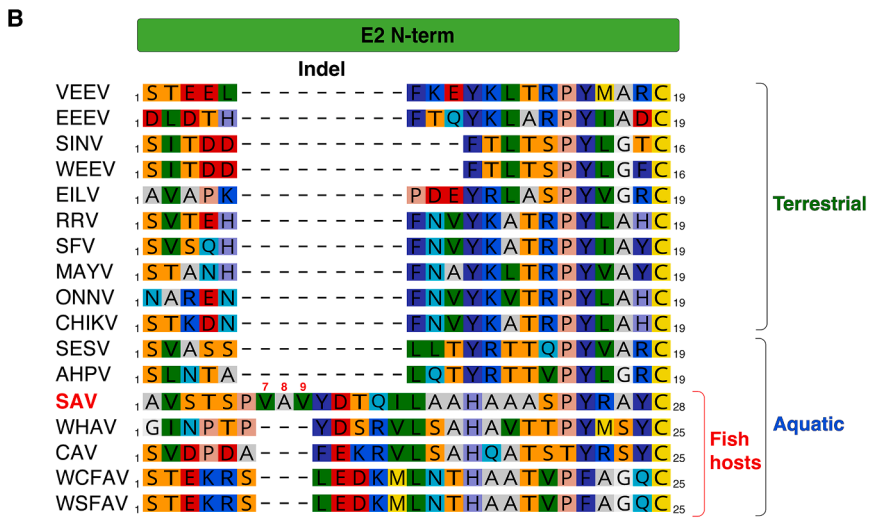
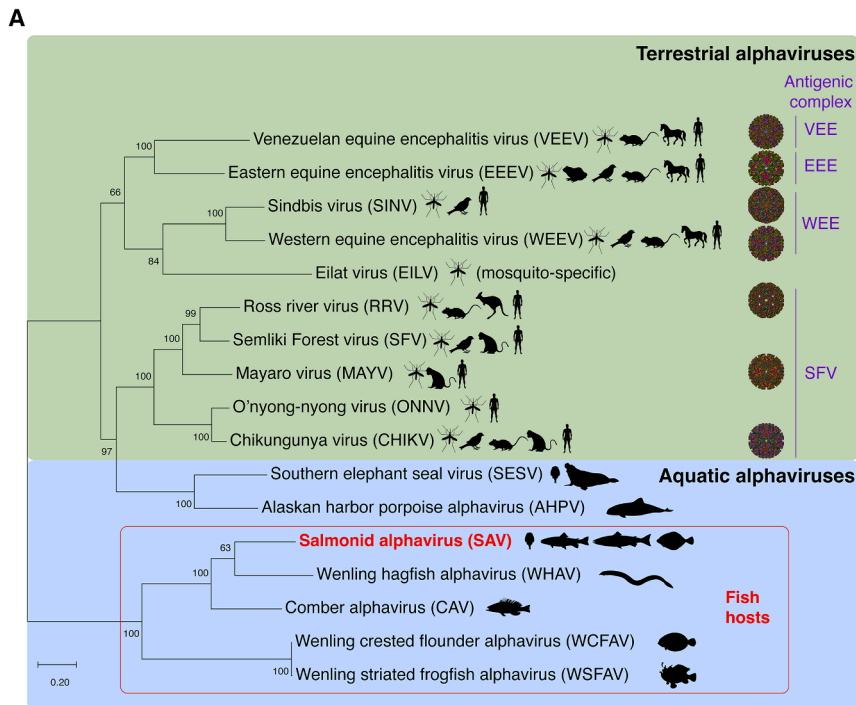


Figure 1. Aquatic and terrestrial alphavirus phylogeny and E2 envelope protein alignment

(A) Phylogenetic tree of alphavirus structural protein sequences. The amino acid sequence of 17 representative terrestrial and aquatic alphaviruses were aligned and a maximum-likelihood (ML) phylogenetic tree was generated (LG [G + I] substitution model). Numbers at nodes indicate percent bootstrap support from 1,000 replicates. Tree drawn to scale with branch lengths measured in number of substitutions per site (scale bar). For each alphavirus, the known host range and vectors are indicated by animal diagrams. Experimentally determined virion structures are depicted for the following terrestrial alphaviruses: VEEV (PDB: 3J0c); EEEV (PDB: 6odf); SINV (PDB: 6imm); WEEV (PDB: 8dec); RRV (PDB: 6vyv); MAYV (PDB: 7ko8); CHIKV (PDB: 3j2w).

(B) Amino acid sequence alignment of the N-term region of alphavirus E2. Protein sequences of the 17 aquatic and terrestrial alphaviruses in (A) were aligned and the region corresponding to E2 N-term is shown. The alignment reveals an amino acid insert present only in fish alphavirus E2 proteins (Indel, top). Salmonid alphavirus (SAV) is unique in harboring an additional stretch of three residues herein referred to as 7-8-9 triplet based on SAV E2 amino acid numbering.

determined structural data available for alphaviruses infecting fish hosts.

The initial genomic characterization of SAV from rainbow trout (subtype 2), then named SD virus (SDV), revealed that it diverged substantially compared to its terrestrial counterparts, notably by the length of its individual proteins, in particular the E1 and E2 envelope glycoproteins which contain large inserts.⁵ As displayed in the protein alignment focusing on alphavirus E2 N-term presented in Figure 1B and a large indel is present close to the N-term with a clear demarcation between alphaviruses infecting fish from the other members of the genus. The alignment highlights the shared insert that is present in fish-infecting alphaviruses. For SAV, the insert is

composed of nine residues with the sequence ₆PVAVYDTQI₁₄. Among fish alphaviruses, the SAV insert is unique as it contains an additional stretch of three residues ₇VAV₉, designated hereafter as triplet 7-8-9. Remarkably, our group has previously demonstrated that a single amino acid substitution at the N-term of SAV2 E2 could dramatically alter virulence *in vivo* in rainbow trout.¹¹ The published work reported that a single substitution, alanine (A) to valine (V) at position 8 (A8V), was responsible for 90% of the attenuation observed in a recombinant clone previously obtained in the laboratory by reverse genetics. However, upon sequencing verifications performed as part of this current study,

we have uncovered that the substitution was erroneously attributed to position 8 (Figure S1). Retrospective RT-PCR and sequencing analyses performed on RNA extracted from preparations of the virulent S49P viral isolate, recombinant SAV (rSAV2) strains, as well as plasmid constructs used for SAV2 reverse genetics all confirm that the attenuating substitution actually concerns position 9 (i.e., A9V) and not position 8, with A₉ associated with virulence and V₉ associated with attenuation (Figure S1). Of note, the above-mentioned error does not concern the methionine (M) to threonine (T) substitution at position 136 (M136T), which was previously found to have a minor role in modulating virulence.¹¹ Since position 9 is part of the triplet 7-8-9 triplet within the SAV E2 N-term insert, we decided

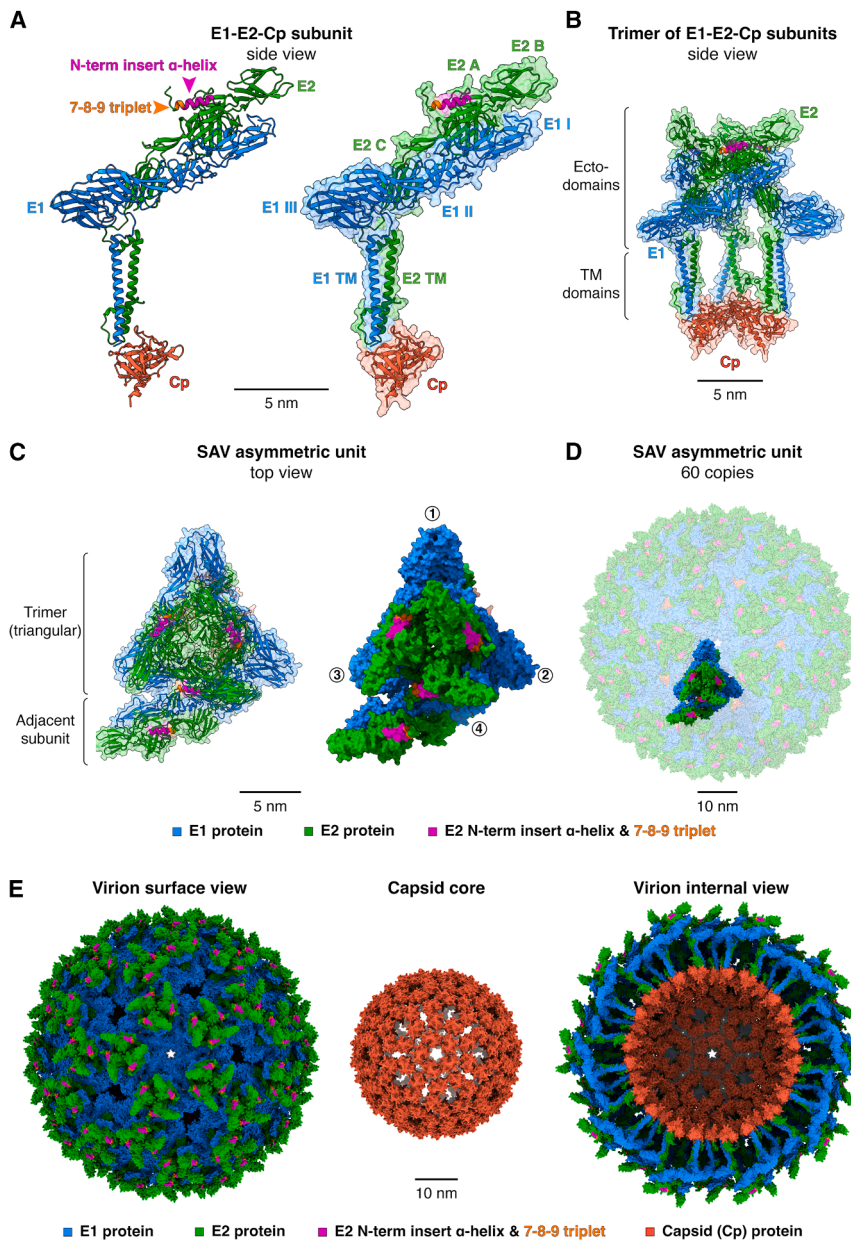


Figure 2. 3D reconstruction and architecture of a SAV virion by predictive *in silico* approaches

(A) E2-E1-Cp subunit shown as ribbons and with surface transparency. The structures of E1, E2, and the ordered C-term half of Cp were predicted using AlphaFold 3. The domains of E1 and E2 proteins are indicated along with the location of E2 N-term insert α -helix and the 7-8-9 triplet (arrows). (B) Side-view of a trimer of E2-E1-Cp subunits represented as ribbons with surface in transparency.

(C) Details of an SAV asymmetric unit (top view) shown as ribbon and surface representations. Circled numbers indicate individual E2-E1-Cp subunits forming the asymmetric unit.

(D) 60 copies of SAV asymmetric units forming an icosahedral virion.

(E) Predicted structure of a complete icosahedral SAV virus particle with outer surface, capsid core, and internal views. The SAV virion is composed of 240 E2-E1-Cp subunits and assembled using symmetry operations of the reference VEEV icosahedral structure as template (PDB: 3j0c). For video animations of the virion assembly, see Videos S1, S2, and S3. For panels A to E, structure labeling, representation, and scale bars were visualized using ChimeraX.

the two proteins were modeled together, while Cp was modeled separately. For Cp, only the structurally ordered C-terminal (C-term) domain containing a chymotrypsin-like fold was modeled (aa 122–283), since the N-term tail which binds to genomic RNA is disordered.^{9,20} The predicted models of SAV E1, E2, and Cp proteins were then mapped by structural alignment onto a template E2-E1-Cp subunit of VEEV (PDB 3j0c)⁹, a reference alphavirus with a robust cryo-EM-determined structure.

AlphaFold 3 generated a model that captures well the architecture of alphavirus E2-E1 heterodimer with similar domain organization (Figures 2A and S2). The overall predicted template

to investigate this position along with the 7-8-9 triplet further, as detailed below.

***In silico* structural protein modeling and 3D reconstruction of an SAV virion**

To gain more insights into salient structural features of SAV, we set out to model its structural proteins using AlphaFold 3, one of the leading and most accurate deep learning tools for predicting protein structures and complexes.¹⁸ The polyprotein sequence of the prototype SAV (GenBank: AJ316246.1) was chosen to predict the structures of E1 and E2 envelope glycoproteins and the Cp protein. Since alphavirus E1 and E2 form tight heterodimers and because AlphaFold 3 is able to predict protein complexes,

modeling (pTM) score, a measure of the accuracy of the entire structure, was high at 0.76 (0–1 range; the higher the better) and the interface pTM (ipTM) score, which measures the accuracy of the relative position of proteins within a complex, was also high at 0.76 (Table 1). The predicted aligned error (PAE) is an estimate of the error in the pairwise relative position and orientation of protein residues. PAE gives an indication of how accurately AlphaFold positions a given protein subdomain relative to others (Figure S2). For E2-E1 heterodimer, the PAE estimates were generally low except for the TM domains of the proteins. Further, the E2-E1 model was associated with high to very high predicted local distance difference test scores (pLDDT; 0–100 range, the higher the

Table 1. AlphaFold 3 confidence scorings for SAV E1, E2, and Cp

Protein/complex	Region (aa)	pTM	ipTM	$C_{\alpha}pLDDT$
E2-E1 heterodimer	E1 1–461 & E2 1–438	0.76	0.76	–
E1 _{ecto}	E1 1–397	–	–	91
E2 _{ecto}	E2 1–352	–	–	88
E1 _{TM}	E1 398–461	–	–	66
E1 _{TM} internal region (GenBank: AJ316246.1)	E1 429–441	–	–	51
E1 _{TM} internal region (GenBank: NC_003930.1)	E1 429–440	–	–	72
E2 _{TM}	E2 353–438	–	–	63
E2 _{N-term} α -helix (without E3)	E2 7–20	–	–	63
Cp	Cp 122–283	0.87	–	–
Cp	Cp 122–283	–	–	90
E3 in uncleaved p62	E3 1–71	–	–	75
E3	E3 1–71	–	–	80
E2 _{N-term} α -helix without E3	E2 7–20	–	–	63
E2 _{N-term} α -helix with E3 and cleaved furin loop	E2 7–20	–	–	85
E2 _{N-term} α -helix with E3 and uncleaved furin loop	E2 7–20	–	–	86

The table summarizes the different scores generated by AlphaFold 3 for the different regions of the structural proteins of SAV. Scores shown: predicted template modeling, pTM; interface predicted template modeling, ipTM; and the average per-residue carbon α (C_{α}) predicted local distance difference test, $C_{\alpha}pLDDT$. pTM and ipTM are within the 0 to 1 range (the closer to 1 the better the prediction). $C_{\alpha}pLDDT$ was calculated by averaging the pLDDT scores of the C_{α} of each residue within a given protein region (score range 0–100, the closer to 100 the better).

better) that provide a residue-level confidence metric (Figure S2). The confidence scoring was particularly robust for residues of the beta sheet-rich ectodomains of E1 (aa 1–397) and E2 (aa 1–352) with associated average per-residue Carbon α (C_{α}) pLDDT ($C_{\alpha}pLDDT$) scores of 91 and 88, respectively (Table 1). For both E1 and E2, the transmembrane (TM) domains are predicted to be structured as long α -helices with relatively lower average $C_{\alpha}pLDDT$ scores of 66 to 63, respectively (Table 1).

The ordered region of Cp (aa 122–283) was modeled by AlphaFold with high accuracy. It scored a pTM of 0.87, generally low PAE error values, and with a very high average $C_{\alpha}pLDDT$ score of 90 (Figures 2A, S2, and Table 1).

Intriguingly, the E2 N-term insert and 7-8-9 triplet identified in the protein alignment (Figure 1B) are located in proximity to E2 domain A in a linker region known as the β -ribbon.⁸ Part of the SAV E2 insert and the triplet 7-8-9 are predicted to form an extended α -helix located near the apex of the E2-E1 heterodimer (Figure 2A), with an average $C_{\alpha}pLDDT$ score of 63 (Table 1). The presence of this α -helix is unexpected as alphavirus E2 ectodomains are considered to be all- β proteins belonging to the immunoglobulin superfamily, as typified by the CHIKV E2 crystal structure.⁸ To obtain further validation of this atypical α -helical secondary structure, we used RoseTTAFold, another highly accurate protein structure prediction program based on a distinct three-track neural network,²¹ to predict the structure of SAV E2 (Figure S3A). Overall, RoseTTAFold generated a structure of E2 that is highly similar to the one predicted by AlphaFold 3 (Figure S3A). RoseTTAFold also predicts that part of the residues of the N-term insert form an α -helix (E2 aa 9–19). These predictions are further confirmed by PsiPred, a sequence-based secondary structure prediction program which also predicts the

presence of an α -helix at E2 N-term (E2 aa 9–20)^{22,23} (Figure S3B).

Since Alphavirus virions assemble into particles with a regularly organized architecture adhering to a $T = 4$ icosahedral symmetry,²⁰ we reasoned that it could be feasible to reconstruct *in silico* a complete SAV virion from the modeled SAV E1, E2, and Cp structural proteins. To build the SAV virion, the E1, E2, and Cp protein models were first mapped onto a reference alphavirus asymmetric unit (VEEV, PDB: 3j0c) which is composed of four E2-E1-Cp subunits arranged as a triangular trimer (Figure 2B) with an adjacent subunit at its base (Figure 2C). The symmetry operations of the reference VEEV structure allowing to generate a complete icosahedral virion from 60 copies of the asymmetric unit were then applied to the SAV asymmetric unit assembly (Figure 2D). This approach successfully reconstructed an entire SAV virion based on the modeled E1, E2, and Cp proteins. Each modeled proteins fit well within the $T = 4$ reference icosahedral lattice, revealing a typical alphavirus structural organization made of two concentric icosahedral shells with the outer shell composed of the envelope glycoproteins E1 and E2 and an enclosed shell made of Cp proteins (Figures 2E, Videos S1, S2, and S3). Importantly, the reconstructed SAV virion is composed of 240 copies of each of the structural modeled proteins with the E2 N-term α -helix and 7-8-9 triplet prominently exposed at the surface. Such exposure and repetition suggest that even subtle variations in E2 N-term could potentially lead to measurable phenotypic effects, such as for virulence and neutralizing antibody binding as shown for SAV domain B.¹⁴ The SAV virion model constitutes a powerful tool to generate testable hypotheses for structure-function studies and for assessing the phenotypic impacts of variations.

Investigating SAV p62 (E3E2) furin-mediated maturation process

The alphavirus E2 envelope glycoprotein is first synthesized as a precursor called p62 (pE2) which on a sequence level is composed of E2 with a small N-term extension called E3 (71 aa for SAV). E3 plays a pivotal role for p62-E1 heterodimer formation and stabilization in the ER, as its presence avoids premature release of the E1 fusion loop upon exposure to low pH in the compartments of the secretory pathway.^{20,24,25} E3 possesses at its C-term extremity a conserved furin cleavage site (Figure 3A). This basic stretch of residues (consensus sequence R-X-R/K-R) is conserved for all alphaviruses underscoring its functional importance. The maturation of p62 into E2 takes place in the *trans*-golgi network (TGN) and involves the cellular protease furin which cleaves off E3. In some alphaviruses E3 can remain non-covalently associated with E2, such as for SFV,²⁶ VEEV,⁹ and CHIKV.⁸ In the case of SAV, because the N-term E2 insert α -helix predicted by AlphaFold 3, RoseTTAFold, and PsiPred is located proximal to the C-term of E3 in the protein sequence of p62, we set out to investigate its maturation process and potential consequences for E2 (Figure 3A).

We first examined biochemically the cleavage status of p62 by performing western blot analysis of SAV-infected cell lysates and sucrose cushion-purified SAV virions using specific antibodies directed against E2 (mAb 17H23) and E1 (78k5) proteins (Figure 3B).^{14,27} Both proteins were readily detected in cell lysates and purified virions (Figure 3B, lanes 2–4). The analysis further confirms that in SAV-infected cells where p62 is synthesized and protein maturation occurs, E2 is detected as two distinct bands with one band corresponding to the furin-processed p62 cleaved into E2 and migrating slightly above the 50 kDa marker, and the uncleaved p62 precursor which migrates at a higher molecular weight because of the E3 N-terminal extension (Figure 3B, lanes 2 and 3). In contrast, in purified virions only the ~50 kDa E2 band is present indicating that only furin-processed E2 and not p62 assembles and buds into newly formed SAV particles (Figure 3B, lane 4).

Next, we computationally simulated three distinct stages of SAV p62 maturation by modeling with AlphaFold 3 the ectodomains of E1 and E2 in presence or absence of E3 (Figure 3C). In a first stage, E3 is modeled as covalently linked to E2 corresponding to the p62 precursor prior to cleavage by furin at the ₆₈R-K-K-R₇₁↓₁A-V-S-T₄ cleavage site (Figure 3C, top). In a second stage, E3 is present but non-covalently-linked to E2 corresponding to E3 in complex with E2-E1 ectodomains post furin cleavage (Figure 3C, middle). In a final third stage, E3 is absent and this corresponds to the post-furin cleavage stage and release of E3 from the E2-E1 heterodimer. Remarkably, AlphaFold 3 was able to accurately model E3 with average C_{α} PLDDT scores of 75 and 80 for the uncleaved (p62) and cleaved form, respectively, with lower scores for the N- and C-termini (Figure 3C, Table 1). SAV E3 is predicted to be structured very similarly to VEEV and CHIKV E3, notably with the presence of two parallel short rod-like α -helices.^{8,9} The predicted models generated by AlphaFold 3 also predict that E3 is positioned near the apex of the E2-E1 heterodimer, on top of the E2 β -ribbon connector where the E2 N-term insert α -helix is located (arrow) and in proximity to E2 domain A, irrespective of whether

p62 is cleaved or not (Figure 3C, compare top and middle). Structural alignment of uncleaved SAV p62 with the crystal structure of uncleaved CHIKV p62 reveals that the E3 proteins of both viruses superimpose very well (RMSD between 136 pruned atom pairs, 1.34 Å) with a close fit of the two defining α -helices and β -strand secondary structures (Figure 3D). The predicted positioning of SAV E3 relative to E2 in the AlphaFold model matches very closely with the one observed for CHIKV p62.⁸ Thus, the positioning of E3 relative to E2 appears to be an evolutionarily conserved feature shared between terrestrial and aquatic alphaviruses. Previous structural work demonstrated that E3 and E2 residues responsible for the interaction at the E3-E2 interface were conserved between SINV (E3 Y26, L34, and Y46; E2 K219) and CHIKV (E3 Y27, L35, and Y47; E2 K255).²⁸ Remarkably, SAV E3 Y54 (C_{α} PLDDT score 94) aligns structurally with CHIKV E3 Y47 with side chains aligning almost perfectly (Figure 3D). For E2, we found that position SAV E2 R268 (C_{α} PLDDT score 97) aligns structurally with CHIKV E2 K255. SAV E2 R268 and CHIKV E2 K255 are both found near the middle of an E2 β -strand of the β -ribbon connector region with very close alignment of their side chains. R and K residues are both basic amino acids with positively charged side chains and similar biochemical properties. Overall, of the four conserved E3 and E2 residues identified as important for mediating the E3-E2 interaction for CHIKV and SINV, we show here that SAV has an identical residue in E3 (Y54) and an E2 residue that is highly similar (R268), both with a high degree of prediction confidence scores (C_{α} PLDDT above 95) and remarkable alignment of their side chains.

The *in silico* predictive models suggest that for SAV p62, the furin cleavage site forms an exposed, unstructured loop associated with an average C_{α} PLDDT score below 50, indicative of a disordered and likely flexible region (Figure 3C, top and 3D). The furin loop directly connects with SAV E2 N-term and the first residues of the insert α -helix and the 7–8–9 triplet (Figure 3C, top). Remarkably, in all three p62 maturation stages the secondary structure of E2 insert α -helix is predicted to be maintained (Figure 3C). While E3 partially covers the E2 insert α -helix, the triplet residues are exposed during all stages of the maturation process of p62 (Figure 3C top, middle, bottom). Of note, the E2 N-term insert α -helix is predicted with a much higher average C_{α} PLDDT score of 85 when E3 is modeled together with the E2-E1 heterodimer ectodomains compared to when E3 is absent, average C_{α} PLDDT score of 63 (Figure 3C middle and bottom; Table 1). This is also the case when the furin loop remains uncleaved (Figure 3 middle with top; Table 1), average C_{α} PLDDT score of 86. These analyses suggest that the N-term insert α -helix may play an important role in making the furin loop accessible to proteolytic processing and suggests that E3 could also potentially play a stabilizing role locally at the N-term of E2 and the insert α -helix. For SAV, the extended α -helix makes the flexible furin loop protrude more outwardly from p62 compared to the CHIKV counterpart, perhaps making this site more accessible to furin proteolytic processing (Figure 3D).

Structural and evolutionary relationships among alphaviruses

To obtain a more comprehensive picture of the structural-evolutionary relationships between aquatic and terrestrial alphaviruses,

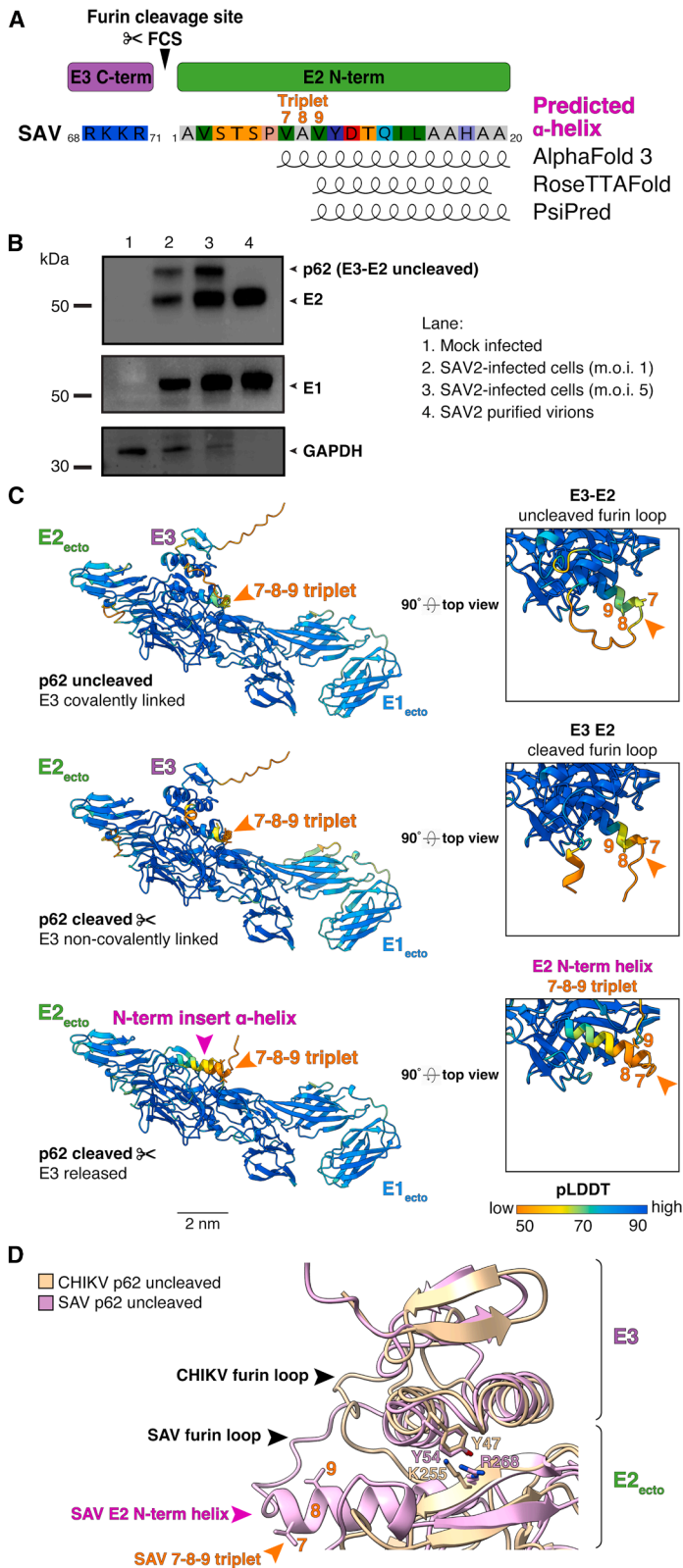


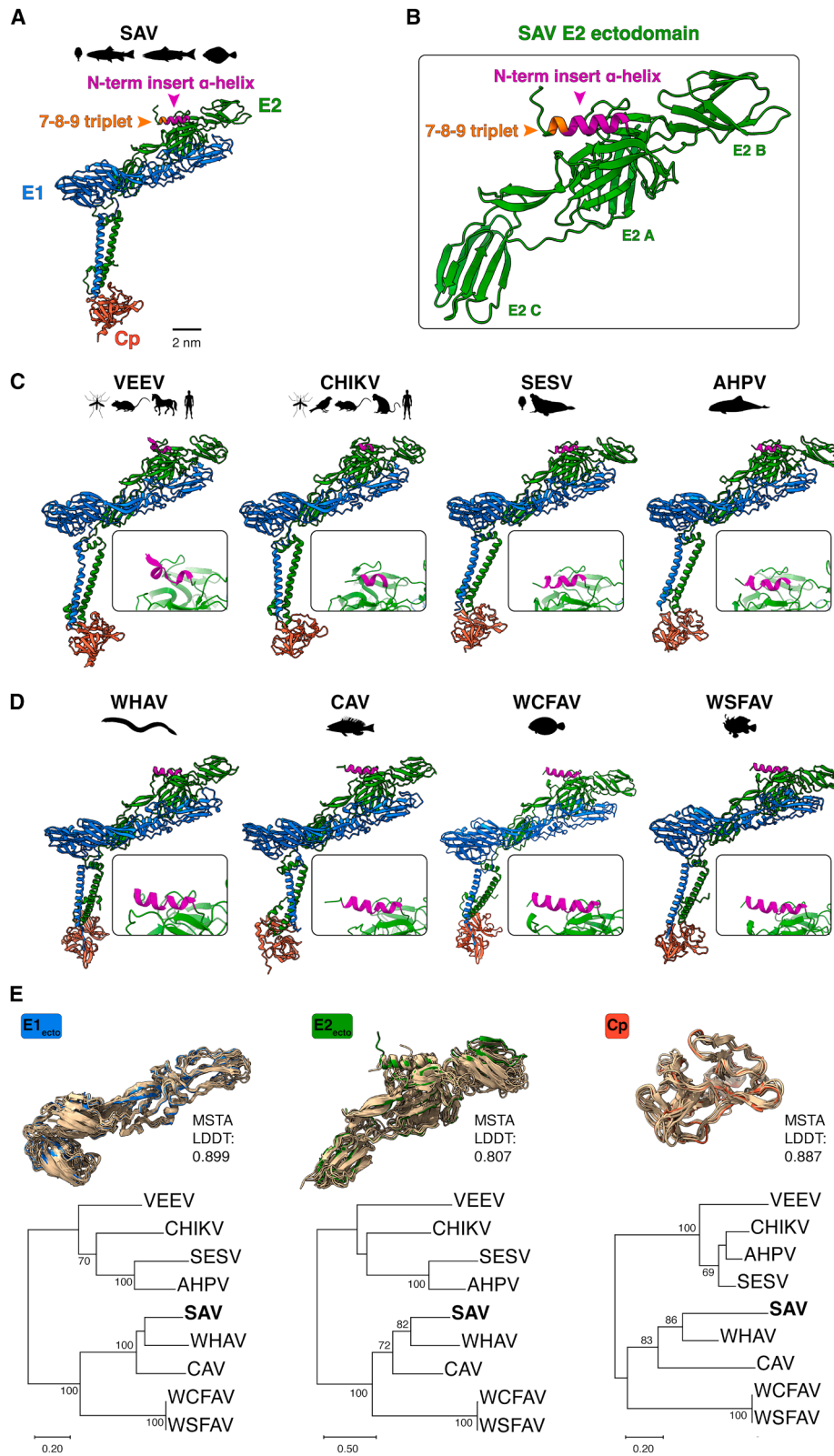
Figure 3. *In vitro* biochemical analysis and *in silico* modeling of SAV E3E2 furin cleavage and comparison with CHIKV

(A) SAV E3 C-term furin cleavage site (FCS) and E2 N-term. Position of the FCS, amino acid composition, triplet 7-8-9 and E2 N-term insert α -helix are shown. The residues involved in the E2 N-term α -helix predicted by AlphaFold3, RoseTTAFold, and PsiPred are depicted as helical cartoons.

(B) Biochemical analysis of cleavage status of p62 in infected cells and purified SAV virions. Western blot analysis was performed on lysates of mock-infected (lane 1) or SAV2-infected BF-2 cells at an m.o.i. of 1 (lane 2) or 5 (lane 3) or on sucrose cushion-purified SAV2 virions (lane 4). Western blot analyses performed using anti-E2 (17H23) or anti-E1 (78K5) mAbs. GAPDH detection was used as loading control for cell lysates.

(C) Structural modeling of SAV p62 (E3E2) cleavage by furin. SAV E1_{ecto}, E2_{ecto}, and E3 were modeled as complexes using AlphaFold 3. Top corresponds to uncleaved p62 (E3E2), middle corresponds to E3 cleaved from E2 with E3 remaining non-covalently linked, and bottom corresponds to E1 and E2 ectodomains after E3 release. Arrows indicate location of E2 N-term α -helix and 7-8-9 triplet. The models are colored using AlphaFold pLDDT confidence scoring.

(D) Structural alignment of CHIKV and SAV uncleaved p62. The uncleaved model of SAV p62 (pink) was structurally aligned with the crystal structure of CHIKV p62 (gold, PDB: 3n40) using MatchMaker in ChimeraX. For panels C and D, structure labeling, representation, and scale bars were visualized using ChimeraX.



(legend on next page)

we examined the E2-E1-Cp subunit of SAV and compared it with that of terrestrial alphaviruses (VEEV and CHIKV) for which experimentally determined structures are available (PDB: 3j0c and 8fcg, respectively), and AlphaFold 3-predicted models of mammalian aquatic alphaviruses SESV and AHPV, as well as fish-infecting alphaviruses WHAV, CAV, WCFAV, and WSFAV (Figures 4A–4D). This comparative analysis reveals that the overall organization of the E2-E1-Cp subunit for all alphaviruses appears to be well conserved. However, a closer inspection of the N-term of the different alphavirus E2 proteins reveals important variations in the N-term insert α -helix that was identified for SAV. Indeed, for the distantly related terrestrial VEEV and CHIKV, for which cryo-EM structures are displayed, only very short turns are present (Figure 4C). For CHIKV, this short turn was also observed in the X-ray crystal structure of its furin cleaved p62 protein (PDB: 3n42).⁸ For the E2 of aquatic alphaviruses infecting mammals, SESV and AHPV, a short α -helix is predicted to be present (Figure 4C). Similarly to SAV, the closely related fish-infecting alphaviruses, WHAV, CAV, WCFAV, and WSFAV, are all predicted to share a longer α -helix at their N-term, similar to the one predicted for SAV (Figure 4D). These results suggest that the N-term insert and its predicted extended α -helix is a conserved feature of aquatic alphaviruses which may indicate an important functional and adaptive role for this structural feature for alphaviruses in aquatic environments.

Next, we assessed the structural relatedness of the ectodomains of E1 and E2, and Cp (structured C-term) of terrestrial and aquatic alphaviruses using FoldMason, a powerful multiple structural alignment algorithm (MSTA) that leverages the structural alphabet developed in FoldSeek.^{29,30} Importantly, FoldMason can handle both experimentally and predicted protein structures. Thus, the E1, E2, and Cp protein domains of terrestrial and aquatic alphaviruses were structurally aligned using FoldMason (Figure 4E top). The structural alignments confirm that protein domains (E1_{ecto}, E2_{ecto}, and Cp) of terrestrial and aquatic alphaviruses all fold very similarly, with few deviations in structural features. For each alignment FoldMason generates an MSTA LDDT score, a measure of similarity of protein structures based on their local structural neighborhoods on a scale ranging from 0 to 1. FoldMason calculated MSTA LDDT scores of 0.899 for E1_{ecto}, 0.807 for E2_{ecto}, and 0.887 for Cp, indicating that members of the *Alphavirus* genus from aquatic to terrestrial environments share structural similarities for the three analyzed proteins. From the MSTA FoldMason generated a set of three

structure-informed amino acid alignments. To examine whether these structure-based alignments could recapitulate phylogenetic relationships, each FoldMason structure-based alignment was used to infer an ML phylogenetic tree (Figure 4E, bottom). Overall, the structure-based trees revealed similar branch topologies as sequence alignment-based trees (Figure S4) and recapitulates the main alphavirus groupings obtained in the sequence-based alignment (Figure 1A) with generally high bootstrap support for the major branches (Figure 4E). Of note, for both structure- and sequence-based phylogenetic analyses, the overall tree topologies were similar, but the branches for CHIKV, AHPV, and SESV differed slightly in Cp trees compared to E1_{ecto} or E2_{ecto} trees.

Phylogenetic & structural analysis of SAV E2 and triplet 7-8-9 variations

We next concentrated on naturally occurring variations and the sequence diversity found at the N-term insert α -helix and triplet 7-8-9 motif. To obtain a sufficiently wide array of SAV isolates from various host species, SAV E2 sequences were retrieved by several methods including performing searches in NCBI GenBank and by conducting Blast analysis using SAV2 E2 reference sequence as input (GenBank: AJ316246.1). Also included were sequences from SAV field strains made available in published work.^{31–33} This allowed to conduct an extensive phylogenetic analysis of E2 sequences on a diverse set of 90 isolates with representatives from the main subtypes (SAV1–6) from a variety of hosts including rainbow trout, Atlantic salmon, common dab and European plaice (Table S2). As subtype information of some sequences were not provided in the sequence metadata, a nucleotide sequence alignment based on an internal 357-nucleotide E2 fragment used for subtype demarcation was generated.³⁴ This allowed to construct an ML phylogenetic tree of 90 SAV isolates (Figure 5A). When available, the host species from which the isolate originated is shown. The E2-based tree allowed to assign subtypes for all 90 isolates. SAV2 and SAV3 isolates, the most represented in this tree, formed a strongly supported clade (bootstrap support of 99). Likewise, representative isolates from SAV1, SAV4, and SAV5 subtypes clustered into a subclade which was also well supported (bootstrap support of 97) while the SAV6 subtype representative formed a distinct outlier branch from all other subtypes.

For the triplet 7-8-9 site, the most frequently observed motif for all six subtypes analyzed was γ AAV₉ (A, alanine; V, valine; 77/90),

Figure 4. Structural and evolutionary analyses of aquatic and terrestrial alphavirus proteins

(A) E2-E1-Cp subunit of SAV predicted by AlphaFold 3.

(B) Detailed view of E2 ectodomain with N-term insert helix and 7-8-9 triplet highlighted (arrows).

(C) E2-E1-Cp subunit of VEEV (PDB: 3j0c) and CHIKV (PDB: 8fcg) terrestrial alphaviruses and E2-E1-Cp subunit of mammalian aquatic alphaviruses SESV and AHPV predicted by AlphaFold 3.

(D) E2-E1-Cp subunit of WHAV, CAV, WCFAV, and WSFAV fish alphaviruses predicted by AlphaFold 3. For Cp, only the ordered C-term region of the respective proteins was modeled. The known host range and vectors are indicated by animal diagrams. For (C and D), a focused view of the E2 N-term of each model is shown in each box. Structure labeling, representation, and scale bars were visualized using ChimeraX.

(E) Comparative analyses between alphavirus proteins using structure-based protein alignment and phylogenetic analysis. FoldMason was used to perform multiple protein structure alignments (MSTA) of alphavirus E1 and E2 ectodomains and Cp protein (C-term domain). In each structural alignment, the blue (E1), green (E2), and red (Cp) protein corresponds to SAV while the superimposed proteins in gold are from the other species analyzed. The structure-based protein sequence alignments were used to build Maximum-Likelihood phylogenetic trees using WAG+G+I substitution model for E1 and WAG+G for E2 and Cp. Numbers at nodes indicate percent bootstrap support from 1,000 replicates. Trees drawn to scale with branch lengths measured in number of substitutions per site (scale bars).

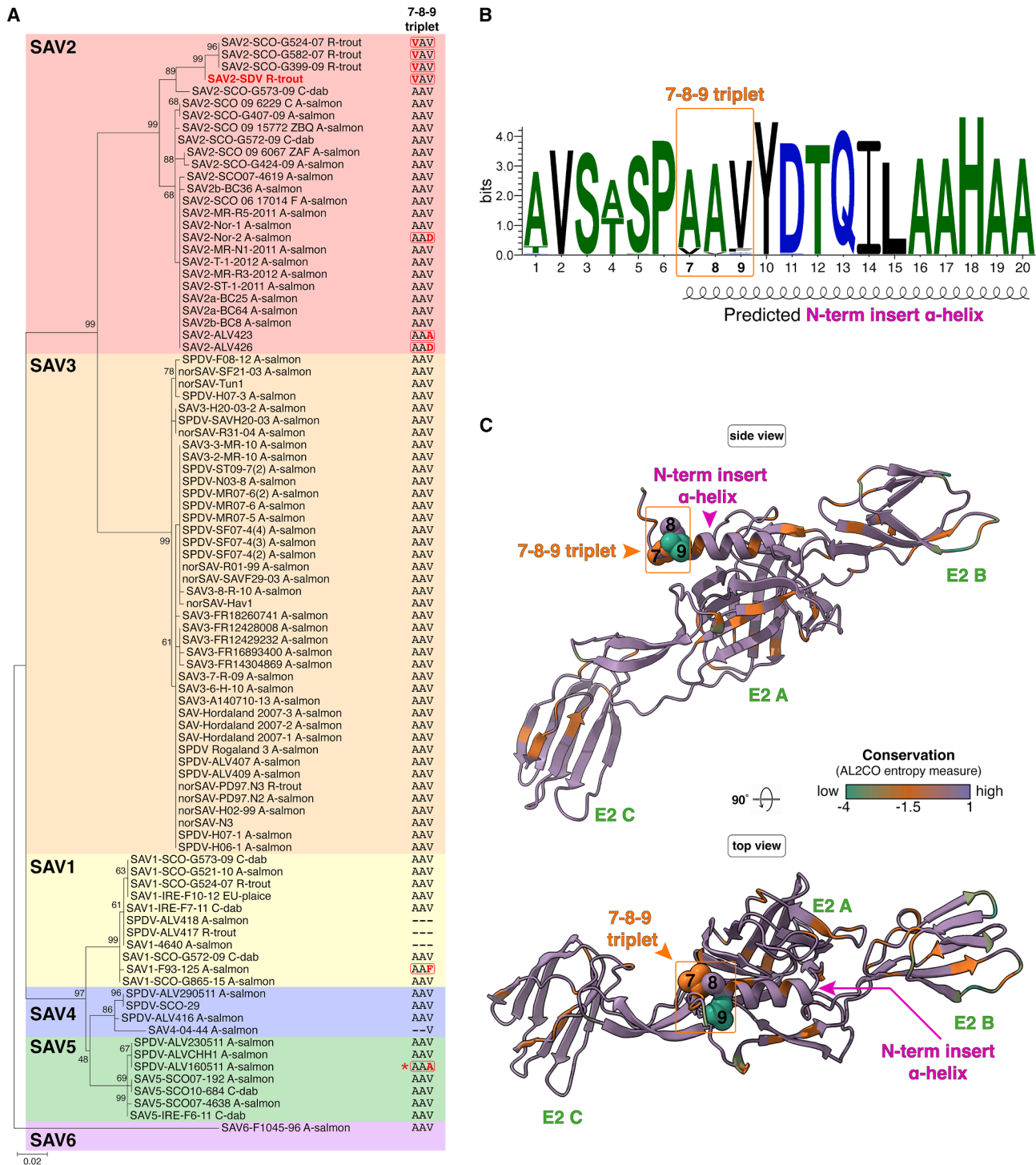


Figure 5. Phylogenetic structural analysis of SAV E2 and triplet 7-8-9 variations

(A) Phylogenetic tree based on partial E2 nucleotide sequences of 90 SAV strains. SAV E2 sequences (357-nucleotide region used for subtype demarcation) from 90 SAV strains belonging to subtypes 1–6 were retrieved from various sources (Table S2). The sequences were aligned and a phylogenetic tree was generated (GTR [G + I] substitution model). Numbers at nodes indicate percent bootstrap support from 1,000 replicates. Tree drawn to scale with branch lengths measured in number of substitutions per site (scale bar). When known, the host from which the viral sequence originated is indicated. A-salmon, Atlantic salmon (*Salmo salar*); C-dab, common dab (*Limanda limanda*); EU-plaice, European plaice (*Pleuronectes platessa*); R-trout, rainbow trout (*Oncorhynchus mykiss*). For each sequence, the amino acid composition for the 7–8–9 triplet is indicated with variations from the AAV consensus motif highlighted in red/bold. The strain shown in

(legend continued on next page)

hereafter designated as the consensus motif. The AAV consensus clearly stands out in the Weblogo made by aligning the 90 sequences analyzed in the phylogenetic tree (Figure 5B). Interestingly, the AAV triplet 7-8-9 motif was invariant for all SAV3 isolates. The divergent SAV6 subtype sequence also harbors this conserved sequence. Variations in residue composition were identified for several isolates, notably in seven isolates of the SAV2 subtype. All variations affected the residues at position 7 or 9, while residue position 8 remained invariant for all isolates analyzed (A₈). Interestingly, the γ VAV₉ motif, containing the alanine (A) to valine (V) substitution at position 7 compared to the consensus, was found for SAV2 reference sequence from rainbow trout (GenBank: [AJ316246.1](#)) along with three more recently characterized field isolates from Scottish rainbow trout (SCO-G524-07, SCO-582-07, and SCO-G399-09). Importantly, this specific variation is only found in SAV2 strains infecting rainbow trout, suggesting it may be a host species adaptation within a defined geographic range (Scotland/continental Europe). Interestingly, other variations found for SAV2 strains affect position 9: γ AAD₉ (D, aspartic acid) Nor-2 from Atlantic salmon and ALV426 and γ AAA₉ for ALV423 strain. Of note, three isolates from subtype 1 SAV, ALV-418 from Atlantic salmon and ALV-417 from rainbow trout as well as strain 4640 from Atlantic salmon presented a six-residue deletion encompassing the triplet 7-8-9 motif (E2 residues 4 to 9, Δ_{4-9}). In addition, the reference SAV1 isolate (F93-125) from Atlantic salmon harbored a variant motif γ AAF₉ (F, phenylalanine). All SAV4 strains analyzed harbored the conserved γ AAV₉ sequence with the exception of strain 04-44 from Atlantic salmon which had a large 13 amino acid deletion encompassing the E3 furin cleavage site and the E2 N-term up to the first two residues of the 7-8-9 triplet motif. For SAV5 subtype strains only ALV160511 harbored a variant motif γ AAA₉. For this strain the motif was associated with another variation at position 1 with an alanine (A) to threonine (T) substitution shared by all other members of subtype SAV5 (7 representatives). This N-term proximal variation is reflected in the Weblogo which reveals that position 4 also displays similar variation (A/T).

Next, sequence conservation was mapped onto the E2 ectodomain structure from the reference sequence (GenBank: [AJ316246.1](#); γ VAV₉ motif) (Figure 5C). The protein sequence alignment of the 90 isolates was used as input to obtain per-residue conservation scores that are mapped onto the SAV E2 ectodomain structure (AL2CO algorithm), an approach that allows to visualize spatial conservation features within a given protein structure (Figure 5C).³⁵ Overall, most of E2 ectodomain amino acids were found to be relatively well conserved (Figure 5C, purple). However, position 9 within the triplet 7-8-9 was associated

with a score among the lowest for conservation (green), with position 7 also associated with a low score (orange). This analysis suggests that variations of residue composition within these exposed sites could be of functional and/or structural significance. It is noteworthy to point out that other sites with low conservation are found in unstructured loops of E2 domain B, which is one of the main target sites for neutralizing antibodies (Figure 5C).¹⁴

In vitro investigation of the biological effects of variations in SAV E2 N-term

To assess the biological impact of variations in the structural features identified at the SAV E2 N-term, we used SAV reverse genetics developed in our laboratory to generate recombinant SAV2 (rSAV2) harboring selected mutations identified in the phylogenetic analyses.³⁶ In addition to previously obtained triplet 7-8-9 rSAV2 variants which harbor E2 VAV and VAA 7-8-9 triplet motifs (Figure S1),¹¹ other mutations were introduced by site-directed mutagenesis enabling to obtain the following rSAV2 E2 triplet 7-8-9 motif variants: AAV corresponding to the consensus motif, AAA, AAF, and AAD. We also generated the T1-AAA variant which corresponds to the AAA E2 triplet 7-8-9 with an A1T substitution at its N-term, to match the N-term sequence of the ALV16051 subtype 5 strain identified in the phylogenetic analysis (Figure 5A). Of note, despite several attempts, a rSAV2 with a six-residue deletion encompassing the triplet 7-8-9 motif (rSAV2 E2 Δ_{4-9}) was not recoverable (Table S4). rSAV2 harboring the substitutions were recovered and used to infect susceptible BF-2 cells (Table S4). Infections by these rSAV2 were analyzed by immunofluorescence assay at 7- and 10-day post-infection (d.p.i.) to detect SAV-infected cells (Figure 6A).

The rSAV2 with the consensus 7-8-9 triplet motif AAV displayed robust infection at 7 and 10 d.p.i. with most cells being positive for infection, a result comparable to infection with the rSAV2 with VAV and VAA triplet motifs (Figure 6A, top). In contrast, infection by rSAV2 with the AAA triplet motif resulted in very limited infection at both time points (Figure 6A, bottom). Strikingly, the introduction of A1T substitution in the triplet AAA background (rSAV2 E2 T1-AAA) restored infectivity to levels similar to those of the consensus sequence (rSAV2 E2 AAV), suggesting a possible compensatory effect. Infection by rSAV2 E2 AAD was associated with poor *in vitro* replication fitness with only patches of infected cells observed at 7 and 10 d.p.i. The rSAV2 E2 AAF variant appeared to be defective and unable to infect cells, with no positively infected cells observed at both time points (Figure 6, bottom). These immunofluorescence

red font corresponds to the SAV reference (GenBank: [AJ316246.1](#)). * denotes the presence of a threonine (T) residue at E2 position 1 for the ALV160511 strain instead of the consensus alanine (A).

(B) Sequence logo of the first 20 amino acids of SAV E2. An alignment of the first 20 amino acids of SAV E2 sequences from the 90 strains shown in (A) was used as input to generate the sequence logo using Weblogo 3.7.12. The total height of a stack at each position indicates sequence conservation, while the height of a symbol indicates the relative frequency of each amino acid. The position of the 7-8-9 triplet is indicated (top).

(C) Mapping of residue conservation within SAV E2 ectodomain. Visualization of residue conservation was performed using the SAV E2 ectodomain (aa 1–352) predicted by AlphaFold 3 and AL2CO entropy measure analysis which is based on a protein alignment of SAV E2 sequences from the 90 strains shown in (A). Residues within the E2 ectodomain were color-coded according to their conservation with purple indicating a conserved site, orange and green indicating sites associated with increasing variability. The 7-8-9 triplet residues are depicted as spheres. Structure representation and AL2CO analysis were performed using ChimeraX.

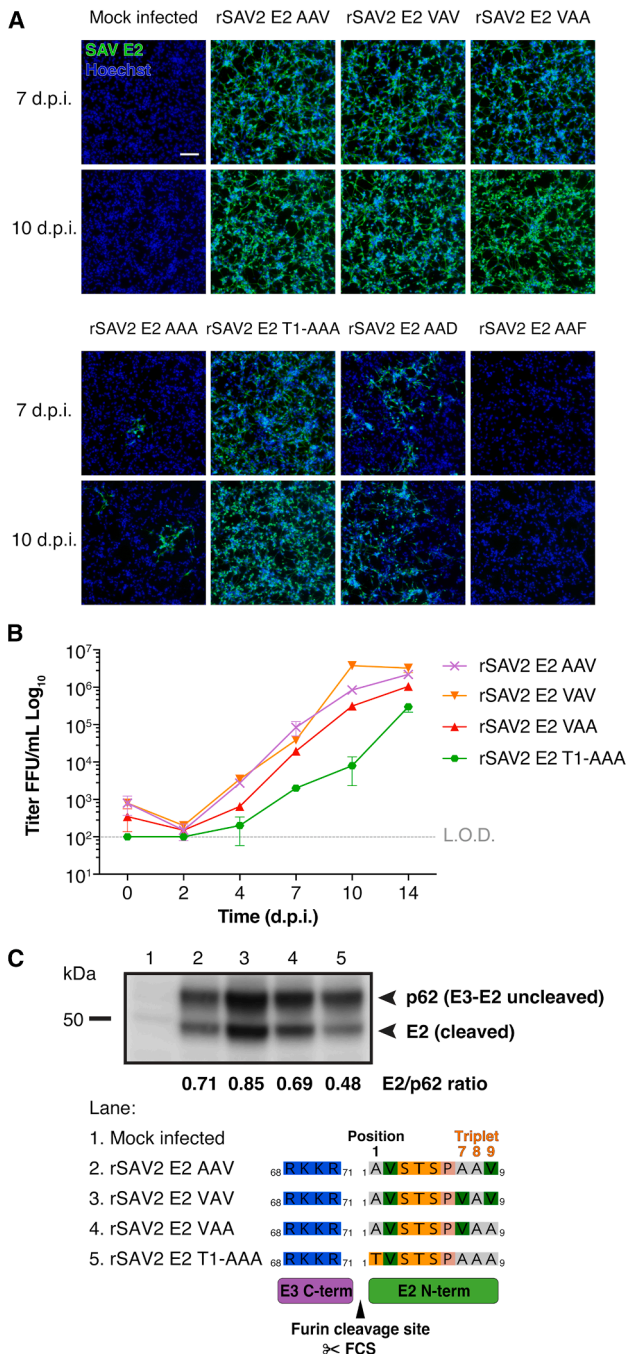


Figure 6. Impact of SAV E2 N-term substitutions on rSAV2 infection in cell culture and on p62 (E3E2) furin cleavage

(A) Immunofluorescence assay of BF-2 cells infected with rSAV2 harboring E2 envelope proteins with distinct 7-8-9 triplet motifs: AAV consensus, VAV, VAA, AAF, AAD, AAA, and T1-AAA. E2 N-term variations were introduced using site-directed mutagenesis starting from the VAA variant and recombinant viruses with variant 7-8-9 triplet motif were recovered and used to infect BF-2 cells. Immunofluorescence assays were performed at 7 and 10 d.p.i. Cells were fixed and immunolabeled using anti-SAV E2 mAb (clone 17H23) to visualize infected cells. Scale bars represent 100 μ m.

(B) Viral growth kinetics of rSAV2 harboring E2 proteins with variant 7-8-9 triplet motifs. The rSAV2 variants with AAF, AAD, and AAA E2 triplet 7-8-9

analysis results are confirmed by the titers obtained for each variant (Table S4) and indicate that the triplet 7-8-9 motif is highly sensitive to certain substitutions particularly at position 9 which can have a drastic impact on viral infectivity and replication fitness *in vitro*.

Of the above recovered variant rSAV2, only rSAV2 E2 AAV, VAV, VAA, and T1-AAA yielded sufficient viral titers to conduct viral growth kinetics analyses (Table S4). The replication kinetics measurements were conducted using BF-2 cells infected with the above rSAV2 at a multiplicity of infection (m.o.i.) of 0.1. Supernatants of infected cells were recovered over a period ranging from 0 to 14 d.p.i. and titered by fluorescent focus assay (Figure 6B). This analysis shows that the rSAV2 with E2 AAV consensus triplet motif replicated as efficiently as rSAV2 E2 VAV, reaching similar final titers at 14 d.p.i. of 2×10^6 and 3×10^6 FFU/mL, respectively. The rSAV2 E2 VAA variant followed a similar growth kinetic and reached a slightly lower final titer of 8×10^5 FFU/mL. In contrast, the rSAV2 E2 T1-AAA variant displayed slower replication kinetics, especially at 4, 7, and 10 d.p.i. However, it did reach a final titer of 3×10^5 FFU/mL at 14 d.p.i. which is close to that obtained for rSAV2 VAA.

To investigate whether the slower growth kinetics of the rSAV2 E2 T1-AAA variant could be due to a less efficient proteolytic processing of its p62 protein, we carried out a western blot analysis on lysates from BF-2 cells infected with the 4 recombinant SAVs (rSAV2) with E2 variations at either position 1 or the 7-8-9 triplet: AAV, VAV, VAA, and T1-AAA (Figure 6C). This assay is based on specific detection of SAV E2 allowing to distinguish between the uncleaved p62 (immature) form with furin-cleaved (mature) E2 (Figure 3B). As shown in Figure 6C, infected cell lysates all contained both uncleaved (p62) and cleaved E2 with varying proportions between the two bands for each variant. A densitometry analysis of the ratio between cleaved E2 over uncleaved p62 shows that variants containing an A residue at position 1 of E2 (the closest to the cleavage site and corresponding to P1') had relatively high levels of furin cleavage (ratios range between 0.69 and 0.85, lanes 2–4). However, for the T1-AAA E2 variant (lane 5), the ratio drops to 0.48, with presence of relatively more uncleaved form (p62) than cleaved form (E2) compared

motifs were not included in the analysis due to low virus recovery titers. BF-2 cells were infected at an m.o.i. of 0.1. At 0, 2, 4, 7, 10, and 14 d.p.i. infected cell supernatants were collected, clarified and viral titers were determined by fluorescent focus assay. BF-2 cells were infected with 10-fold serial dilutions of clarified supernatants. After 7 d.p.i. cells were fixed and immunolabeled using anti-SAV E2 mAb (clone 17H23). Fluorescent foci were examined and counted using a fluorescence microscope. Titers are expressed as fluorescent focus units (FFU) per mL. Results are expressed as averages of two replicate experiments ($n = 2$) \pm S.D. L.O.D., limit of detection.

(C) Biochemical analysis of the cleavage status of p62 in BF-2 cells infected with rSAV2 E2 variants. Western blot analysis was performed on lysates of mock-infected (lane 1. Mock infected) or rSAV2-infected BF-2 cells (lanes: 2. rSAV2 E2 AAV; 3. rSAV2 E2 VAV; 4. rSAV2 E2 VAA; 5. rSAV2 E2 T1-AAA). Western blot analyses were performed using anti-E2 (17H23) mAb. Densitometry analyses of E2/p62 ratio were performed using ImageJ. Shown below is a protein alignment of the E3E2 furin cleavage site with the amino acid composition for each rSAV2 E2 variant, which includes E2 position 1 (P1') and the triplet 7-8-9 motif.

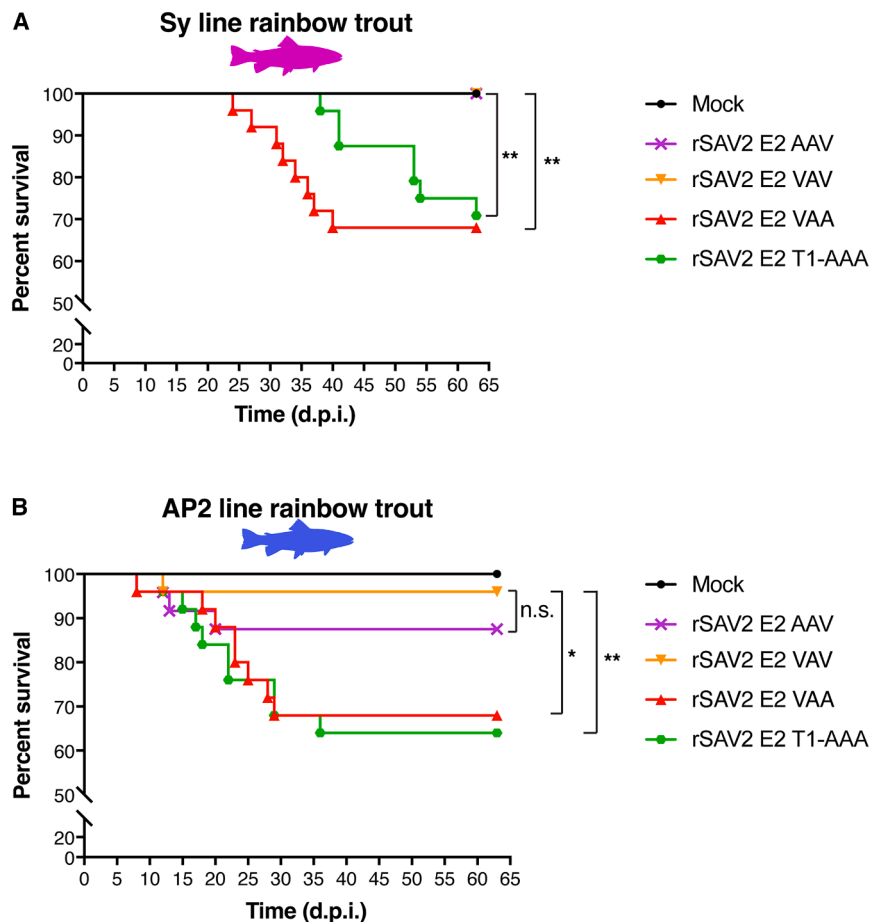


Figure 7. Comparative analysis of the effect of SAV E2 N-term variations in rSAV2 infection in two rainbow trout lines

(A) Cumulative mortalities induced by rSAV2 infection with variant 7-8-9 triplet motif in the rainbow trout Sy line. Twenty-five juvenile trout (average weight of 2.64 g) were infected by intraperitoneal (i.p.) injection with 10^4 FFU per fish of the following recombinant viruses: rSAV2 E2 AAV (consensus); rSAV2 E2 VAV; rSAV2 E2 VAA; rSAV2 E2 T1-AAA or were mock infected. Mortalities were recorded daily after injection over a total period of 63 days.

(B) Cumulative mortalities induced by rSAV2 infection with variant 7-8-9 triplet motif in the rainbow trout AP2 line. Twenty-five juvenile trout (average weight of 2.95 g) were infected by i.p. injection with 10^4 FFU per fish of the same rSAV2 variants as above or were mock infected. Mortalities were recorded daily after injection over a total period of 63 days. For both (A) and (B), a comparison of survival between indicated groups was performed using a Log rank Mantel-Cox test on the Kaplan-Meier survival data. Statistical significance convention used: n.s. $p > 0.05$; * $p \leq 0.05$; ** $p \leq 0.01$.

to the other variants. This result indicates lower furin cleavage efficiency likely due to the presence of a T residue at E2 position 1 (P1'). This is confirmed by our bioinformatics analyses using ProP 1.0 and PiTou 3 (Table S5) with both algorithms giving the lowest furin cleavability scores for the T1-AAA E2 variant and that variations at triplet 7-8-9 had little effect, most likely because these residues are located more distally to the cleavage site. Thus, both biochemical (Figure 6C) and bioinformatics (Table S5) analyses concur, demonstrating that the T1 residue could be an important contributing factor to rSAV2 E2 T1-AAA's slowed growth kinetics (Figure 6B).

In vivo testing of the impact of E2 triplet 7-8-9 in rainbow trout

To examine the impact on virulence of the variations at E2 N-term and 7-8-9 triplet motif *in vivo* during the course of an infection, the rSAV2 analyzed for their replication kinetics were used to infect two distinct lines of juvenile rainbow trout: the INRAE Synthetic (Sy) and AP2 lines (Figure 7). The Sy line was used in previous SAV infection studies allowing for comparative analyses and represents a genetically diverse population while the AP2 line, an isogenic line derived from the Sy line, was chosen because of its high sensitivity to viral infections in particular to hemorrhagic septicemia virus-induced mortality (viral hemorrhagic septicemia

virus [VHSV]).³⁷ For each infection group, 25 rainbow trout were infected with each rSAV2 by intraperitoneal injection (i.p.) at a dose of 10^4 FFU/fish (average weights, Sy line 2.64 g; AP2 line 2.95 g). For both Sy and AP2 lines, a mock infection control group was included. Mortalities were then recorded over a period of 63 days. For the Sy line infection trial, as expected, infection with the rSAV2 E2 VAA variant, which contains an alanine at position 9 (A₉) and previously shown to be associated with a virulent phenotype *in vivo*¹¹ (Figure S1) induced mortalities starting at day 25 reaching a cumulative percent mortality (CPM) of 32% at the end of the trial (Figure 7A, Table 2). Infection with rSAV2 E2 VAV variant, with a valine residue at position 9, a substitution previously shown to be attenuating in the Sy line of rainbow trout^{11,36} (Figure S1), was found to be completely attenuated with no recorded mortalities during the course of the trial. Interestingly, infection by the rSAV2 AAV variant, containing the consensus triplet 7-8-9 motif resulted in zero recorded mortalities, indicating it is also fully attenuated, like the rSAV2 E2 VAV variant (Figure 7A, Table 2). In contrast, infection with the rSAV2 E2 T1-AAA variant resulted in mortalities, appearing later than infection by rSAV2 E2 VAA (start of mortalities at 38 d.p.i.) but reaching a similar final CPM of 29.2% at the end of the trial. rSAV2 E2 VAA and rSAV2 E2 T1-AAA infections were grouped in the same log rank Mantel-Cox statistical group (group A, Table 2) while rSAV2 E2 VAV and rSAV2 E2 AAV infections formed another statistical group (group B, Table 2). These results point to a critical role played by position 9 (alanine to valine substitution) on virulence *in vivo*, in contrast with position 7 where the same substitution has no impact.

Table 2. Statistical analyses of rSAV2 infection trials in rainbow trout Sy and AP2 lines

Sy line						
rSAV2 E2	AAV	VAA	VAV	T1-AAA	Stat. group	CPM (%)
AAV	–	0.0022 **	>0.9999 n.s.	0.0038 **	A	0
VAA	0.0022 **	–	0.0022 **	0.5708 n.s.	B	32
VAV	>0.9999 n.s.	0.0022 **	–	0.0038 **	A	0
T1-AAA	0.0038 **	0.5708 n.s.	0.0038 **	–	B	29.2
AP2 line						
rSAV2 E2	AAV	VAA	VAV	T1-AAA	Stat. group	CPM (%)
AAV	–	0.1317 n.s.	0.2677 n.s.	0.0762 n.s.	AB	12.5
VAA	0.1317 n.s.	–	0.0099 **	0.7556 n.s.	B	32
VAV	0.2677 n.s.	0.0099 **	–	0.0048 **	A	4
T1-AAA	0.0762 n.s.	0.7556 n.s.	0.0048 **	–	B	36

Pairwise *p* values (Log-rank Mantel-Cox test) calculated from Kaplan-Meier survival curves. CPM, cumulative percent mortality (%); Stat. group: Statistical group. The following convention was used for describing statistical significance: not significant (n.s.) $p > 0.05$; significant (*) $p \leq 0.05$; highly significant (**), $p \leq 0.01$.

The infection trial with the AP2 line generally gave similar results as those obtained with the Sy line but with rSAV2 E2 VAA and rSAV2 E2 T1-AAA infections giving rise to mortalities at earlier time points, with a start of mortalities beginning at 8 and 12 d.p.i., respectively (Figure 7B). Both infection with rSAV2 E2 VAA and rSAV2 E2 T1-AAA were associated with similar virulence levels, with 32% and 36% CPM at the end of the trial, forming a single statistical group (group A, Table 2) and (Figure 7B). Moreover, limited mortality was observed for both the attenuated rSAV2 E2 VAV (statistical group B) and rSAV2 E2 AAV infections (statistical group AB) with 4% and 12.5% CPM at the end of the trial, respectively (Figure 7B, Table 2). These results reveal overall higher virulence observed when rSAV2 infect AP2 line compared to the Sy line, a likely consequence of the higher sensitivity of the AP2 line to viral infections as observed in previous infection trials.

To test the genetic stability of recombinant viruses, AP2 rainbow trout that had succumbed post-challenge were sampled to extract viral RNA from their head kidney and spleen. RT-PCR amplification was performed and the E2 N-term region was analyzed by sequencing for positive samples (Figure S5). This confirms that the mutations introduced in rSAV2 E2 AAV, rSAV2 E2 VAA, and rSAV2 E2 T1-AAA remained stable as the respective E2 N-term variations introduced in each recombinant virus were all found to be present with no other mutations found (Figures S5A–S5C). In contrast, no positive amplification could be achieved for the only fish that died in the rSAV2 E2 VAV challenge group, putting into question whether the animal succumbed to causes other than rSAV2 infection.

In summary, the *in vivo* infection trials in both Sy and AP2 lines of rainbow trout demonstrate the essential role played by position 9 within the 7–8–9 triplet motif in regulating virulence via an alanine to valine substitution, with A9 associated with virulence and V9 associated with attenuation. The same substitution at the proximal position 7 had no impact on virulence. Finally, the rSAV2 E2 T1-AAA variant was found to be virulent in both rainbow trout lines despite having delayed growth kinetics and inefficient furin cleavage at the E3E2 junction.

DISCUSSION

Compared to their terrestrial counterparts, aquatic alphaviruses have been understudied, with SAV being the only representative species for which a substantial body of research is available, principally due to its disease impact on aquaculture fish species and as a WOAHL-listed viral pathogen.³⁸ This has left the field with scarce resources to tackle fundamental questions regarding fish alphavirus biology, evolution as well as the mechanisms underpinning host-pathogen interactions, virulence, and interspecies host jumping. Through an integrative approach combining *in silico* 3D reconstruction of an entire SAV virion, phylogenetic and structure-based comparative analyses with *in vitro* and *in vivo* experimental validations, this study lays the groundwork to tackle such questions. As fish alphavirus proteins lack experimentally determined structures, previous efforts to gain insights were based on classical homology modeling approaches like I-Tasser³⁹ and Swiss-Model.⁴⁰ For example, the former was used to locate a naturally occurring mutation in SAV3 E2 (P206S) involved in viral fitness and host adaptations⁴¹ and the latter enabled to map deletions and residue variations in SAV3 envelope proteins that arise during the course of experimental infection in susceptible hosts.⁴² Similarly, we have previously modeled the non-structural protein 2 (nsP2) of SAV2 using RaptorX,⁴³ a template-based protein structure modeling server, in an effort to determine its functional subdomains and investigate its inhibitory function on the RIG-I sensing pathway of the innate antiviral immune response.⁴⁴

In recent years, the field of structural bioinformatics has seen the rise of revolutionary approaches, such as AlphaFold,^{18,19} RoseTTAFold,^{21,45} and ESMFold⁴⁶ which are based on artificial intelligence, deep-learning, and protein language models. These advanced programs are able to predict the structures of proteins with high accuracy and at scale. This has led to the establishment of vast databases of predicted structures, such as the AlphaFold Protein Structure Database with over 200 million predicted structures (AFDB, <https://alphafold.ebi.ac.uk>) and ESM Metagenomic Atlas with over 700 million predicted structures

(<https://esmatlas.com>). However, AFDB excludes viral proteins as it does not handle viral polyproteins like the alphavirus structural and non-structural polyproteins and ESM Metagenomic Atlas lacks taxonomic data severely limiting the study of viral proteins. The Big Fantastic Virus Database (BFVD, <https://bfvd.foldseek.com>)⁴⁷ and the Viral AlphaFold Database (VAD, <https://data-sharing.atkinson-lab.com/vad/>)⁴⁸ have recently emerged to address this critical gap in predicted viral protein structural data and will greatly benefit advances in virology.

The above tools have already proven their power to study the function and evolution of viral proteins and to shed light on the so-called “viral dark matter.”⁴⁹ This is exemplified in a structure modeling study by Mutz and colleagues on Orthopoxvirus proteins demonstrating multiple cases of exaptation (recruitment) and inactivation of host cellular enzymes by this family of DNA viruses.⁵⁰ In a study on carp edema poxvirus (CEV) for which we took part, a combination of approaches involving structure prediction and structural alignment using the search tool FoldSeek³⁰ was undertaken and revealed the surprising structural relationship between one of its proteins, encoded by the *cds46* gene, and host cell endonucleases of fish.⁵¹ This finding hints on a possible cellular origin of *cds46* by a gene capture event in a common viral ancestor. The current work underscores the advantage of performing structural alignments, as was done here using FoldMason²⁹ with alphavirus E1, E2, and Cp predicted protein structures to reconstruct structure-based viral phylogenies. Indeed, the rapid evolution of viruses, especially RNA viruses, can be an obstacle to establish clear evolutionary relationships using sequence analysis alone. Because protein structures and folds tend to be more conserved than sequences,^{52,53} structure-based alignments can overcome this limitation and inform on evolutionary relationships over vast time scales. Mifsud and collaborators have recently combined predictive structural bioinformatics with phylogenetic analyses to survey the landscape of *Flaviviridae* family glycoprotein structures.⁵⁴ The authors reveal a complex evolutionary history punctuated by bacterial gene capture and possible recombination events between members of distinct genera.

Thanks to these major advances, the current work takes a holistic approach by modeling entire structures of SAV envelope proteins (E1, E2, and E3), as well as the ordered region of Cp leveraging on the power and accuracy of AlphaFold 3.¹⁸ A complete model of an SAV virion was obtained by assembling individual protein models into higher-order hierarchical structures by mapping E2-E1-Cp subunits onto the asymmetric unit and icosahedral lattice of VEEV, a reference alphavirus for which a high-resolution cryo-EM structure has been determined. This constitutes a powerful blueprint for further structure-function studies for SAV and other fish alphaviruses, in particular to examine the functional consequences of mutations and to better understand the biology of aquatic alphaviruses and their evolutionary relationships with terrestrial alphaviruses.

Unexpectedly, modeling of the entire SAV virion revealed that the E2 protein harbored a singular surface-exposed N-term α -helix located in the β -ribbon connecting linker. This helix was confirmed by RoseTTAFold²¹ and PsiPred.²² It stands out because high-resolution X-ray crystallography studies have previously established that CHIKV and SINV E2 ectodomains were

structured as all- β proteins belonging to the immunoglobulin superfamily.^{8,55} Further evidence for the N-term α -helix comes from the comparative structural-evolutionary analysis performed in this study showing that the E2 protein of all fish alphaviruses are predicted to contain this feature, which corresponds in part to the shared N-term insert identified in the initial alphavirus E2 protein alignment. Intriguingly, while an N-term α -helix was found to be absent in VEEV and CHIKV terrestrial alphaviruses, a shorter α -helix was predicted for the aquatic alphaviruses infecting sea mammals SESV and AHPV. Since alphaviruses are considered to be of marine origin,² these findings point toward a hypothesis in which the N-term α -helix could represent an ancestral trait conserved in alphaviruses from the aquatic environment. It would be of interest to explore whether the absence of the N-term α -helix in extant representatives of terrestrial alphaviruses was the result of a gradual loss during alphavirus evolution from their original marine environment to their adaptation to terrestrial life cycle infecting new mammalian, bird, and invertebrate hosts. To examine the impact of introducing the SAV E2 N-term α -helix in a terrestrial alphavirus, we modeled with AlphaFold 3 a chimeric construct composed of the ectodomains of CHIKV p62-E1 with the stretch of CHIKV E2 residues spanning positions 6–13 replaced with the residues composing the SAV E2 N-term insert α -helix (aa 6–22) (Figure S3C). The modeled structure is quasi-identical to the X-ray-determined crystal structure (superimposed with PDB: 3n40, further validating the use of AlphaFold 3 as a robust structure prediction tool), with the notable exception of the N-terminus of E2 which displays the same N-term α -helical secondary structure as the one observed in SAV models and not seen in the CHIKV crystal structure. Similarly to SAV p62, the modified CHIKV p62 protein displays a furin loop that extends more outwardly from p62. Since E2 is a major determinant of viral entry and tropism, the distinctive α -helix structural feature identified in SAV E2 N-term may offer clues to a unique aspect of its life cycle namely its mode of transmission that does not necessitate a vector. Indeed, SAV can transmit from fish hosts directly via water, in contrast to terrestrial alphaviruses infecting mammals and birds that require mosquito vectors for transmission.

The E2 protein alignment highlights that SAV is unique among fish alphaviruses in harboring the 7-8-9 triplet motif corresponding to the first three residues of its E2 N-term α -helix as no other fish alphavirus included in our work harbor this motif (WHAV, CAV, WCFVAV and WSFAV). Whether the residues surrounding the 7-8-9 triplet insert in WHAV, CAV, WCFVAV and WSFAV play an equivalent role is an open question. However, to our knowledge the reverse genetics tools that would allow to answer this question are only available for SAV (SAV2 and SAV3). Furthermore, the identification and genomic characterization of new species of fish alphaviruses may shed light on these questions. Previous work conducted in our laboratory using SAV reverse genetics and *in vivo* rainbow trout challenge demonstrated that an A9V substitution in E2 resulted in attenuation (incorrectly assigned to position 8 in Merour et al.,¹¹ see Figure S1). This substitution likely arose spontaneously during the establishment of the reverse genetics system for SAV in our laboratory (based on a subtype 2 SAV).³⁶ One of the main challenges toward building a phylogenomic framework is the

availability of genomic and metagenomic data that are solidly connected to phylogenetic analyses, epidemiological and experimental data of virulence and fitness.⁵⁶ In a first attempt to apply this framework to SAV, we have undertaken a focused phylogenetic-based approach examining SAV E2 to investigate whether variations found in field isolates and occurring on the triplet 7-8-9 motif of isolates could also play a role in regulating virulence *in vivo*. While clinical metadata such as disease symptoms and severity associated with isolates are still largely lacking, renewed efforts in characterizing SAV genomes have substantially increased the sequence diversity landscape of isolates. This presents an opportunity to bridge the gap between *in silico*, *in vitro* and *in vivo* data on virulence-impacting mutations.⁵⁶

Examples of how such approaches can lead to insights in the evolution of viral pathogenicity are found in studies on the fish novirhabdovirus infectious hematopoietic necrosis virus (IHNV) that have revealed that cross-species transmission from sockeye salmon to rainbow trout was first accompanied with decreased virulence in its new host followed by a recovery of virulence for some genotypes, which was associated with increased transmissibility.^{57,58} More recently, molecular virulence markers of rainbow trout infection have been characterized scattered throughout the genome of the related VHSV.⁵⁹ Much like for SAV, the study showed that single amino acid changes could drastically alter pathogenicity. Phylogenetic analyses also revealed the evolution toward increased virulence for certain subtypes.⁶⁰

The E2 phylogenetic analysis performed here is based on an extensive set of sequences from 90 SAV isolates. The most frequent triplet 7-8-9 motif observed was γ AAV₉ (77/90), which was found in all SAV3 sequences analyzed. For subtypes 1, 2, and 5, variations were observed at positions 7 and 9 but never at position 8 (A₈) which remained invariant for all isolates analyzed. The E2 protein alignment also revealed that sequences from SAV1 and SAV4 isolates had N-term deletions spanning either a part or the entire triplet motif. Deletions in SAV structural proteins (E1 and E2) appear to be fairly common occurrences as reported by Roh and colleagues in a study based on Nanopore sequencing that followed variations occurring in SAV3 genomes during the course of infection in Atlantic salmon and brown trout.⁴² The majority of the deletions arising spontaneously during infection are believed to be deleterious for the virus. Similar deletion variants were reported by a previous study which was also based on Nanopore sequencing.⁶¹ These findings echo our inability to recover rSAV2 E2 Δ_{4-9} .

Since disease status and severity metadata was lacking for the isolates with triplet 7-8-9 variations, we introduced these variations by site-directed mutagenesis in our rSAV2 system and tested them *in vitro* in susceptible BF-2 cells and *in vivo* in rainbow trout. *In vitro*, the rSAV2 E2 with the γ AAV₉ (rSAV2 E2 AAV) triplet, found in the majority of field isolates, had similar growth characteristics as rSAV2 E2 VAA and rSAV2 E2 VAV variants which were obtained previously. However, the capacity to infect BF-2 cell cultures was severely limited for rSAV2 E2 AAA, AAD, AAF variants, possibly due to a loss in viral fitness. Specifically, introduction of a bulky hydrophobic residue (F) or the change to a negatively charged residue (D) at position 9 (AAF

and AAD) led to a strong decrease in infectivity. Surprisingly, a more subtle change from AAV to AAA involving alanine and valine small hydrophobic amino acids can also profoundly impact infectivity (AAA), a deleterious effect that is compensated by substituting the small hydrophobic alanine with a polar uncharged threonine at position 1 (T1-AAA).

In contrast to the rSAV2 E2 AAA variant, the rSAV2 E2 T1-AAA was able to infect BF-2 cells albeit with a delay in its growth kinetics compared to rSAV2 E2 AAV. Biochemical analyses revealed that the T residue at position 1 (T₁) led to less efficient furin proteolytic processing of its p62 protein. We have further analyzed the conservation of E2 position 1 from the alignment of 90 SAV sequences used to generate the Weblogo. In this analysis, one sequence (SAV4 04-44) was omitted since it contained a deletion of its furin cleavage site. We find that A was the dominant residue at this position (present in 91% of sequences), followed by T residue (7.9%, all from SAV5) and one sequence with a D residue (1.1%, SAV6). The biochemical analysis clearly showed that T₁ decreases E3E2 furin cleavage compared to A₁, confirming that this position impacts p62 cleavage and viral growth. This also suggests that SAV5 and SAV6 may have evolved distinct proteolytic processing profiles. More broadly, these results suggest possible epistatic effects between triplet 7-8-9 motif residues and the N-term of E2. The recombinants which could replicate efficiently *in vitro* were then tested *in vivo* in two lines of rainbow trout (Sy and AP2). This analysis confirmed that for both rainbow trout lines an alanine to valine substitution at position 9 was accompanied by attenuation and that the same alanine to valine substitution at position 7 had no effects on virulence. From a structure-function vantage point, position 9 can be viewed as a mutationally sensitive virulence switch. Further testing of other variations at this site either alone or in combination, as found in the different SAV subtypes, could inform whether other residues can alter virulence. As our reverse genetics system is based on the SAV2 subtype, it would also be interesting to perform such analyses using other SAV genetic backgrounds, such as with the rSAV3 backbone.

While our study focused on the triplet 7-8-9 and the first position of E2, it is important to emphasize other sites in E2 as well as other proteins can also impact fitness and virulence. In previous work from our group, another spontaneous E2 substitution, T136M, was shown to have an attenuating effect during *in vivo* rainbow trout challenge, albeit to a lesser degree than the A9V substitution as the latter accounted for 90% of the attenuation.¹¹ Close inspection of the protein alignment of the 90 isolates in the current study revealed that none contained a methionine at position 136 (M136) and that T₁₃₆ is quasi-invariant with only a single isolate, SAV3 8-R10 harboring an alanine, A₁₃₆ (GenBank: [KC122921.1](https://www.ncbi.nlm.nih.gov/nuclot/KC122921.1)). However, the impact of the T136A substitution on virulence remains to be elucidated.

During the course of this work, Braaen and colleagues reported to have generated and tested *in vivo* in Atlantic salmon a recombinant virus based on a SAV3 backbone and harboring the A8V and T136M substitutions in its E2 protein.⁶² The authors showed that the recombinant virus was attenuated compared to the parental strain suggesting that an alanine to valine substitution at position 8 could potentially also alter pathogenicity. However, the authors have not tested the contribution of each

substitution. Also, the infection models (Atlantic salmon and rainbow trout) and reverse genetics backbones (rSAV3 and rSAV2) are different so the results of Braaen et al. may not be directly transposable to this current work. Another case where a single residue change can have a profound phenotypical impact has been studied by Karlsen and collaborators who have identified the P206S substitution in SAV3 E2, which is located in domain B of the protein and appeared during an epizootic episode from a wild reservoir species to farmed salmon in Norway.⁴¹ The authors tested the substitution using reverse genetics and found that *in vitro* the recombinant virus with the substitution had slower replication kinetics and lower fitness compared to the parental strain but was found to be shed and transmitted more rapidly in the water in an *in vivo* Atlantic salmon cohabitation transmission model.

While we still lack a mechanistic picture for understanding the reasons why substitutions in the N-term of E2 can lead to drastic changes in virulence, a few hypotheses can be put forward. Indeed, the modeling of the entire SAV virion clearly shows that the predicted N-term α -helix containing the triplet 7-8-9 is surface-exposed and repeated 240 times on the icosahedral structure of the virus. As such, it becomes evident that the effects of even a modest modification, such as a single substitution, can be amplified by its repetition across the entire icosahedral surface. In addition, while the receptor of SAV has yet to be identified, the location of these E2 structural features, which are exposed to the external environment, could allow them to be directly involved in host cell receptor interactions.

Furthermore, the close proximity of the triplet 7-8-9 motif to the N-term of E2 suggests it could play a role in interactions with the E3 companion protein, the pH-dependent binding and dissociation of which can impact stability of the E2-E1 heterodimer and exposure of the fusion loop.⁶³ During our examination of the different stages of p62 maturation by furin cleavage at the E3E2 junction (₆₈RKKR₇₁↓₁AVST₄), we found that the triplet 7-8-9 residues located at the tip of the N-term α -helix were predicted to remain exposed in the presence of E3, whether covalently linked to E2 or not. It is noteworthy to highlight that the furin cleavage site, which is composed of the canonical ₆₈RKKR₇₁ basic motif at the C-term of E3, is predicted to form an unstructured and exposed loop, quite similarly to furin activation loops observed for other viral envelope glycoproteins such as coronavirus spike proteins.^{64,65} Comparative analysis with CHIKV showed that SAV E2 N-term α -helix does not change E3 position's relative to E2 but makes the furin loop protrude more outwardly from p62. To test the effects of substitutions in the E2 N-term region on furin cleavability, we have conducted an *in silico* analysis using ProP 1.0⁶⁶ and PiTou 3⁶⁷ furin prediction programs (Table S5). This analysis shows that the A1T substitution (T1-AAA variant), which is the closest to the furin cleavage scissile bond (involving residue P1-P1' of the cleavage site, P1 corresponding to the last residue of E3 and P1' to the first residue of E2), had a noticeable inhibitory effect on furin cleavability as it scored the lowest prediction values for both ProP and PiTou algorithms. These bioinformatics analyses confirm the biochemical assays on SAV p62 proteolytic processing by furin. The other mutations involving triplet 7-8-9 residues had little effect on furin cleavage, likely because they are located more distally to the cleavage site.

Regarding the substitution within the triplet 7-8-9 motif, it is remarkable that even subtle changes like A9V can severely impact pathogenicity. Both alanine and valine are uncharged hydrophobic amino acids and have similar biochemical profiles on first approximation. However, they differ in hydropathy scoring, with alanine having the lowest hydropathy score (1.8) among hydrophobic amino acids and valine having the second highest score (4.2), after isoleucine. Furthermore, a notable difference between alanine and valine lies in their side chain structures. While alanine has a straight-chain aliphatic sidechain, valine has a β -branched side chain.⁶⁸ This key difference explains why alanine is considered to be among the best helix-forming residues, while valine is invariably a poor helix former. These distinct properties were used by Gregoret and Sauer to biochemically measure the tolerance of an α -helix of a native protein to combinatorial substitutions with alanine and valine residues. The authors showed that tertiary interactions within the protein mitigate helix destabilization caused by valine substitutions.⁶⁸ Taking into account the biochemical properties of these amino acids, it is reasonable to posit that alanine to valine substitution at position 9 could destabilize the predicted N-term α -helix.

As AlphaFold is unable to predict the effect of single-residue changes on protein structure and stability,⁶⁹ we have performed an additional *in silico* analysis using Pythia, a recently developed deep learning model capable of rapid and accurate estimation of mutation-associated change in Gibb's free energy of folding ($\Delta\Delta G$), a measure of the outcome of an amino acid change on protein stability (Figure S6).⁷⁰ The AlphaFold 3 predicted structure of the E2 ectodomain (E2_{ecto} aa 1–352, GenBank: AJ316246.1 reference sequence, VAV 7-8-9 triplet motif) was used as input and the program generated a comprehensive heatmap encompassing all residues of E2_{ecto}, providing scores at each position and for every amino acid change (Figure S6). A lower score (darker, blue shading) indicates that the substitution is likely to stabilize the protein, while a higher energy value (lighter, yellow shading) indicates that the substitution is more likely to be destabilizing. When focusing on the triplet 7-8-9 motif, mutations on positions 7 and 9 were on the whole more stabilizing (lower scores) while mutations at position 8 were more destabilizing (higher scores). This suggests that position 8 may be less tolerant to variations compared to positions 7 and 9, which is in agreement with the E2 protein alignment and AL2CO scoring showing that position 8 had an invariant alanine, while variations were observed for positions 7 and 9. Regarding the latter positions, Pythia scores for valine to alanine substitutions were negative, -2.3974 and -0.4327 , respectively, indicating that these amino acid changes were predicted to be stabilizing. These results are in line with alanine being a helix former and valine as a helix destabilizer. This warrants further biochemical examination of SAV E2 and the N-term α -helix to confirm these results.

This work harnesses the power of structural bioinformatics and lays the foundation for future studies on the evolution of alphaviruses. This is particularly timely as the effects of climate change on sea and freshwater temperatures have already taken by surprise many experts⁷¹ and the consequences on the evolutionary trajectories of aquatic viruses remain to be explored. Fish alphaviruses infect ectotherm hosts that have body temperatures that vary according to their environment.

Whether these viruses will adapt to increasing temperatures due to climate change is an open question. Our group has already obtained evidence that Novirhabdoviruses, including VHSV and IHNV, can adapt to higher temperatures (unpublished). For fish alphaviruses, indirect evidence that such adaptation is possible can be found in the known temperature ranges of terrestrial alphaviruses such as CHIKV that infect both ectothermic mosquitoes and endothermic mammalian hosts *in vitro*,⁷² indicating a high degree of plasticity in their permissive temperature range. A more direct and tantalizing argument comes from our group's previous work showing that a variant recombinant virus (rSDV14) was successfully serially passaged in cell culture at 14°C, well above the temperature of the parental strain originally grown at 10°C.³⁶ Intriguingly, the thermo-adapted strain was shown to gain virulence in a rainbow trout challenge model. The molecular basis for this temperature adaptation and the link with virulence remains to be elucidated. The innovative and integrative approaches used in the current study, notably by mapping virulence-associated mutations on predicted structures combined with *in vitro* and *in vivo* functional testing, can thus be applied to shed light on these phenomena. This would address one of the most challenging threats facing aquaculture globally and, more broadly, would have important ramifications for our understanding of alphavirus ecology and evolution.

Limitations of the study

A limit of using AlphaFold is that it generates single conformation protein structures. Alphavirus E1 and E2 proteins adopt different conformations during endosomal entry, where low pH triggers conformational changes. Despite this, AlphaFold coherently generated the prefusion conformations of both SAV E1 and E2. These models fitted well with the VEEV icosahedral lattice used for the virion reconstruction and correspond to the virion conformation before cellular entry.

The E2 alignment and phylogenetic analyses introduce an inherent sampling bias due to variations in the number of sequences available depending on the subtype, location, and host. As such, the phylogenetic and sequence variation analyses performed may not closely reflect the frequencies of variations that exist in the field. However, such analysis can provide a qualitative appreciation of substitutions at key positions that can be of functional and evolutionary importance, as highlighted in our study.

We have taken a targeted approach focusing on variations found in SAV E2 N-term. It is important to consider that other sites in E2 and other proteins likely play important roles in modulating virulence, such as E2 A8V, as performed recently by Braaen et al.⁶² By leveraging the powerful structural model of SAV we have developed, future studies can extend the functional analyses to other sites for SAV and possibly other related fish alphaviruses.

Here, *in vivo* rSAV2 challenges were performed on rainbow trout, using two distinct lines, Sy and AP2, to assess the effects of host genetic background. For future work, it would be of interest to use a salmon model and other rSAV subtypes to gain insights into how E2 N-term variations are linked with host adaptation.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Jean K. Millet (jean.millet@inrae.fr).

Materials availability

Reagents (including plasmids) generated in this study are available from the [lead contact](#).

Data and code availability

- The 3D reconstructed SAV virion model was deposited in the ModelArchive (Swiss Institute of Bioinformatics, SIB; <https://www.modelarchive.org>) with the following identifier: ModelArchive: [ma-gkhkz](#). This deposit also includes the individual models for SAV E2-E1 heterodimer and Cp. The aquatic alphavirus E2-E1-Cp heterotrimer models were deposited with the following identifiers: AHPV, ModelArchive: [ma-xgzne](#); SESV, ModelArchive: [ma-zwpvs](#); WHAV, ModelArchive: [ma-md6n0](#); CAV, ModelArchive: [ma-zvp1t](#); WCF, ModelArchive: [ma-6kpi5](#); and WSAV, ModelArchive: [ma-jt78w](#).
- This article does not report original code.
- Any additional information required to reanalyze the data reported in this article is available upon request from the [lead contact](#).

ACKNOWLEDGMENTS

We would like to thank the late Michel Brémont for his guidance and support in the initial stages of this work and acknowledge the members of our team (équipe Virologie Moléculaire des Poissons) for helpful discussions. We gratefully acknowledge Florence Phocas (GABI Unit, INRAE) for proofreading our manuscript. We are very thankful for the valuable advice given by Gerardo Tauriello (Swiss Institute of Bioinformatics) for the SAV virion structure modeling and deposition in the ModelArchive. We warmly acknowledge the support from INRAE's rodent and fish experimentation unit (IERP, doi.org/10.15454/1.5572427140471238E12), part of the EMERG'IN national research infrastructure (doi.org/10.15454/90CK-Y371), and would like to thank Damien Degallaix, Pénélope Simon, Dimitri Rigaudeau, and Christelle Langevin for their excellent assistance for the *in vivo* studies. Calvin Fauvet is the recipient of a Ph.D. fellowship from the ABIES (Agriculture Food Biology Environment Health) Doctoral School, AgroParisTech. This work was funded in part by young researcher grants awarded to J.K.M. by the Animal Health Department of INRAE, France.

AUTHOR CONTRIBUTIONS

Conceptualization, S.B. and J.K.M.; methodology, S.B., C.F., E.M., J.B., A.L., D.L., and J.K.M.; investigation, S.B., E.M., and J.K.M.; writing – original draft, S.B. and J.K.M.; writing – review and editing, S.B., C.F., E.M., D.L., and J.K.M.; funding acquisition, J.K.M.; supervision, S.B. and J.K.M.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- [KEY RESOURCES TABLE](#)
- [EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS](#)
 - Animals used in this study
 - Ethics statement
 - Virus strain
 - Cell line
- [METHOD DETAILS](#)
 - Phylogenetic analyses

- AlphaFold-based fish alphavirus protein structure predictions
- *In silico* reconstruction of a whole SAV virion
- Structure-based phylogenetic alignment and trees
- Visualization of SAV E2 sequence conservation
- Predicted impact of N-term variations on SAV E2 protein stability
- Plasmid constructs and site-directed mutagenesis
- Recombinant SAV2 (rSAV2) recovery
- Western blot analyses of SAV E3E2 furin cleavage
- Virus titration and growth kinetics
- Immunofluorescence and fluorescence focus assays
- *In vivo* analyses of rSAV2 infection in rainbow trout
- **QUANTIFICATION AND STATISTICAL ANALYSES**

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.isci.2026.115070>.

Received: September 12, 2025

Revised: December 17, 2025

Accepted: February 13, 2026

Published: February 18, 2026

REFERENCES

1. Deperasinska, I., Schulz, P., and Siwicki, A.K. (2018). Salmonid Alphavirus (SAV). *J. Vet. Res.* 62, 1–6. <https://doi.org/10.2478/jvetres-2018-0001>.
2. Forrester, N.L., Palacios, G., Tesh, R.B., Savji, N., Guzman, H., Sherman, M., Weaver, S.C., and Lipkin, W.I. (2012). Genome-scale phylogeny of the alphavirus genus suggests a marine origin. *J. Virol.* 86, 2729–2738. <https://doi.org/10.1128/JVI.05591-11>.
3. Shi, M., Lin, X.D., Chen, X., Tian, J.H., Chen, L.J., Li, K., Wang, W., Eden, J.S., Shen, J.J., Liu, L., et al. (2018). The evolutionary history of vertebrate RNA viruses. *Nature* 556, 197–202. <https://doi.org/10.1038/s41586-018-0012-7>.
4. Zamperin, G., Milani, A., Gastaldelli, M., Quartesan, R., Fortin, A., Cappelozza, E., Patarnello, P., and Toffan, A. (2020). Genomic Sequence of a New Alphavirus Detected in Comber (*Serranus cabrilla*). *Microbiol. Resour. Announc.* 9, e01294.19. <https://doi.org/10.1128/MRA.01294-19>.
5. Villoing, S., Béarzotti, M., Chilmonczyk, S., Castric, J., and Brémont, M. (2000). Rainbow Trout Sleeping Disease Virus Is an Atypical Alphavirus. *J. Virol.* 74, 173–183. <https://doi.org/10.1128/JVI.74.1.173-183.2000>.
6. Weston, J., Villoing, S., Brémont, M., Castric, J., Pfeffer, M., Jewhurst, V., McLoughlin, M., Rødseth, O., Christie, K.E., Koumans, J., and Todd, D. (2002). Comparison of two aquatic alphaviruses, salmon pancreas disease virus and sleeping disease virus, by using genome sequence analysis, monoclonal reactivity, and cross-infection. *J. Virol.* 76, 6155–6163. <https://doi.org/10.1128/jvi.76.12.6155-6163.2002>.
7. Petterson, E., Sandberg, M., and Santi, N. (2009). Salmonid alphavirus associated with *Lepeophtheirus salmonis* (Copepoda: Caligidae) from Atlantic salmon, *Salmo salar* L. *J. Fish. Dis.* 32, 477–479. <https://doi.org/10.1111/j.1365-2761.2009.01038.x>.
8. Voss, J.E., Vaney, M.C., Duquerroy, S., Vonnrhein, C., Girard-Blanc, C., Crublet, E., Thompson, A., Bricogne, G., and Rey, F.A. (2010). Glycoprotein organization of Chikungunya virus particles revealed by X-ray crystallography. *Nature* 468, 709–712. <https://doi.org/10.1038/nature09555>.
9. Zhang, R., Hryc, C.F., Cong, Y., Liu, X., Jakana, J., Gorchakov, R., Baker, M.L., Weaver, S.C., and Chiu, W. (2011). 4.4 Å cryo-EM structure of an enveloped alphavirus Venezuelan equine encephalitis virus. *EMBO J.* 30, 3854–3863. <https://doi.org/10.1038/emboj.2011.261>.
10. Kim, A.S., and Diamond, M.S. (2023). A molecular understanding of alpha-virus entry and antibody protection. *Nat. Rev. Microbiol.* 21, 396–407. <https://doi.org/10.1038/s41579-022-00825-7>.
11. Merour, E., Lamoureux, A., Bernard, J., Biacchesi, S., and Brémont, M. (2013). A fully attenuated recombinant Salmonid alphavirus becomes pathogenic through a single amino acid change in the E2 glycoprotein. *J. Virol.* 87, 6027–6030. <https://doi.org/10.1128/JVI.03501-12>.
12. Gorchakov, R., Wang, E., Leal, G., Forrester, N.L., Plante, K., Rossi, S.L., Partidos, C.D., Adams, A.P., Seymour, R.L., Weger, J., et al. (2012). Attenuation of Chikungunya virus vaccine strain 181/clone 25 is determined by two amino acid substitutions in the E2 envelope glycoprotein. *J. Virol.* 86, 6084–6096. <https://doi.org/10.1128/JVI.06449-11>.
13. Tsetsarkin, K.A., Vanlandingham, D.L., McGee, C.E., and Higgs, S. (2007). A single mutation in chikungunya virus affects vector specificity and epidemic potential. *PLoS Pathog.* 3, e201. <https://doi.org/10.1371/journal.ppat.0030201>.
14. Merour, E., Lamoureux, A., Biacchesi, S., and Brémont, M. (2016). Fine mapping of a salmonid E2 alphavirus neutralizing epitope. *J. Gen. Virol.* 97, 893–900. <https://doi.org/10.1099/jgv.0.000411>.
15. Klimstra, W.B., Ryman, K.D., and Johnston, R.E. (1998). Adaptation of Sindbis virus to BHK cells selects for use of heparan sulfate as an attachment receptor. *J. Virol.* 72, 7357–7366. <https://doi.org/10.1128/JVI.72.9.7357-7366.1998>.
16. Kinney, R.M., Chang, G.J., Tsuchiya, K.R., Sneider, J.M., Roehrig, J.T., Woodward, T.M., and Trent, D.W. (1993). Attenuation of Venezuelan equine encephalitis virus strain TC-83 is encoded by the 5′-noncoding region and the E2 envelope glycoprotein. *J. Virol.* 67, 1269–1277. <https://doi.org/10.1128/JVI.67.3.1269-1277.1993>.
17. Vashishtha, M., Phalen, T., Marquardt, M.T., Ryu, J.S., Ng, A.C., and Kielian, M. (1998). A single point mutation controls the cholesterol dependence of Semliki Forest virus entry and exit. *J. Cell Biol.* 140, 91–99. <https://doi.org/10.1083/jcb.140.1.91>.
18. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A.J., Bambrick, J., et al. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 630, 493–500. <https://doi.org/10.1038/s41586-024-07487-w>.
19. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Židek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589. <https://doi.org/10.1038/s41586-021-03819-2>.
20. Vaney, M.C., Duquerroy, S., and Rey, F.A. (2013). Alphavirus structure: activation for entry at the target cell surface. *Curr. Opin. Virol.* 3, 151–158. <https://doi.org/10.1016/j.coviro.2013.04.003>.
21. Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G.R., Wang, J., Cong, Q., Kinch, L.N., Schaeffer, R.D., et al. (2021). Accurate prediction of protein structures and interactions using a three-track neural network. *Science* 373, 871–876. <https://doi.org/10.1126/science.abj8754>.
22. McGuffin, L.J., Bryson, K., and Jones, D.T. (2000). The PSIPRED protein structure prediction server. *Bioinformatics* 16, 404–405. <https://doi.org/10.1093/bioinformatics/16.4.404>.
23. Combet, C., Blanchet, C., Geourjon, C., and Deléage, G. (2000). NPS@: Network Protein Sequence Analysis. *Trends Biochem. Sci.* 25, 147–150. [https://doi.org/10.1016/S0968-0004\(99\)01540-6](https://doi.org/10.1016/S0968-0004(99)01540-6).
24. Uchime, O., Fields, W., and Kielian, M. (2013). The role of E3 in pH protection during alphavirus assembly and exit. *J. Virol.* 87, 10255–10262. <https://doi.org/10.1128/JVI.01507-13>.
25. Fields, W., and Kielian, M. (2013). A key interaction between the alphavirus envelope proteins responsible for initial dimer dissociation during fusion. *J. Virol.* 87, 3774–3781. <https://doi.org/10.1128/JVI.03310-12>.
26. Wu, S.R., Haag, L., Sjöberg, M., Garoff, H., and Hammar, L. (2008). The dynamic envelope of a fusion class II virus. E3 domain of glycoprotein E2 precursor in Semliki Forest virus provides a unique contact with the fusion protein E1. *J. Biol. Chem.* 283, 26452–26460. <https://doi.org/10.1074/jbc.M801470200>.

27. Moriette, C., LeBerre, M., Boscher, S.K., Castric, J., and Brémont, M. (2005). Characterization and mapping of monoclonal antibodies against the Sleeping disease virus, an aquatic alphavirus. *J. Gen. Virol.* *86*, 3119–3127. <https://doi.org/10.1099/vir.0.81030-0>.
28. Chen, L., Wang, M., Zhu, D., Sun, Z., Ma, J., Wang, J., Kong, L., Wang, S., Liu, Z., Wei, L., et al. (2018). Implication for alphavirus host-cell entry and assembly indicated by a 3.5Å resolution cryo-EM structure. *Nat. Commun.* *9*, 5326. <https://doi.org/10.1038/s41467-018-07704-x>.
29. Gilchrist, C.L.M., Mirdita, M., and Steinegger, M. (2024). Multiple Protein Structure Alignment at Scale with FoldMason. Preprint at bioRxiv. <https://doi.org/10.1101/2024.08.01.606130>.
30. van Kempen, M., Kim, S.S., Tumescheit, C., Mirdita, M., Lee, J., Gilchrist, C.L.M., Söding, J., and Steinegger, M. (2024). Fast and accurate protein structure search with Foldseek. *Nat. Biotechnol.* *42*, 243–246. <https://doi.org/10.1038/s41587-023-01773-0>.
31. Hjortaa, M.J., Fringuelli, E., Monjane, A.L., Mikalsen, A.B., Jonassen, C.M., Savage, P., and Sindre, H. (2021). Emergence of Salmonid Alphavirus Genotype 2 in Norway-Molecular Characterization of Viral Strains Circulating in Norway and Scotland. *Viruses* *13*, 1556. <https://doi.org/10.3390/v13081556>.
32. Macqueen, D.J., Eve, O., Gundappa, M.K., Daniels, R.R., Gallagher, M.D., Alexandersen, S., and Karlsen, M. (2021). Genomic Epidemiology of Salmonid Alphavirus in Norwegian Aquaculture Reveals Recent Subtype-2 Transmission Dynamics and Novel Subtype-3 Lineages. *Viruses* *13*, 2549. <https://doi.org/10.3390/v13122549>.
33. Gallagher, M.D., Matejusova, I., Ruane, N.M., and Macqueen, D.J. (2020). Genome-wide target enriched viral sequencing reveals extensive 'hidden' salmonid alphavirus diversity in farmed and wild fish populations. *Aquaculture* *522*, 735117. <https://doi.org/10.1016/j.aquaculture.2020.735117>.
34. Fringuelli, E., Rowley, H.M., Wilson, J.C., Hunter, R., Rodger, H., and Graham, D.A. (2008). Phylogenetic analyses and molecular epidemiology of European salmonid alphaviruses (SAV) based on partial E2 and nsP3 gene nucleotide sequences. *J. Fish. Dis.* *31*, 811–823. <https://doi.org/10.1111/j.1365-2761.2008.00944.x>.
35. Pei, J., and Grishin, N.V. (2001). AL2CO: calculation of positional conservation in a protein sequence alignment. *Bioinformatics* *17*, 700–712. <https://doi.org/10.1093/bioinformatics/17.8.700>.
36. Moriette, C., LeBerre, M., Lamoureux, A., Lai, T.L., and Brémont, M. (2006). Recovery of a recombinant salmonid alphavirus fully attenuated and protective for rainbow trout. *J. Virol.* *80*, 4088–4098. <https://doi.org/10.1128/JVI.80.8.4088-4098.2006>.
37. Quillet, E., Dorson, M., Le Guillou, S., Benmansour, A., and Boudinot, P. (2007). Wide range of susceptibility to rhabdoviruses in homozygous clones of rainbow trout. *Fish Shellfish Immunol.* *22*, 510–519. <https://doi.org/10.1016/j.fsi.2006.07.002>.
38. Costa, V.A., and Holmes, E.C. (2024). Diversity, evolution, and emergence of fish viruses. *J. Virol.* *98*, e0011824. <https://doi.org/10.1128/jvi.00118-24>.
39. Roy, A., Kucukural, A., and Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* *5*, 725–738. <https://doi.org/10.1038/nprot.2010.5>.
40. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., et al. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* *46*, W296–W303. <https://doi.org/10.1093/nar/gky427>.
41. Karlsen, M., Andersen, L., Blindheim, S.H., Rimstad, E., and Nylund, A. (2015). A naturally occurring substitution in the E2 protein of Salmonid alphavirus subtype 3 changes viral fitness. *Virus Res.* *196*, 79–86. <https://doi.org/10.1016/j.virusres.2014.11.011>.
42. Roh, H., Skafnesmo, K.O., Kannimuthu, D., Madhun, A., Patel, S., Kvamme, B.O., Morton, H.C., and Grove, S. (2024). Nanopore sequencing provides snapshots of the genetic variation within salmonid alphavirus-3 (SAV3) during an ongoing infection in Atlantic salmon (*Salmo salar*) and brown trout (*Salmo trutta*). *Vet. Res.* *55*, 106. <https://doi.org/10.1186/s13567-024-01349-z>.
43. Kallberg, M., Wang, H., Wang, S., Peng, J., Wang, Z., Lu, H., and Xu, J. (2012). Template-based protein structure modeling using the RaptorX web server. *Nat. Protoc.* *7*, 1511–1522. <https://doi.org/10.1038/nprot.2012.085>.
44. Jami, R., Mérour, E., Bernard, J., Lamoureux, A., Millet Jean, K., and Biacchesi, S. (2021). The C-Terminal Domain of Salmonid Alphavirus Nonstructural Protein 2 (nsP2) Is Essential and Sufficient To Block RIG-I Pathway Induction and Interferon-Mediated Antiviral Response. *J. Virol.* *95*, e01155. <https://doi.org/10.1128/JVI.01155-21>.
45. Krishna, R., Wang, J., Ahern, W., Sturmfels, P., Venkatesh, P., Kalvet, I., Lee, G.R., Morey-Burrows, F.S., Anishchenko, I., Humphreys, I.R., et al. (2024). Generalized biomolecular modeling and design with RoseTTAFold All-Atom. *Science* *384*, ead12528. <https://doi.org/10.1126/science.ad12528>.
46. Lin, Z., Akin, H., Rao, R., Hie, B., Zhu, Z., Lu, W., Smetanin, N., Verkuil, R., Kabeli, O., Shmuelli, Y., et al. (2023). Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science* *379*, 1123–1130. <https://doi.org/10.1126/science.ade2574>.
47. Kim, R.S., Levy Karin, E., Mirdita, M., Chikhi, R., and Steinegger, M. (2025). BFVD—a large repository of predicted viral protein structures. *Nucleic Acids Res.* *53*, D340–D347. <https://doi.org/10.1093/nar/gkae1119>.
48. Odai, R., Leemann, M., Al-Murad, T., Abdullah, M., Shyrokova, L., Tenson, T., Hauryluk, V., Durairaj, J., Pereira, J., and Atkinson, G.C. (2025). The Viral AlphaFold Database of monomers and homodimers reveals conserved protein folds in viruses of bacteria, archaea, and eukaryotes. *Sci. Adv.* *11*, eadz8560. <https://doi.org/10.1126/sciadv.adz8560>.
49. Mutz, P., Camargo, A.P., Sahakyan, H., Neri, U., Butkovic, A., Wolf, Y.I., Krupovic, M., Dolja, V.V., and Koonin, E.V. (2025). The protein structure of Orthornavirae and its dark matter. *mBio* *16*, e0320024. <https://doi.org/10.1128/mbio.03200-24>.
50. Mutz, P., Resch, W., Faure, G., Senkevich, T.G., Koonin, E.V., and Moss, B. (2023). Exaptation of Inactivated Host Enzymes for Structural Roles in Orthopoxviruses and Novel Folds of Virus Proteins Revealed by Protein Structure Modeling. *mBio* *14*, e0040823. <https://doi.org/10.1128/mbio.00408-23>.
51. Montacq, L., Flores, D., Giummarra, H., Pallandre, L., Angot, A., Thomas, R., Charrier, A., Lamothe, L., Lesne, M., Bellet, C., et al. (2024). cds46, a highly variable carp edema virus gene. *J. Gen. Virol.* *105*, 002048. <https://doi.org/10.1099/jgv.0.002048>.
52. Illergard, K., Ardell, D.H., and Elofsson, A. (2009). Structure is three to ten times more conserved than sequence—a study of structural response in protein cores. *Proteins* *77*, 499–508. <https://doi.org/10.1002/prot.22458>.
53. Abrescia, N.G.A., Bamford, D.H., Grimes, J.M., and Stuart, D.I. (2012). Structure unifies the viral universe. *Annu. Rev. Biochem.* *81*, 795–822. <https://doi.org/10.1146/annurev-biochem-060910-095130>.
54. Mifsud, J.C.O., Lytras, S., Oliver, M.R., Toon, K., Costa, V.A., Holmes, E.C., and Grove, J. (2024). Mapping glycoprotein structure reveals Flaviviridae evolutionary history. *Nature* *633*, 695–703. <https://doi.org/10.1038/s41586-024-07899-8>.
55. Li, L., Jose, J., Xiang, Y., Kuhn, R.J., and Rossmann, M.G. (2010). Structural changes of envelope proteins during alphavirus fusion. *Nature* *468*, 705–708. <https://doi.org/10.1038/nature09546>.
56. Geoghegan, J.L., and Holmes, E.C. (2018). The phylogenomics of evolving virus virulence. *Nat. Rev. Genet.* *19*, 756–769. <https://doi.org/10.1038/s41576-018-0055-5>.
57. Garver, K.A., Batts, W.N., and Kurath, G. (2006). Virulence Comparisons of Infectious Hematopoietic Necrosis Virus U and M Genogroups in Sockeye Salmon and Rainbow Trout. *J. Aquat. Anim. Health* *18*, 232–243. <https://doi.org/10.1577/H05-038.1>.
58. Penaranda, M.M., Purcell, M.K., and Kurath, G. (2009). Differential virulence mechanisms of infectious hematopoietic necrosis virus in rainbow trout (*Oncorhynchus mykiss*) include host entry and virus replication

- kinetics. *J. Gen. Virol.* 90, 2172–2182. <https://doi.org/10.1099/vir.0.012286-0>.
59. Baillon, L., Mérour, E., Cabon, J., Louboutin, L., Vigouroux, E., Alencar, A.L.F., Cuenca, A., Blanchard, Y., Olesen, N.J., Panzarin, V., et al. (2020). The Viral Hemorrhagic Septicemia Virus (VHSV) Markers of Virulence in Rainbow Trout (*Oncorhynchus mykiss*). *Front. Microbiol.* 11. <https://doi.org/10.3389/fmicb.2020.574231>.
 60. Panzarin, V., Cuenca, A., Gastaldelli, M., Alencar, A.L.F., Pascoli, F., Morin, T., Blanchard, Y., Cabon, J., Louboutin, L., Ryder, D., et al. (2020). VHSV Single Amino Acid Polymorphisms (SAPs) Associated With Virulence in Rainbow Trout. *Front. Microbiol.* 11, 1984. <https://doi.org/10.3389/fmicb.2020.01984>.
 61. Gallagher, M.D., Karlsen, M., Petterson, E., Haugland, Ø., Matejusova, I., and Macqueen, D.J. (2020). Genome Sequencing of SAV3 Reveals Repeated Seeding Events of Viral Strains in Norwegian Aquaculture. *Front. Microbiol.* 11, 740. <https://doi.org/10.3389/fmicb.2020.00740>.
 62. Braaen, S., Wessel, Ø., Bjørgen, H., and Rimstad, E. (2025). Long-Term Protection in Atlantic Salmon (*Salmo salar*) to Pancreas Disease (PD) Can Be Achieved Through Immunization with Genetically Modified, Live Attenuated Salmonid Alphavirus 3. *Vaccines (Basel)* 13, 190. <https://doi.org/10.3390/vaccines13020190>.
 63. Holmes, A.C., Basore, K., Fremont, D.H., and Diamond, M.S. (2020). A molecular understanding of alphavirus entry. *PLoS Pathog.* 16, e1008876. <https://doi.org/10.1371/journal.ppat.1008876>.
 64. Jaimes, J.A., Andre, N.M., Chappie, J.S., Millet, J.K., and Whittaker, G.R. (2020). Phylogenetic Analysis and Structural Modeling of SARS-CoV-2 Spike Protein Reveals an Evolutionary Distinct and Proteolytically Sensitive Activation Loop. *J. Mol. Biol.* 432, 3309–3325. <https://doi.org/10.1016/j.jmb.2020.04.009>.
 65. Zehr, J.D., Kosakovsky, S.L., Millet, J.K., Olarte-Castillo, X.A., Lucaci, A.G., Shank, S.D., Ceres, K.M., Choi, A., Whittaker, G.R., Goodman, L.B., et al. (2023). Natural selection differences detected in key protein domains between non-pathogenic and pathogenic feline coronavirus phenotypes. *Virus Evol.* 9, vead019. <https://doi.org/10.1093/ve/vead019>.
 66. Duckert, P., Brunak, S., and Blom, N. (2004). Prediction of proprotein convertase cleavage sites. *Protein Eng. Des. Sel.* 17, 107–112. <https://doi.org/10.1093/protein/gzh013>.
 67. Tian, S., Huajun, W., and Wu, J. (2012). Computational prediction of furin cleavage sites by a hybrid method and understanding mechanism underlying diseases. *Sci. Rep.* 2, 261. <https://doi.org/10.1038/srep00261>.
 68. Gregoret, L.M., and Sauer, R.T. (1998). Tolerance of a protein helix to multiple alanine and valine substitutions. *Fold. Des.* 3, 119–126. [https://doi.org/10.1016/S1359-0278\(98\)00017-0](https://doi.org/10.1016/S1359-0278(98)00017-0).
 69. Buel, G.R., and Walters, K.J. (2022). Can AlphaFold2 predict the impact of missense mutations on structure? *Nat. Struct. Mol. Biol.* 29, 1–2. <https://doi.org/10.1038/s41594-021-00714-2>.
 70. Sun, J., Zhu, T., Cui, Y., and Wu, B. (2025). Structure-based self-supervised learning enables ultrafast protein stability prediction upon mutation. *Innovation* 6, 100750. <https://doi.org/10.1016/j.xinn.2024.100750>.
 71. Ripple, W.J., Wolf, C., Gregg, J.W., Rockström, J., Mann, M.E., Oreskes, N., Lenton, T.M., Rahmstorf, S., Newsome, T.M., Xu, C., et al. (2024). The 2024 state of the climate report: Perilous times on planet Earth. *Bioscience* 74, 812–824. <https://doi.org/10.1093/biosci/biae087>.
 72. Roberts, G.C., Zothner, C., Remenyi, R., Merits, A., Stonehouse, N.J., and Harris, M. (2017). Evaluation of a range of mammalian and mosquito cell lines for use in Chikungunya virus research. *Sci. Rep.* 7, 14641. <https://doi.org/10.1038/s41598-017-15269-w>.
 73. Kumar, S., Stecher, G., Li, M., Niyaz, C., and Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* 35, 1547–1549. <https://doi.org/10.1093/molbev/msy096>.
 74. Stecher, G., Tamura, K., and Kumar, S. (2020). Molecular Evolutionary Genetics Analysis (MEGA) for macOS. *Mol. Biol. Evol.* 37, 1237–1239. <https://doi.org/10.1093/molbev/msz312>.
 75. Jones, D.T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* 292, 195–202. <https://doi.org/10.1006/jmbi.1999.3091>.
 76. Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H., and Ferrin, T.E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* 30, 70–82. <https://doi.org/10.1002/pro.3943>.
 77. Meng, E.C., Pettersen, E.F., Couch, G.S., Huang, C.C., and Ferrin, T.E. (2006). Tools for integrated sequence-structure analysis with UCSF Chimera. *BMC Bioinf.* 7, 339. <https://doi.org/10.1186/1471-2105-7-339>.
 78. Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004). WebLogo: a sequence logo generator. *Genome Res.* 14, 1188–1190. <https://doi.org/10.1101/gr.849004>.
 79. Lallias, D., Bernard, M., Ciobotaru, C., Dechamp, N., Labbé, L., Goardon, L., Le Calvez, J.M., Bideau, M., Fricot, A., Prézélin, A., et al. (2021). Sources of variation of DNA methylation in rainbow trout: combined effects of temperature and genetic background. *Epigenetics* 16, 1031–1052. <https://doi.org/10.1080/15592294.2020.1834924>.
 80. Katoh, K., Misawa, K., Kuma, K.I., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30, 3059–3066. <https://doi.org/10.1093/nar/gkf436>.
 81. Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. <https://doi.org/10.1093/molbev/mst010>.
 82. Le, S.Q., and Gascuel, O. (2008). An improved general amino acid replacement matrix. *Mol. Biol. Evol.* 25, 1307–1320. <https://doi.org/10.1093/molbev/msn067>.
 83. Tighe, A.J., Gallagher, M.D., Carlsson, J., Matejusova, I., Swords, F., Macqueen, D.J., and Ruane, N.M. (2020). Nanopore whole genome sequencing and partitioned phylogenetic analysis supports a new salmonid alphavirus genotype (SAV7). *Dis. Aquat. Organ.* 142, 203–211. <https://doi.org/10.3354/dao03546>.
 84. Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. <https://doi.org/10.1093/nar/gkh340>.
 85. Croll, T.I. (2018). ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr. D Struct. Biol.* 74, 519–530. <https://doi.org/10.1107/S2059798318002425>.
 86. Sourimant, J., Rameix-Welti, M.A., Gaillard, A.L., Chevret, D., Galloux, M., Gault, E., and Eléouët, J.F. (2015). Fine mapping and characterization of the L-polymerase-binding domain of the respiratory syncytial virus phosphoprotein. *J. Virol.* 89, 4421–4433. <https://doi.org/10.1128/JVI.03619-14>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse monoclonal antibody (mAb) anti-SAV E2	The antibody was generated in our laboratory (équipe Virologie Moléculaire des Poissons, unité de Virologie et Immunologie Moléculaires, INRAE) as reported in Moriette et al. ²⁷	17H23
Mouse monoclonal antibody (mAb) anti-SAV E1	The antibody was generated in our laboratory (équipe Virologie Moléculaire des Poissons, unité de Virologie et Immunologie Moléculaires, INRAE) as reported in Moriette et al. ²⁷	78K5
Mouse monoclonal antibody (mAb) anti-GAPDH	Enzo Life Sciences	1D4
AlexaFluor 488-conjugated goat anti-mouse IgG (H + L)	Invitrogen Life Technologies	Cat. No. A11001
Peroxidase-conjugated goat anti-mouse IgG (H + L)	Sera Care	Cat. No. 5450-011
Bacterial and virus strains		
Salmonid Alphavirus subtype 2 (SAV2) formerly named Sleeping Disease Virus (SDV) S49P strain	The virus was first described in our laboratory (équipe Virologie Moléculaire des Poissons, unité de Virologie et Immunologie Moléculaires, INRAE) as reported in Villoing et al. ⁵	SAV2 (SDV) S49P
Recombinant SAV2 (rSAV2) with variant E2 N-term (rSAV2 E2) AAV	This study.	rSAV2 E2 AAV
Recombinant SAV2 (rSAV2) with variant E2 N-term (rSAV2 E2) VAV	Mérour et al. ¹¹	rSAV2 E2 VAV
Recombinant SAV2 (rSAV2) with variant E2 N-term (rSAV2 E2) VAA	Mérour et al. ¹¹	rSAV2 E2 VAA
Recombinant SAV2 (rSAV2) with variant E2 N-term (rSAV2 E2) AAA	This study.	rSAV2 E2 AAA
Recombinant SAV2 (rSAV2) with variant E2 N-term (rSAV2 E2) T1-AAA	This study.	rSAV2 E2 T1-AAA
Recombinant SAV2 (rSAV2) with variant E2 N-term (rSAV2 E2) AAD	This study.	rSAV2 E2 AAD
Recombinant SAV2 (rSAV2) with variant E2 N-term (rSAV2 E2) AAF	This study.	rSAV2 E2 AAF
Biological samples		
Cell lysates of SAV2- or rSAV2-infected BF-2 cells for Western blot analyses	This study.	Please refer to manuscript text for details.
Cell supernatants of rSAV2 infected BF-2 cells for viral growth kinetics	This study.	Please refer to manuscript text for details.
Head kidney and spleen organs of rSAV2-infected rainbow trout (AP2 line).	This study.	Please refer to manuscript text for details.
Chemicals, peptides, and recombinant proteins		
Protease Inhibitor Cocktail (PIC)	Roche/Merck	Cat. No. 04693159001
Pierce Enhanced Chemiluminescence (ECL) Western Blotting substrate	ThermoScientific	Cat. No. 32106
Tricaine	Sigma-Aldrich/Merck	MS-222

(Continued on next page)

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
Hoechst nuclei stain	Sigma-Aldrich/Merck	33342
Critical commercial assays		
QuickChange Multi Site-Directed Mutagenesis Kit	Agilent Technologies	Cat. No. 200514-5
Amaya Human T cell Nucleofector Kit	Lonza	T solution. Cat. No. VCA-1002
QIAamp kit	Qiagen	Cat. No. 52904
RNeasy Mini Kit	Qiagen	Cat. No. 74136
SuperScript IV Reverse Transcriptase	ThermoFisher	Cat. No. 3016673
Phusion High Fidelity DNA Polymerase	Thermo Scientific	F530S
Deposited data		
Models of E1, E2 and CP of Venezuelan Equine Encephalitis Virus TC-83 strain	Protein DataBank (PDB)	3j0c
Cryo-EM structure of Chikungunya virus asymmetric unit	Protein DataBank (PDB)	8fcg
Crystal structure of the immature envelope glycoprotein complex of Chikungunya virus	Protein DataBank (PDB)	3n40
EEEV glycoproteins bound with heparan sulfate	Protein DataBank (PDB)	6odf
Cryo-EM structure of an alphavirus, Sindbis virus	Protein DataBank (PDB)	6imm
Cryo-EM Structure of Western Equine Encephalitis Virus	Protein DataBank (PDB)	8dec
Ross river virus (STRAIN T48)	Protein DataBank (PDB)	6vyv
Cryo-EM structure of the mature and infective Mayaro virus	Protein DataBank (PDB)	7ko8
Electron cryomicroscopy of Chikungunya virus	Protein DataBank (PDB)	3j2w
Model of SAV virion based on E2-E1 heterodimer and capsid (Cp) proteins	This study. Model deposited in ModelArchive https://www.modelarchive.org	ma-gkhkz
AHPV E2-E1-Cp heterotrimer model	This study. Model deposited in ModelArchive https://www.modelarchive.org	ma-xgzne
SESV E2-E1-Cp heterotrimer model	This study. Model deposited in ModelArchive https://www.modelarchive.org	ma-zwpvs
WHAV E2-E1-Cp heterotrimer model	This study. Model deposited in ModelArchive https://www.modelarchive.org	ma-md6n0
CAV E2-E1-Cp heterotrimer model	This study. Model deposited in ModelArchive https://www.modelarchive.org	ma-zvp1t
WCFV E2-E1-Cp heterotrimer model	This study. Model deposited in ModelArchive https://www.modelarchive.org	ma-6kpi5
WSFAV E2-E1-Cp heterotrimer model	This study. Model deposited in ModelArchive https://www.modelarchive.org	ma-jt78w
Experimental models: Cell lines		
Bluegill fry (<i>Lepomis macrochirus</i>) fibroblast-like cells	ATCC	CCL-91

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental models: Organisms/strains		
Rainbow trout (<i>Oncorhynchus mykiss</i>), synthetic line	INRAE	Synthetic (Sy)
Rainbow trout (<i>Oncorhynchus mykiss</i>), AP2 isogenic line	INRAE	AP2
Oligonucleotides		
For detailed information on all primers used in this study, please see Table S3 .	Not applicable (N.A.)	Not applicable (N.A.)
Recombinant DNA		
pSAV2-patho plasmid (originally named pSDV-Struct _{wr})	The plasmid was generated in our laboratory (équipe Virologie Moléculaire des Poissons, unité de Virologie et Immunologie Moléculaires, INRAE) as described in Méroux et al. ¹¹	Not applicable (N.A.)
Software and algorithms		
BLAST Search	NCBI	blastn (nucleotides) and blastp (proteins) https://blast.ncbi.nlm.nih.gov/Blast.cgi
Geneious Prime	Biomatters	2025.0.3
MEGAX	Kumar et al., 2018 ⁷³ and Stecher et al. ⁷⁴	10.1.8
AlphaFold	DeepMind AlphaFold Server. Abramson et al. ¹⁸ and Jumper et al. ¹⁹	Version 3 https://robetta.bakerlab.org
RoseTTAFold	Baek et al. ²¹	Robetta https://robetta.bakerlab.org
FoldSeek Protein Structure Search	Gilchrist et al. ²⁹	Accessible on FoldSeek Search Server https://search.foldseek.com/search
FoldMason Multiple Protein Structure Alignment	Gilchrist et al. ²⁹	Accessible on FoldSeek Search Server https://search.foldseek.com/search
PsiPred	Jones ⁷⁵	4.0 https://bioinf.cs.ucl.ac.uk/psipred
ChimeraX	Pettersen et al. ⁷⁶ and Meng et al. ⁷⁷	1.9
Weblogo	Crooks et al. ⁷⁸	3.7.12
Pythia	Sun et al. ⁷⁰	https://pythiastudio.wulab.xyz/tools/pythia
ProP	Duckert et al. ⁶⁶	1.0
Prism	GraphPad	7
ImageJ	Open Source image processing software. See list of contributors: https://imagej.net/people/	2.0.0-rc-43/1.52n
Zen	Zeiss	2.3

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Animals used in this study

Two lines of juvenile rainbow trout (*Oncorhynchus mykiss*) were used for the infection trials with rSAV2 variants: the INRAE synthetic line (average weight 2.64 g) and the AP2 isogenic line (average weight 2.95 g).^{37,79} Both lines were developed at INRAE and maintained at the Infectiologie Expérimentale des Rongeurs et des Poissons unit (IERP, Rodent and Fish Experimentation unit, part of the EMERG'IN national research infrastructure). The fish were kept in tanks maintained at 10°C. Due to the young age of the juveniles used in the experimentations, it is not possible to report on the influence (or association) of sex, gender, or both on the results of the study. We acknowledge that the lack of reporting constitutes a limitation to our research's generalizability.

Ethics statement

All *in vivo* experimentations involving rainbow trout strictly adhered to the European guidelines and recommendations on animal experimentation and welfare (European Union Directive 2010/63). All experimental procedures involving rainbow trout were approved by the local ethics committee on animal experimentation (Comité d'éthique appliqué à l'expérimentation animale pour le center de recherche INRAE de Jouy-en-Josas; COMETHEA INRAE no. 45) and were authorized by the Ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche (APAFIS authorization no. 29801–2021021110262075 v2). The fish

facilities in which experiments were carried out were also approved (authorization no. D-78-322-720). To minimize animal suffering and distress, all manipulations were carried out under light anesthesia performed by bath immersion with tricaine (0.005%, Sigma-Aldrich MS-222). During the entire duration of experiments, fish were monitored twice daily for clinical signs and survival. Upon display of typical infection symptoms, animals were humanely euthanized by bath immersion using a lethal dose of tricaine (0.015%, Sigma-Aldrich MS-222).

Virus strain

The subtype 2 (SAV2) strain used in this study was S49P. It was formerly named Sleeping Disease Virus (SDV) and was initially described in our laboratory (équipe Virologie Moléculaire des Poissons, UMR Virologie et Immunologie Moléculaires, INRAE).⁵ The virus was grown and titered at 10°C using the BF-2 cell line. This viral strain was used for SAV2 reverse genetics, as detailed below.

Cell line

Bluegill fry-2 fibroblast-like cells (BF-2, ATCC CCL-91) were derived from Bluegill fry (*Lepomis macrochirus*) and were maintained at 14°C in Glasgow Minimum Essential Medium (GMEM, Pan Biotech) supplemented with 23.75 mM Tris-HCl pH 7.6, 4 mM NaHCO₃, 2 mM L-glutamine (Eurobio), 10% fetal bovine serum (Capricorn Scientific), 1% tryptose phosphate broth (Gibco ThermoFisher) and penicillin-streptomycin antibiotics, respectively at 100 U/mL and 100 µg/mL (Biovalley). The cell lines have not been authenticated and were not tested for mycoplasma contamination.

METHOD DETAILS

Phylogenetic analyses

Alphavirus and SAV nucleotide and protein sequences were retrieved from NCBI GenBank, Blast searches using SAV2 reference sequence (GenBank: [AJ316246.1](#)), and from published studies (Tables S1 and S2;^{31–33}).

For the terrestrial and aquatic alphavirus phylogenetic analysis, the amino acid sequences of 17 representative alphavirus structural proteins were used (Table S1). The sequences encoding the capsid protein were removed to improve alignment reliability, as this region contains numerous indels and diverges substantially.² The sequences were aligned using MAFFT (v7.388) algorithm^{80,81} within Geneious software (Biomatters). A part of the protein alignment corresponding to the N-term region of E2 envelope protein was used to highlight conserved and divergent features among aquatic and terrestrial alphaviruses, notably an insert only present in fish alphaviruses and the unique 7-8-9 triplet motif of SAV. Phylogenetic analyses were performed in MEGAX 10.1.8.^{73,74} The choice of the evolutionary model was determined from 56 models based on the Bayesian information criterion (BIC) and a Maximum-Likelihood (ML) tree was generated from the alignment using LG (G + I) substitution model.⁸² Phylogeny testing was undertaken by the bootstrap method with 1000 replicates. The phylogenetic tree was drawn to scale in MEGAX and formatted using FigTree 1.4.4 with branch lengths measured in the number of substitutions per site. Numbers at nodes indicate bootstrap support.

For the SAV phylogenetic analysis, nucleotide sequences corresponding to an internal 357-bp region of E2 used for SAV subtype demarcation were retrieved from 90 representative SAV isolates (Table S2).³⁴ When available, the host species from which the viral sequence was sampled was noted and included in the tree. All major SAV subtypes (1–6) are represented in the analysis. Verifications performed on a SAV7 sequence (isolate F/58/17⁸³) showed that it contained numerous nonsense mutations rendering the encoding proteins non-functional and was not included in the analysis. The nucleotide sequences were aligned (codon-aware) in MEGAX using MUSCLE.⁸⁴ A Maximum-Likelihood (ML) tree was generated from the alignment using the GTR (G + I) substitution model. Phylogeny testing was performed using the bootstrap method with 1000 replicates. The phylogenetic tree was drawn to scale in MEGAX and formatted using FigTree 1.4.4 with branch lengths measured in the number of substitutions per site. Numbers at nodes indicate bootstrap support.

A protein sequence alignment of the E2-encoding region of the above 90 SAV isolates was generated in Geneious Prime 2025.0.3 (Biomatters) using MAFFT (v7.388) algorithm. This protein alignment was used for subsequent analyses, including determining the amino acid composition of the 7-8-9 triplet for each isolate analyzed in the SAV phylogenetic tree, the sequence logo analysis (WebLogo) of E2 N-term, and mapping of conservation on predicted E2 structure (AL2CO entropy measure) (see below).

AlphaFold-based fish alphavirus protein structure predictions

The predicted structures of fish alphavirus E1, E2, and capsid (Cp) proteins were modeled using AlphaFold 3 (<https://alphafoldserver.com>).^{18,19} For capsids, only the C-term portion of the proteins were modeled to match with the structurally ordered domain of experimentally determined alphavirus structures (e.g., VEEV PDB: [3j0c](#)).⁹

For SAV, the predictions were based on the reference SAV2 protein sequences (GenBank: [AJ316246.1](#)). Since alphavirus E1 and E2 envelope proteins form a closely associated heterodimeric complex and as AlphaFold 3 can natively generate multimeric models, SAV E1 and E2 proteins were modeled together as a heterodimer, while the C-term half of the capsid (aa 122–283) was modeled separately. The overall prediction scores for both the E2-E1 heterodimer and capsid were high with pTM scores of 0.76 and 0.87 respectively (Table 1). However, for E1, AlphaFold 3 gave poor pLDDT scores for a stretch of residues in the alpha helical transmembrane (TM) domain (aa ₄₂₉RIVGNPSGPVSSS₄₄₁, average C_αpLDDT of 51, Table 1) introducing a kink and break in the α-helix which are

not present in any of the experimentally determined structures of alphavirus E1 proteins (e.g., VEEV PDB: [3j0c](#) and CHIKV PDB: [8fcg](#)), indicative of a relatively high degree of prediction error. To improve the predicted model of SAV E1, this stretch of amino acids was replaced with that of the SAV2 E1 sequence that was previously cloned in our laboratory (aa₄₂₉GIVGTLVVLFLI₄₄₀) and found in other SAV strains (e.g., GenBank: [NC_003930.1](#)).³⁶ By doing so, AlphaFold 3 was able to predict a TM domain with a linear α -helix with improved pLDDT values (average $C\alpha_{pLDDT}$ of 72 for the internal TM region, [Table 1](#)) and matching experimentally determined alphavirus E1 structures ([Figure S2](#)). This improved predicted model of SAV E1 was subsequently used throughout this study.

To confirm the predicted N-term α -helix predicted by AlphaFold 3 in SAV E2, the protein sequence of SAV E2 was used to generate a full-length protein model by RoseTTAFold, another leading structure-prediction algorithm (<https://robotta.bakerlab.org>).²¹ In addition, the N-term sequence of E2 was analyzed using PsiPred (<https://bioinf.cs.ucl.ac.uk/psipred/>),⁷⁵ a PSI-BLAST-based secondary structure prediction algorithm to predict secondary structures.

For the analysis of E2E3 furin cleavage, the E3 sequence (GenBank: [AJ316246.1](#)) was modeled using AlphaFold 3 together with E2 and the above-mentioned E1 (with modified TM domain). Modeling of E3 was performed either as the uncleaved form (p62, E3E2), or as the cleaved form with E3 non-covalently associated with E2.

For the other aquatic alphaviruses infecting mammalian or fish hosts, the NCBI source accession numbers of protein sequences used for modeling the E2-E1 heterodimer and capsid (C-term structurally ordered portion) were the following: Southern elephant seal virus (SESV) GenBank: [AEJ36233.1](#) sampled from *Mirounga leonine*; Alaskan harbor porpoise alphavirus (AHPV) GenBank: [QJE50388.1](#), sampled from *Phocoena phocoena*; Comber alphavirus (CAV) GenBank: [MN207265.1](#), sampled from *Serranus cabrilla*; Wenling crested flounder alphavirus (WCFAV) GenBank: [MG600127.1](#), sampled from *Plagiopsetta* sp.; Wenling hagfish alphavirus (WHAV) GenBank: [MG600128.1](#), sampled from *Eptatretus burgeri*; and Wenling striated frogfish alphavirus (WSFAV) GenBank: [MG600126.1](#), sampled from *Antennarius striatus*. To generate structural models with AlphaFold 3, E1, E2, and Cp were modeled as a heterotrimer. For capsids, the amino acid boundaries used were based on protein alignments and boundaries used for SAV Cp: SESV aa 108–268; AHPV aa 76–236; CAV aa 142–303; WCFAV aa 136–296; WHAV aa 167–327; and WSFAV aa 136–296. The overall pTM scores obtained for the AlphaFold 3-predicted heterotrimeric complexes were the following: SESV 0.74; AHPV 0.72; CAV 0.68; WCFAV 0.69; WHAV 0.66; and WSFAV 0.68.

The experimentally determined structures of VEEV and CHIKV were retrieved from PDB (PDB: [3j0c](#) and [8fcg](#)). Comparative analysis between uncleaved SAV and CHIKV p62 proteins were performed by structurally aligning the uncleaved SAV p62 model with the crystal structure of CHIKV p62 (PDB: [3n40](#)). A chimeric CHIKV p62 protein containing SAV E2 N-term α -helix was modeled by AlphaFold 3 by replacing CHIKV residues 6–13 with residues 6–22 of SAV E2 N-term. Molecular visualizations and structural alignments (MatchMaker) were performed using UCSF ChimeraX 1.9.⁷⁶

In silico reconstruction of a whole SAV virion

To reconstruct a 3D model of an entire SAV virion, the predicted structures for E1, E2, and the capsid proteins generated by AlphaFold 3 were mapped onto the icosahedral lattice of Venezuelan equine encephalitis virus (VEEV), a reference alphavirus structure which has been experimentally determined by cryo-EM (PDB: [3j0c](#), $T = 4$ icosahedral symmetry).⁹ The TM domains of SAV E1 and E2 envelope proteins, which are connected to the ectodomains by a flexible linker, were structurally aligned to obtain a better alignment fit with the corresponding regions of VEEV E1 and E2 proteins using ChimeraX.^{76,77} Further refinements to detect and avoid clashes within and between SAV E2-E1-Cp subunits were performed using Isolde.⁸⁵ The models of SAV E1, E2, and Cp were then fitted by structural alignment (MatchMaker tool within ChimeraX^{76,77}) onto the asymmetric unit of VEEV virion, which is composed of 4 copies of E2-E1-Cp heterotrimer (PDB: [3j0c](#)). To apply the symmetry operations generating a complete icosahedral virion from a single SAV asymmetric subunit (240 copies of E1, E2, and Cp proteins), the following list categories of the VEEV PDB: [3j0c](#) structure file were copied in the SAV asymmetric unit mmCIF file: `_pdbx_struct_assembly`, `_pdbx_struct_assembly_gen` and `_pdbx_struct_oper`. All molecular visualizations were performed using UCSF ChimeraX 1.9.⁷⁶

Structure-based phylogenetic alignment and trees

FoldMason was used to generate structure-informed phylogenetic trees, using the AlphaFold 3 predicted structures of the ectodomains of E1 and E2, and the C-term portion of Cp of aquatic alphaviruses infecting mammals or fish hosts (<https://search.foldseek.com/foldmason>).²⁹ FoldMason is a fast and accurate multiple structure aligner (MSTA) based on Foldseek's 3Di alphabet.³⁰ For each of the predicted protein structures used as input, the following aa boundaries (defined by protein sequence alignments with terrestrial alphaviruses) were used for improved structural alignment: SAV E1_{ecto} aa 1–397, E2_{ecto} aa 1–352, Cp_{C-term} aa 134–283; SESV E1_{ecto} aa 1–377, E2_{ecto} aa 1–334, Cp_{C-term} aa 121–268; AHPV E1_{ecto} aa 1–377, E2_{ecto} aa 1–336, Cp_{C-term} aa 89–236; CAV E1_{ecto} aa 1–397, E2_{ecto} aa 1–352, Cp_{C-term} aa 155–303; WCFAV E1_{ecto} aa 1–396, E2_{ecto} aa 1–354, Cp_{C-term} aa 148–296; WHAV E1_{ecto} aa 1–397, E2_{ecto} aa 1–350, Cp_{C-term} aa 179–327; and WSFAV E1_{ecto} aa 1–396, E2_{ecto} aa 1–354, Cp_{C-term} aa 148–296. For the terrestrial alphaviruses VEEV and CHIKV (PDB: [3j0c](#) and [8fcg](#)), the following aa boundaries were used: VEEV E1_{ecto} aa 1–380, E2_{ecto} aa 1–342, Cp_{C-term} aa 127–275; and CHIKV E1_{ecto} aa 1–379, E2_{ecto} aa 5–342, Cp_{C-term} aa 111–261. FoldMason generates a guide tree, a graphical representation of the structural alignment, as well as a structure-based alignment (MSTA). For each protein (E1_{ecto}, E2_{ecto}, and Cp_{C-term}), the MSTA was used to generate a Maximum-Likelihood (ML) phylogenetic tree. Phylogenetic analyses were performed in MEGAX 10.1.8.^{73,74} For each protein, the choice of the evolutionary model was determined from 56 models based on the Bayesian information criterion (BIC) and a Maximum-Likelihood (ML) tree was generated from the alignment using WAG+G+I substitution model for the

E1_{ecto} tree and WAG+G substitution model for the E2_{ecto} and Cp_{C-term} trees.⁸² Phylogeny testing was undertaken by the bootstrap method with 1000 replicates. The phylogenetic tree was drawn to scale in MEGAX and formatted using FigTree 1.4.4 with branch lengths measured in the number of substitutions per site. Numbers at nodes indicate bootstrap support. Equivalent sequence-based trees were generated for comparison (see Figure S4).

Visualization of SAV E2 sequence conservation

To obtain a representation of the sequence conservation at the N-term of SAV E2, a sequence logo of the protein was generated using Weblogo 3.7.12 (<https://weblogo.threeplusone.com/create.cgi>).⁷⁸ A protein alignment focused on the first 20 amino acids of SAV E2 was derived from the full-length protein alignment of the 90 SAV E2 sequences previously generated (see above). This alignment was used as input for the Weblogo server. The total height of stacked symbols at each position indicates the degree of conservation at that position and the height of each symbol within a given position indicates the relative frequency of each amino acid.

To map conservation onto the structure of the SAV E2 ectodomain (aa 1–352) predicted by AlphaFold 3, the ALCO2 entropy measure program was implemented in ChimeraX 1.9.^{35,76} Based on the previously obtained protein alignment of 90 SAV E2 sequences (see above), AL2CO calculates a conservation index for each residue of the alignment, which can be mapped and color-coded onto each position of SAV E2 ectodomain.³⁵

Predicted impact of N-term variations on SAV E2 protein stability

Calculation of the predicted effect of substitutions on protein stability was performed using Pythia (<https://pythiastudio.wulab.xyz/tools/pythia>), a robust deep-learning model able to rapidly and accurately predict the effect of mutations on protein stability using protein structural information as input.⁷⁰ The input structure used is the E2 ectodomain (aa 1–352) of the reference SAV (GenBank: AJ316246.1, with A₁ and γ VAV₉ triplet 7–8–9 motif). Pythia generates a heatmap and numeric table covering all positions of a given protein and tests at each position the effect of substitutions with each of the 20 amino acids. For each position and residue substitution, Pythia attributes a score proportional to $\Delta\Delta G$, which corresponds to the difference in the Gibbs free energy of unfolding (ΔG) between two proteins (non-substituted vs. substituted). A lower score indicates that the substitution is more likely to stabilize the protein, while a higher score indicates that the substitution is more likely to destabilize the protein.

Plasmid constructs and site-directed mutagenesis

The plasmids pSAV2-patho (pathogenic, originally named pSDVStruct_{wt}¹¹) and pJET1.2-*BsrGI*_patho were generated in our laboratory and were used here to recover recombinant SAV2 (rSAV2) with variant E2 N-term amino acid composition.^{11,14,36} pSAV2-patho is an infectious plasmid construct containing full-length SAV2 cDNA (subtype 2) and the plasmid pJET1.2-*BsrGI*_patho consists of a *BsrGI*-digested DNA fragment encompassing the complete E2 coding sequence and cloned into the pJET1.2 vector. Both pSAV2-patho and pJET1.2-*BsrGI*_patho contain the E2 substitutions found in the virulent strain of SAV2, SDV S49P (Figure S1). Sanger sequencing (Eurofins Scientific) with SEQ13 and pJET1.2For primers, respectively, demonstrates that pSAV2-patho and pJET1.2-*BsrGI*_patho plasmids contain an E2 sequence with VAA triplet 7–8–9 (Figure S1, Table S3). The first reverse genetics plasmid pSAV2 (originally pSDV) contains the VAV triplet 7–8–9 (Figure S1). Targeted mutagenesis of SAV E2 N-term was accomplished using QuickChange Site-Directed Mutagenesis Kit (Agilent Technologies) according to the manufacturer's guidelines with the pJET1.2-*BsrGI*_patho as template and the mutagenesis primers SDM64–68 (Table S3) to generate the following E2 variants: AAV, AAF, AAD, AAA and T1-AAA. To generate an internal 6-residue E2 N-term deletion mutant (E2 Δ_{4-9}), a PCR amplification-based strategy using a 5'-phosphorylated internal primer pair (Table S3) was used as described by Sourimant et al.⁸⁶ Targeted mutations and the deletion were verified by Sanger sequencing (Eurofins Scientific) using the sequencing primer SEQ13 (Table S3). Each of the mutated *BsrGI* DNA fragments were cloned back into the *BsrGI*-digested pSAV2-patho backbone.

Recombinant SAV2 (rSAV2) recovery

For recovery of rSAV2 with variant E2 N-term, BF-2 cells were trypsinized (Difco Trypsin, BD Biosciences), counted and seeded at a density of 5×10^6 cells/well in six-well plates and transfected by electroporation (Amaxa, Lonza) with 2 μ g of each of the mutated pSAV2 constructs to generate rSAV2 E2 AAV, AAF, AAD, AAA, and T1-AAA variants using the T solution and the T-020 electroporation program. Cells were incubated at 20°C overnight in GMEM medium supplemented with 10% FBS. 24 h post-transfection, the supernatants were replaced with serum-free GMEM medium and the cells were incubated at 10°C for 6–10 days rSAV2s with variant E2 N-term were amplified through two successive passages on BF-2 cells seeded in 24-well plates. Viruses were titrated by immunofluorescence assays as described below.

Verification of the E2 N-term sequences of SAV2 (SDV) S49P, rSAV2 and rSAV2-patho (previously named rSAVStruct_{wt}¹¹) were carried out by performing viral RNA extractions from cell culture supernatants using QIAamp kit (Qiagen). Viral RNAs were retrotranscribed using SuperScript IV reverse transcriptase (Invitrogen) and the *BsrGI*-Rev primer, then PCR-amplified (E2 N-term region) followed by Sanger sequencing (Eurofins Scientific) using *BsrGI*-For primer (Figure S1, Table S3). To streamline recombinant virus naming and following sequence verification results, rSAV2 and rSAV2-patho were renamed rSAV2 E2 VAV, and rSAV2 E2 VAA respectively.

Western blot analyses of SAV E3E2 furin cleavage

For analysis of SAV E3E2 furin cleavage, 2.5×10^6 BF-2 cells were seeded in six-well plates. The cells were either mock-infected or infected with SAV2 (SDV, S49P) at an m.o.i. of 1 or 5 and incubated at 15°C. 72 h after infection the cells were lysed in lysis buffer (150 mM NaCl, 50 mM Tris-HCl pH 8.5, and 0.5% NP-40) with the addition of protease inhibitor cocktail (Roche). For virus purification, wild-type SAV2 was mass produced in BF-2 cells, clarified by low-speed centrifugation at 4,000 rpm for 10 min at 4°C, then concentrated 10-fold by ultracentrifugation at 25,000 rpm in an SW28 Beckman rotor for 90 min at 4°C and finally purified by ultracentrifugation at 36,000 rpm in an SW41 Beckman rotor for 4 h at 4°C on a 25% (w/v) sucrose cushion in TEN buffer (10 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA pH 8). The viral pellet was then resuspended in TEN buffer. For analysis of p62 cleavage of rSAV2 E2 variants, infection of BF-2 cells was performed as previously described but using 6.25×10^5 BF-2 cells seeded in 24-well plates and infected the next day at an m.o.i. of 0.1 for 7 days. Aliquots of cell lysates or sucrose-cushion purified virions were incubated with 5× cracking buffer composed of 42% glycerol, 21% β-mercaptoethanol, 250 mM Tris, 12.5% SDS, and bromophenol blue and heated at 100°C for 5 min. The samples were then separated by SDS-PAGE gel electrophoresis on a 4 to 12% gradient polyacrylamide Bis-Tris gel (NuPage, Invitrogen). After transfer on a polyvinylidene difluoride (PVDF) membrane (Immobilon-P; Millipore), separated proteins were detected using either a mouse monoclonal antibody (mAb 17H23) against SAV E2 diluted to 1:8000, or a mAb against SAV E1 (78K5) diluted to 1:1000,²⁷ or against the cellular housekeeping protein GAPDH diluted to 1:2000 (loading control, Enzo Life Sciences). Immunolabeled proteins were visualized with Peroxidase-conjugated goat anti-mouse antibodies using an enhanced chemiluminescence (ECL) detection system (Pierce), and image acquisitions of Western blot signals were performed using a ChemiDoc Imaging System (Bio-Rad) or Amersham ImageQuant 800 Western Blot Imaging System (Cytiva). For analysis of p62 proteolytic cleavage in rSAV2 infected lysates, densitometry analyses were performed on p62 (uncleaved) and E2 (cleaved) bands using ImageJ software (Measure tool, average gray value), with background subtraction. For each rSAV2 E2 variant, the ratio between cleaved E2 and uncleaved p62 was calculated by dividing the net inverted pixel density value for each E2 band with the corresponding value for the p62 band.

Virus titration and growth kinetics

BF-2 cells were infected with rSAVs at an m.o.i. of 0.1. Supernatants from rSAV2-infected cells were collected at 0, 2-, 4-, 7-, 10-, and 14-d.p.i. 10-fold dilutions of supernatants were used to infect fresh BF-2 cells seeded in 96-well plates. Viral titrations were performed in duplicate experiments. At 7 d.p.i., virus titers were determined by immunofluorescence assays and counting of fluorescent foci as described below (fluorescence focus assay).

Immunofluorescence and fluorescence focus assays

rSAV2-infected BF-2 cells were fixed at 7 d.p.i. with a cold solution of ethanol:acetone (1:1, v/v) for 20 min at -20°C and then air dried at room temperature. Immunofluorescence assays were performed using a mAb directed against SAV E2 protein (17H23) previously obtained in our laboratory.²⁷ Fixed BF-2 cells were incubated for 45 min at room temperature with primary mouse mAb 17H23 diluted 1:10000 in Tween 0.05% (PBS-T). The cells were subsequently washed three times in PBS-T. The cells were then incubated for 45 min at room temperature with Alexa Fluor (AF) 488-conjugated goat anti-mouse antibody (Invitrogen ThermoFisher) diluted 1:800 in PBS-T. The cells were then washed three times in PBS-T. For fluorescence focus assays, positively infected cell foci were visualized and counted under an Eclipse TE 200 inverted fluorescence microscope (Nikon). Viral titers were calculated taking into account the dilution factors and are expressed as fluorescent focus units per mL (FFU/mL). Immunofluorescence assays were also undertaken at 7- and 10-d.p.i. for microscopy image acquisitions as detailed above with the addition of Hoechst for labeling of cell nuclei in the secondary antibody solution. The immunolabeled cells were imaged using an Axio Observer.Z1 inverted fluorescence microscope with a 10× objective (Zeiss) equipped with a CoolSNAP HQ² CCD camera (Photometrics Roper Scientific). Microscopy image acquisitions were performed using Zen 2.3 software (Zeiss).

In vivo analyses of rSAV2 infection in rainbow trout

For each rainbow trout line, groups of 25 virus-free juveniles housed in individual tanks were infected by intraperitoneal injection (i.p.) with 10^4 FFU/fish of rSAV2 E2 AAV, VAV, VAA and T1-AAA or were mock infected (serum-free GMEM medium). The rSAV2 with AAF, AAD, and AAA triplet motif were not included in the infection trials because of the low virus titers obtained after recovery (Table S4). The fish were kept in separate groups in 30 L tanks at 10°C. Mortalities were recorded daily over a period of 63 days.

To verify rSAV2 sequence stability during the course of the infection trial, the SAV E2 N-term sequences of the rSAV2 variants isolated from infected AP2 rainbow trout were analyzed. Sampling of head kidney and spleen was performed on fish that had succumbed to infection for each group. The organs (from 14 fishes) were collected, kept individually (not pooled), and homogenized using ceramic beads and the Precellys 24 Touch tissue homogenizer (Bertin Technologies). RNA was extracted using the RNeasy Mini Kit (Qiagen) according to manufacturer's guidelines. cDNA synthesis and RT-PCR were performed using SuperScript IV reverse transcriptase (Invitrogen) and the SEQ172 and SEQ173 primer pair allowing to amplify a region encompassing the E2 N-term region (Table S3). The PCR-amplified region was sequenced by Sanger sequencing (Eurofins Scientific) using each of the above-mentioned specific primers.

QUANTIFICATION AND STATISTICAL ANALYSES

Experimental data were analyzed and plotted using GraphPad Prism (version 7). For the *in vivo* rSAV2 infection trials, cumulative mortalities were plotted as Kaplan-Meier survival curves. All of the of the statistical details of the infection trial experiments can be found in the legends of [Figure 7](#) and below. For each experimental group, the number of animals was $n = 25$. Statistical groupings were inferred by performing a Log rank Mantel-Cox test on the Kaplan-Meier survival data. The following convention was used for describing statistical significance: not significant (n.s.) $p > 0.05$; significant (*) $p \leq 0.05$; highly significant (**), $p \leq 0.01$.