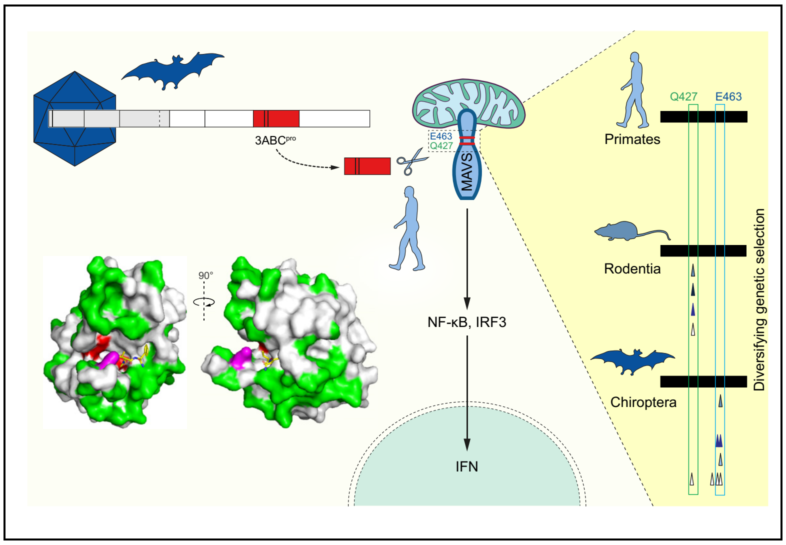
Research Article
Viral Hepatitis

JOURNAL OF HEPATOLOGY

Hepatovirus 3ABC proteases and evolution of mitochondrial antiviral signaling protein (MAVS)

Graphical abstract



Highlights

- 3ABC proteases expressed by bat hepatoviruses cleave human MAVS at Glu^{463}/Gly^{464}.
- Bat hepatovirus 3ABC proteases disrupt viral activation of interferon-β in human cells.
- MAVS orthologs from rodents and bats resist cleavage by cognate hepatovirus 3ABC proteases.
- Bat and rodent MAVS orthologs have been subject to diversifying selection at 3ABC cleavage sites.
- Cross-species disruption of MAVS signaling by 3ABC may facilitate hepatovirus host species shifts.

Authors

Hui Feng, Anna-Lena Sander,
Andrés Moreira-Soto, Daisuke Yamane,
Jan Felix Drexler, Stanley M. Lemon

Correspondence

smlemon@med.unc.edu
(S.M. Lemon)

Lay summary

Hepatitis A virus, a common cause of acute hepatitis globally, is likely to have evolved from a virus that jumped from a rodent species to humans within the last 10–12 thousand years. Here we show that distantly related hepatoviruses, that infect bats and rodents today, express proteases that disrupt innate antiviral responses in human cells. This conserved attribute of hepatoviruses may have contributed to that ancient host species shift.

http://dx.doi.org/10.1016/j.jhep.2019.02.020
J. Hepatol. 2019, 71, 25–34

Research Article
Viral Hepatitis

JOURNAL
OF HEPATOLOGY

Hepatovirus 3ABC proteases and evolution of mitochondrial antiviral signaling protein (MAVS)

Hui Feng[1,2], Anna-Lena Sander[3], Andrés Moreira-Soto[3], Daisuke Yamane[1,2], Jan Felix Drexler[3,4], Stanley M. Lemon[1,2,5,*]

[1]Lineberger Comprehensive Cancer Center, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA; [2]Department of Medicine, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA; [3]Charité-Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Institute of Virology, 10117 Berlin, Germany; [4]German Centre for Infection Research (Deutsches Zentrum für Infektionsforschung), Berlin, Germany; [5]Department of Microbiology & Immunology, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Background & Aims: Unlike other hepatitis viruses that have infected primates for millions of years, hepatitis A virus (HAV) likely entered human populations only 10–12 thousand years ago after jumping from a rodent host. The phylogeny of modern hepatoviruses that infect rodents and bats suggest that multiple similar host shifts have occurred in the past. The factors determining such shifts are unknown, but the capacity to overcome innate antiviral responses in a foreign species is likely key.

Methods: We assessed the capacity of diverse hepatovirus 3ABC proteases to cleave mitochondrial antiviral signaling protein (MAVS) and disrupt antiviral signaling in HEK293 and human hepatocyte-derived cell lines. We also applied maximum-likelihood and Bayesian algorithms to identify sites of diversifying selection in MAVS orthologs from 75 chiropteran, rodent and primate species.

Results: 3ABC proteases from bat, but not rodent hepatoviruses efficiently cleaved human MAVS at $\text{Glu}^{463}/\text{Gly}^{464}$, disrupting virus activation of the interferon-β promoter, whereas human HAV 3ABC cleaved at $\text{Gln}^{427}/\text{Val}^{428}$. In contrast, MAVS orthologs from rodents and bats were resistant to cleavage by 3ABC proteases of cognate hepatoviruses and in several cases human HAV. A search for diversifying selection among MAVS orthologs from all 3 orders revealed 90 of ~540 residues to be under positive selection, including residues in chiropteran MAVS that align with the site of cleavage of human MAVS by bat 3ABC proteases.

Conclusions: 3ABC protease cleavage of MAVS is a conserved attribute of hepatoviruses, acting broadly across different mammalian species and associated with evidence of diversifying selection at cleavage sites in rodent and bat MAVS orthologs. The capacity of hepatoviruses to disrupt MAVS-mediated innate immune responses has shaped evolution of both hepatoviruses and their hosts, and facilitates cross-species transmission of hepatitis A.

Lay summary: Hepatitis A virus, a common cause of acute hepatitis globally, is likely to have evolved from a virus that jumped from a rodent species to humans within the last 10–12 thousand years. Here we show that distantly related hepatoviruses, that infect bats and rodents today, express proteases that disrupt innate antiviral responses in human cells. This conserved attribute of hepatoviruses may have contributed to that ancient host species shift.

Introduction

Despite effective vaccines, hepatitis A virus (HAV) remains a common cause of acute viral hepatitis in many regions of the world.[1] The continued presence of this hepatotropic virus is dependent upon unbroken chains of fecal-oral transmission, since persistent infection is rare-to-nonexistent and chronic HAV shedders are unknown. Accordingly, HAV disappears at times from small, isolated, human populations, only to return with a vengeance when re-introduced.[2,3] Thus, unlike hepatitis B virus (HBV) and hepatitis C virus (HCV) that cause long-term persistant infections and co-evolved with humans and other primate species over millions of years,[4–6] HAV likely became established among human populations only when groups living together became large enough to sustain chains of transmission 10–12 thousand years ago. Multiple, distantly related hepatoviruses have recently been discovered among bats, rodents and other small mammals.[7,8] These viruses are hepatotropic, and although distantly related phylogenetically, share antigenic determinants with human HAV. Their phylogeny provides evidence for multiple past host species shifts across different mammalian orders, and ancestral reconstructions suggest a rodent origin for human HAV.[7]

Given their proclivity for *de novo* generation of diversity through error-prone genome replication, RNA viruses such as HAV are disproportionately involved in host species shifts.[9,10] The determinants of such host shifts are not fully understood, but innate immune surveillance represents a substantial biological barrier to overcome in a new host. HAV is a stealthy pathogen in humans and chimpanzees, inducing little innate immune response in the liver.[11] This is due in part to cleavage of mitochondrial antiviral signaling protein (MAVS), a critical innate

---

Keywords: hepatitis A virus; Chymotrypsin-like protease; Innate immunity; Virus evolution; Cross-species transmission.

Received 24 November 2018; received in revised form 31 January 2019; accepted 19 February 2019; available online 13 March 2019

* Corresponding author. Address: 8.034 Burnett-Womack Building CB#7292, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7292, USA. Tel.: +1 919 843 1848; fax: +1 919 843 7240.

E-mail address: smlemon@med.unc.edu (S.M. Lemon).

immune adaptor protein, by a precursor of the mature HAV $3C^{pro}$ protease, 3ABC.[12] Although the natural host range of HAV appears limited to primates,[13] the virus robustly infects *Mavs*$^{-/-}$ knockout mice.[14] Unlike the human MAVS protein (*hs*MAVS), murine MAVS (*mm*MAVS) is not cleaved by the 3ABC protease. These data suggest that HAV must overcome MAVS-mediated innate immunity to achieve cross-species transmission, and led us to ask whether the 3ABC proteases of contemporary bat and rodent viruses are capable of cleaving *hs*MAVS, and whether this host-pathogen interaction has shaped the past evolution of hepatoviruses and their hosts.

Materials and methods

Cells

HEK293FT cells and human hepatocyte-derived cell lines including Huh-7.5,[15] PH5CH8,[16] and CRISPR/Cas9-edited MAVS-deficient PH5CH8 cells (PH5CH8-KO)[17] were mycoplasma-free. All cells were cultured in DMEM with 10% FBS (Gibco) at 37 °C.

Virus stocks

The cell culture-adapted HM175/18f strain of human HAV[18] was propagated in Huh-7.5 cells for 7 days following electroporation of *in vitro*-transcribed viral RNA (GenBank accession number KP879216.1). Several cell culture-adaptive mutations exist within the 3ABC sequence of HM175/18f virus, but they are not known to influence 3ABC cleavage of MAVS. Lentiviruses were prepared as described previously.[17] Sendai virus (SeV, Cantell strain) was purchased from Charles River Laboratory.

Virus infection and transduction

Cultures were infected with HAV (10^6 genome equivalents/culture) or SeV (600 HA units/ml) at 37 °C for 4 h and 6 h, respectively. Transduction was facilitated by the addition of 8 μg/ml polybrene, and the resulting cells were subjected to selection with blasticidin (6 μg/ml).

Quantification of HAV genomic RNA by RT-qPCR

Total RNA was extracted with the RNeasy Mini Kit (Qiagen). First-strand cDNA was synthesized with Superscript III reverse transcriptase (Invitrogen). HAV cDNA was quantified by PCR with Universal SYBR supermix using cDNA from *in vitro*-transcribed HAV RNA as a standard (Bio-Rad). Primers are listed in Table S1.

Plasmids

Hepatovirus 3ABC protease sequences were cloned into pCMV-(N)HA, whereas pcDNA3-(N)FLAG[12] and pcDNA3-(N)FLAG/eGFP were used to express MAVS sequences. Lentivector pLOCΔGFP[17] was used to stably express HA-tagged 3ABC and GFP-fused MAVS proteins. All constructs were tagged N-terminally, with cloning carried out using a PCR-based strategy[19] that generated a non-relevant 2.4 kDa carboxy-terminal tail. Products were verified by DNA sequencing. Oligonucleotide primers are listed in Table S1.

Transfection

HEK293FT and Huh-7.5 cells preseeded in 12-well plates were transfected with vectors expressing GFP-*hs*MAVS (100 ng) and either empty vector (EV) or non-primate hepatoviral 3ABC proteases (HA-3ABC, 100 ng). GFP-fused non-primate MAVS vectors (500 ng) were similarly transfected with hepatovirus 3ABC proteases (250 and 500 ng). MAVS-knockout PH5CH8-KO cells preseeded in 12-well plates were transfected with vectors expressing GFP-fused mammalian MAVS vectors (500 ng). The interferon-β (*IFNB1*) promoter-driven firefly luciferase plasmid (40 ng IFNβ-luc) and the *Renilla* luciferase plasmid (4 ng pRL-TK) were co-transfected with 3ABC (40 ng) and/or MAVS constructs (40 ng) into naïve or PH5CH8-KO cells preseeded in 96-well plates.

Immunoblots

Twenty-four (HEK293FT cells) and 48 h (Huh-7.5 and PH5CH8 cells) after transfection, cells were lysed and immunoblotting carried out as described.[17] Detailed information regarding the primary antibodies used can be found in the supplementary CTAT table. All secondary antibodies (LI-COR Biosciences; Cat# 926-32211, 926-32212, 926-68020 and 926-68073) were used at a 1:12,000 dilution.

Dual luciferase reporter assay

Luciferase activity was determined 30 h after transfection using the Luciferase Assay System (Promega). To compare the impact of expressing catalytically active *vs.* inactive proteases, we normalized *IFNB1*-driven firefly luciferase activity to that of an internal control, *Renilla* luciferase (expressed from a co-transfected vector), and expressed the result as relative luciferase activity (Fig. 3D and 3F). In experiments comparing SeV *vs.* mock virus challenge, we similarly normalized luciferase activities but report it as fold-increase due to the virus infection (Fig. 1C). In Fig. 3A, where we show that overexpression of MAVS generates an innate immune response, we reported raw firefly luciferase light units together with immunoblots showing the expression of IFIT1 (an interferon-stimulated gene) and GAPDH as a control for cell mass.

Phylogeny analysis

Bayesian phylogeny was generated at the amino acid level from translation alignments excluding all ambiguous data or gaps using MrBayes V3.1[20] and a WAG substitution model. Trees were run for 2 million generations, sampled every 100 steps. After excluding 5,000 of the total 20,000 trees as burn-in, final trees were annotated with TreeAnnotator from the BEAST package[21] and visualized with FigTree. Genome annotations and translation alignments were done using Geneious 6.1.8 (https://www.geneious.com).

Evolutionary analyses

Pressure analyses were performed using translation alignments of all available MAVS sequences from the 3 different orders done in Genious 6.1.8. The CodeML program implemented in the pamlX 1.3.1 software package[22,23] was used to investigate positive selection pressure in different MAVS datasets. Codon substitution models M0, M1a, M2a, M3, M7 and M8 were tested using the codon frequency model F61. Likelihood ratio tests (LRTs) on 2 pairs of site-specific models, M1a *vs.* M2a and M7 *vs.* M8, were performed to determine whether MAVS genes from each of the 3 orders are evolving under positive selection using the chi-square calculator implemented in the PAMLX package, a user interface for phylogenetic analysis by maximum likelihood (PAML). Fixed effect likelihood (FEL), single likelihood ancestor counting (SLAC), and random effect likelihood (REL),[24] and single bayesian approximation (FUBAR)[25] from the HyPhy package

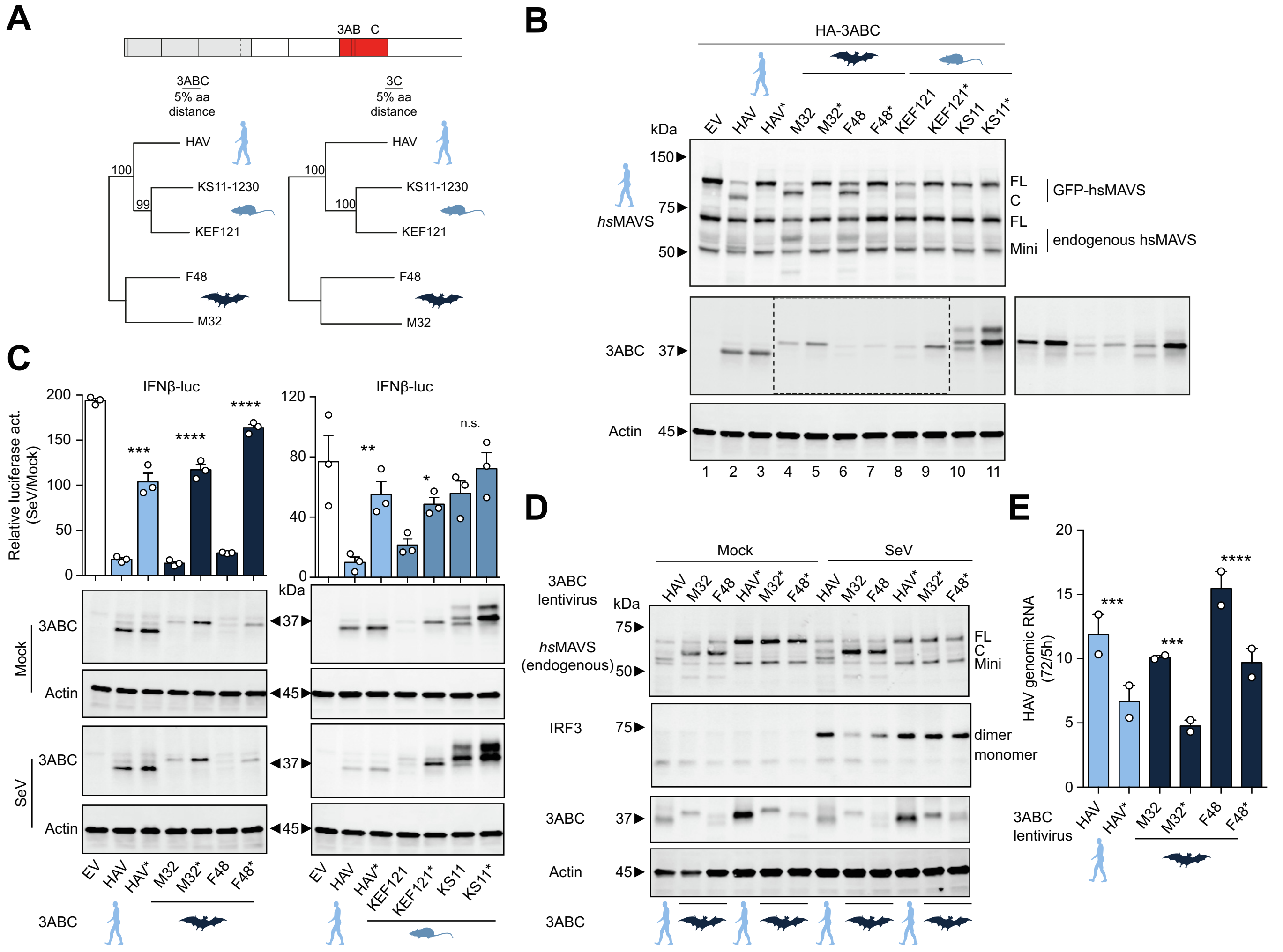


Fig. 1. Bat hepatovirus 3ABC proteases cleave *hs*MAVS. (A) Hepatovirus 3ABC and $3C^{pro}$ protease phylogeny. Nodal values are percentage support for grouping from 1,000 bootstrap replicates. Scale bar indicates genetic distance. (B) Immunoblots of HEK293FT cells co-expressing *hs*MAVS N-terminally fused to GFP and rodent or bat hepatovirus 3ABC proteases N-terminally tagged with HA. Actin included as a loading control. The 3ABC panel shows the noncleaved 3ABC proteases detected with anti-HA. The indicated section of the central panel is shown with high sensitivity detection on the right. (C) SeV activation of the *IFNB1* promoter in PH5CH8 cells expressing rodent or bat hepatovirus 3ABC proteases. (D) Immunoblots of endogenous MAVS and IRF3 dimerization in PH5CH8 cells stably expressing bat virus 3ABC proteases, with and without SeV stimulation. (E) Relative HAV RNA abundance in infected PH5CH8 cells expressing bat hepatovirus 3ABC proteases. Bars show mean ± SEM from 2–3 technical replicates (separate cultures in 1 representative experiment) and were analyzed by *t* test. $^{***}p<0.001$, $^{****}p<0.0001$. All western blots were representative of at least 3 independent experiments. *Catalytically inactive 3ABC proteases. C, cleavage product; EV, empty vector; FL, full length GFP-tagged MAVS and endogenous MAVS; HAV, hepatitis A virus; *hs*MAVS, human MAVS; Mini, a truncated MAVS variant identified previously;[45] SeV, Sendai virus.

(Datamonkey.org) were also employed to detect sites under selection pressure using the HKY85 substitution mode, as suggested by automatic model selection in Datamonkey. The branch-site unrestricted statistical test for episodic diversification (BUSTED) was used to detect gene-wide positive selection in the MAVS datasets.

Amino acid sequence distances were calculated with MEGA7[26] using the pairwise deletion option and all species per order for which MAVS sequences were available. MAVS pairwise distances of species were normalized by mean cytochrome b pairwise distance per order. Sequence distances were plotted using SSE V1.2.[27] Representative viruses from each host order were tagged and sequence identities within families were plotted using a fragment length of 200 and a step size of 100 amino acid residues. Alignment gaps were excluded from the analysis.

Data availability

GenBank accession numbers for hepatovirus sequences are listed in Table S2. GenBank accession numbers for MAVS sequences are listed in Table S3.

Statistical analysis

Comparisons were carried out by 2-sided *t* test using Prism 6.0 (GraphPad Software, Inc.).

Results

3ABC proteases from bat hepatoviruses cleave *hs*MAVS

The 3A polypeptides of human, rodent and bat hepatoviruses all contain a transmembrane domain like that in human HAV 3A that anchors the 3ABC protease to mitochondria and is essential for MAVS cleavage.[7,12] However, the 3C^{pro} sequences themselves share only ~50% amino acid identity (Fig. 1A). We were surprised therefore to find that N-terminally HA-tagged 3ABC proteases from 2 bat hepatoviruses, M32 (*Eidolon helvum*) or FO1A-F48 ('F48', *Rhinolophus landeri*), cleaved *hs*MAVS fused at its amino terminus to GFP and co-expressed in HEK293 cells (Fig. 1B). Replacing the active-site Cys residues in these proteases with Ala (M32* and F48*, respectively) abolished *hs*MAVS cleavage (Fig. 1B), confirming that the 3ABC cysteine protease was responsible. By contrast, 3ABC proteases from viruses recovered from the grey dwarf hamster (*Cricetulus migratorius*) KS11-1230 ('KS11') and Jaliscan cotton rat (*Sigmodon mascotensis*) KEF121, either failed to cleave *hs*MAVS or did so less efficiently (Fig. 1B), although these rodent virus proteases are phylogenetically closer to human HAV 3ABC than the bat hepatovirus proteases (Fig. 1A). Ablating the active-site Cys residues in the rodent proteases (KS11* and KEF121*, respectively) enhanced their levels of expression (Fig. 1B), consistent with the elimination of autocatalysis at the internal 3A/B and 3B/C cleavage sites and providing evidence that the proteases were in fact functionally active.

Overexpression of either bat virus protease in PH5CH8 cells, which are derived from normal adult human hepatocytes and have a robust MAVS signaling pathway, disrupted SeV-induced activation of the *IFNB1* promoter, whereas the catalytically inactive mutants did not (Fig. 1C). Importantly, stable lentivirus-based expression of the bat virus proteases in PH5CH8 cells resulted in cleavage of endogenous MAVS and impaired SeV-triggered dimerization of the transcription factor, interferon regulatory factor 3 (IRF3) (Fig. 1D). Consistent with these data, HAV replication was significantly enhanced in lentivirus-transduced PH5CH8 cells expressing the wild-type *vs.* mutant bat proteases, despite the fact that the human viral 3ABC protease was also expressed by the HAV replicating in these cells (Fig. 1E). We conclude from these data that MAVS cleavage is a conserved attribute of hepatoviruses found in bats, and that proteases expressed by these viruses are capable of disrupting innate immune responses to virus infection in human cells.

Additional experiments revealed that the M32 and F48 proteases preferentially cleave *hs*MAVS at Glu^{463}/Gly^{464}, rather than Gln^{427}/Val^{428}, where *hs*MAVS is targeted by the human virus[12] (Fig. 2A). Both sites in *hs*MAVS are consistent with the dipeptide substrate specificities of human HAV 3C^{pro}. Previous peptide cleavage studies show Gln is strongly preferred by HAV 3C^{pro} over Glu at the p1 position (the first residue on the amino side of the site of scission), favoring the Gln^{427}/Val^{428} site[28] (Fig. 2B). However, peptides with Gly at the p'1 position (as in the Glu^{463}/Gly^{464} site) are cleaved significantly more efficiently than those with p'1 Val. Overall, the preference for Glu^{463}/Gly^{464} demonstrated by the bat 3ABC protease indicates a subtle difference in substrate specificities, and is consistent with differences in the amino acid side chains abutting the substrate cleft in the crystallographic structure of the protease[29] (Fig. 2C).

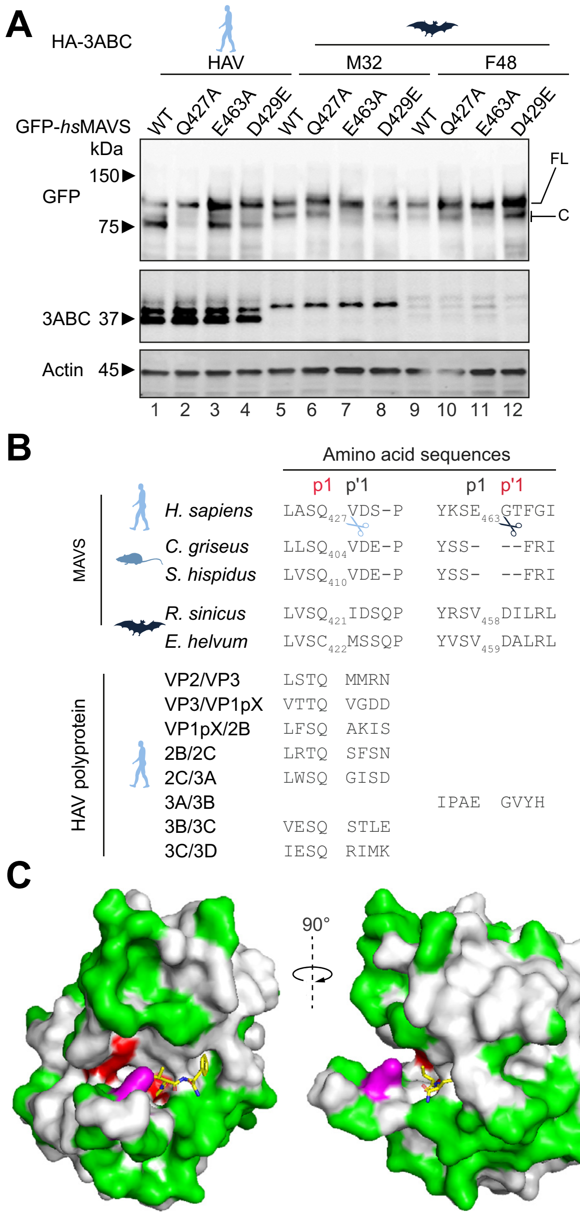


Fig. 2. Substrate specificities of bat and human hepatoviral 3ABC proteases. (A) Bat hepatovirus 3ABC cleavage of *hs*MAVS mutants with amino acid substitutions at p1 positions of possible cleavage sites. The D429E mutation ablates caspase cleavage in apoptotic cells.[46] See legend to Fig. 1B. (B) Alignments of human, rodent and bat MAVS amino acid sequences at 3ABC cleavage sites in *hs*MAVS. Cleavage sites targeted by 3C^{pro} in the HAV polyprotein (NC001489) are shown at the bottom, and suggest a strong preference for Gln over Glu at p1 and a preference for Gly over Val at p'1. (C) Surface-rendered image of the 3D crystallographic structure of human HAV 3C^{pro} protease (1QA7) with residues that differ in the *E. helvum* M32 viral protease colored green or magenta (Lys^{147}, which abuts the substrate cleft near the S3 specificity pocket). Active-site His^{44} and Cys^{172} residues are colored red. HAV, hepatitis A virus; *hs*MAVS, human MAVS; WT, wild-type *hs*MAVS.

Chiropteran MAVS proteins are resistant to cleavage by human or bat 3ABC proteases

We next asked whether MAVS proteins expressed by small mammals harboring hepatoviruses are equally subject to cleavage by human or bat viral proteases. Since the MAVS sequence is not known for *R. landeri*, the horseshoe bat from which F48 virus was recovered, we expressed MAVS from a related species, *R. sinicus* (*rs*MAVS). Similarly, we expressed MAVS from the hispid cotton rat (*S. hispidus*, *sh*MAVS) and the Chinese hamster (*C. griseus*, *cg*MAVS) in lieu of *S. mascotensis* and *C. migratorius*, which harbor KEF121 and KS11 viruses respectively. The *IFNB1* promoter was strongly activated by expressing GFP fusions of these MAVS orthologs, as well as MAVS from the M32 virus

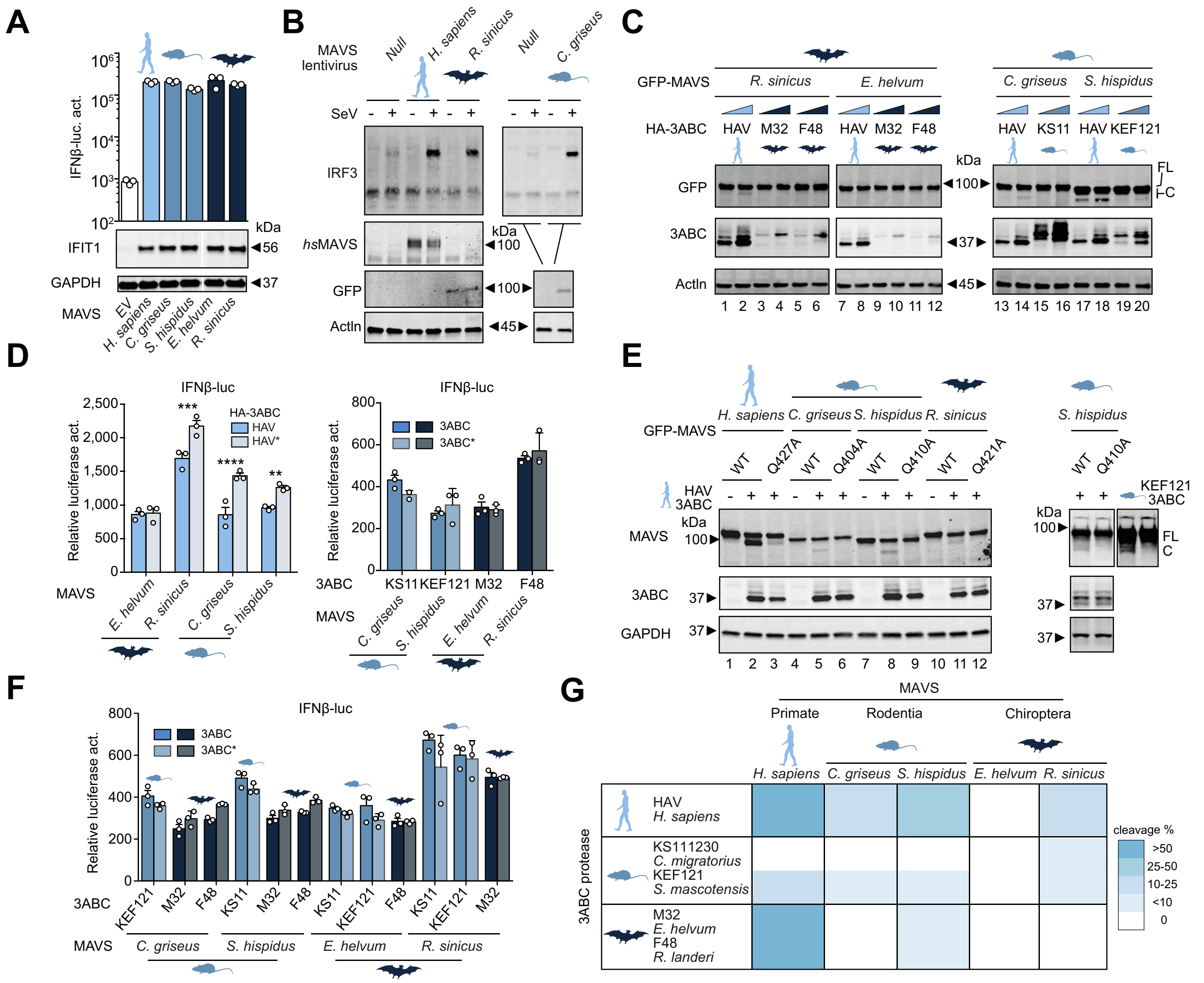


Fig. 3. Cleavage of mammalian MAVS proteins by hepatovirus 3ABC proteases. (A) *IFNB1* promoter activity and immunoblots of IFIT-1 and GAPDH (loading control) in PH5CH8-KO cells expressing MAVS othologs. (B) SeV-triggered IRF3 dimerization in PH5CH8-KO cells reconstituted with bat and rodent MAVS orthologs. (C) Immunoblots of N-terminally GFP-tagged MAVS in HEK293FT cells co-expressed with HA-tagged 3ABC proteases from human, bat or rodent hepatoviruses. See legend to Fig. 1B. (D, F) Mean *IFNB1* promoter activity ± SEM in PH5CH8-KO cells co-expressing rodent or bat MAVS and wild-type or catalytically inactive (*) 3ABC proteases. $^{**}p$ <0.01, $^{***}p$ <0.001, $^{****}p$ <0.0001 by *t* test. (E) Immunoblots of HEK293FT cells co-expressing human, rodent or bat GFP-MAVS, with or without Ala substitutions at Gln residues aligning with Gln^{427} in *hs*MAVS, and human HA-3ABC proteases from HAV and the rodent KEF121 virus. (G) Summary of mammalian MAVS cleavage by 3ABC proteases. The heat map was generated from quantitative analyses of immunoblots (see Fig. 3C). HAV, hepatitis A virus; hsMAVS, human MAVS; PH5CH8-KO, MAVS-knockout PH5CH8; SeV, Sendai virus.

host, *E. helvum* (*eh*MAVS), in PH5CH8 cells knocked out for endogeneous MAVS expression by CRISPR/Cas9 gene editing (Fig. 3A). The synthesis of IFIT1, an interferon-stimulated gene product, was also stimulated (Fig. 3A). Moreover, RIG-I signaling could be efficiently reconstituted in these cells by representative bat (*rs*MAVS) and rodent (*cg*MAVS) MAVS proteins, resulting in SeV-induced IRF3 dimerization (Fig. 3B). These data reveal a surprising degree of functional conservation among MAVS proteins from distantly related mammalian species, and indicate that these non-primate proteins were expressed in physiologically relevant compartments.

Under these conditions, we observed only minimal cleavage of *rs*MAVS, and no cleavage of *eh*MAVS by ectopically expressed human HAV 3ABC (Fig. 3C, lanes 1–2 and 7–8). Consistent with this, *IFNB1* promoter activation induced by expression of *rs*MAVS, but not *eh*MAVS, was reduced by co-expressing HAV 3ABC *vs.* the catalytically inactive mutant 3ABC* (Fig. 3D). Importantly, the p1 Gln residue is preserved in *rs*MAVS at the position aligning with Gln^{427} in hsMAVS (Gln^{421}), whereas it is replaced by Cys in *eh*MAVS (Fig. 2B). By contrast, there is no conservation of the p1 and p'1 residues aligning with Glu^{463}/Gly^{464} in either bat MAVS ortholog. Surprisingly, HAV 3ABC robustly cleaved *sh*MAVS (cotton rat), and less efficiently, *cg*MAVS (hamster) (Fig. 3C, lanes 13–14 and 17–18). These cleavages were functionally important, as co-expressing the active protease reduced *IFNB1* promoter activation induced by either rodent MAVS ortholog (Fig. 3D). In line with these results, the amino acid sequence around the Gln^{427}/Val^{428} cleavage site in *hs*MAVS is well preserved in alignments of the rodent orthologs, whereas the Glu^{463}/Gly^{464} site targeted by bat hepatovirus

3ABC proteases is ablated by deletion of 3 residues (Fig. 2B). Further experiments confirmed that the HAV 3ABC protease cleaves each of these MAVS orthologs at residues aligning with Gln^{427}/Val^{428} in *hs*MAVS (Fig. 3E).

Despite cleaving human *hs*MAVS (Fig. 1A), 3ABC proteases from the bat hepatoviruses, F48 (*R. landeri*) and M32 (*E. helvum*), failed to cleave the cognate MAVS proteins expressed by the bats (Fig. 3C, lanes 3–6 and 9–12). This is consistent with the Glu^{463}/Gly^{464} site targeted by these proteases in *hs*MAVS being poorly conserved in bat MAVS (Fig. 2B). Similarly, the cotton rat *sh*MAVS was only cleaved minimally, and the hamster *cg*MAVS was not cleaved at all, by the cognate rodent 3ABC proteases (Fig. 3C, lanes 19–20 and 15–16, respectively). As might be expected from these results, there was no evidence of cleavage of chiropteran MAVS orthologs by proteases from rodent viruses, or vice versa (data not shown). Consistent with these results, none of the 3ABC proteases from non-primate viruses was capable of significantly disrupting *IFNB1* promoter signaling transduced through either cognate or non-cognate rodent or bat MAVS orthologs (Fig. 3D, F).

Bat and rodent MAVS proteins have undergone diversifying selection at sites of 3ABC cleavage

The data presented above suggest that bat and rodent MAVS orthologs are relatively resistant to cleavage by their cognate 3ABC proteases, whereas the bat viral proteases retain the ability to cleave human *hs*MAVS (Fig. 3G). This led us to seek evidence for genetic diversifying selection in the MAVS proteins expressed by these mammalian orders that have likely hosted hepatoviruses over millions of years.[7] Two-thirds of ~5,400 extant mammalian species belong to the orders Rodentia and Chiroptera,[30] which together host the greatest diversity of hepatoviruses (Fig. 4A). Because MAVS sequences were available from only 11 bat and 19 rodent species, none known hosts of hepatoviruses, we determined MAVS coding sequences in 4 bat (*E. helvum*, *Hypsignathus monstrosus*, *Pipistrellus pipistrellus*, *R. ferrumequinum*) and 2 rodent (*Myodes glareolus* and *S. hispidus*) hepatovirus host species, providing a total of 75 MAVS sequences for analysis (Table S3). Although primates were overrepresented in this database, MAVS sequences from Chiroptera and Rodentia were significantly more diversified than those from primates (Fig. 4B). Surprisingly, MAVS diversity was not related to overall host genetic relationships. *hs*MAVS shared less identity with rodent than bat MAVS (Fig. 4C), although rodents and primates are genetically closer than bats and primates.[31] Sequence distances were unevenly distributed across the MAVS reading frame (Fig. 4D). These data suggest that MAVS from distinct mammalian orders in general, and some regions of MAVS in particular, may have been shaped by differential evolutionary forces.

The ratio (ω) of non-synonymous (*dN*) to synonymous (*dS*) substitution rates was less than unity across the entire MAVS protein in all 3 mammalian orders, regardless of the method used for calculation (Fig. 5A). However, significant episodic diversifying selection was evident in primate, rodent and chiropteran MAVS ($p < 0.001$) according to both LRTs (Table 1) and BUSTED.[32] Maximum-likelihood (FEL, SLAC and PAML) and Bayesian (REL and FUBAR) methods[24,25] identified 90 positively selected residues, mostly in bats and rodents (Fig. 5B, Table S4). Notably, 41 of these (45.6%) aligned in a contiguous fashion with other residues under positive selection. Fourteen of the 17 (87.5%) contiguous peptide segments thus identified had evidence of diversifying selection in at least 2 different orders, suggesting similar evolutionary pressures shaping MAVS across different orders.

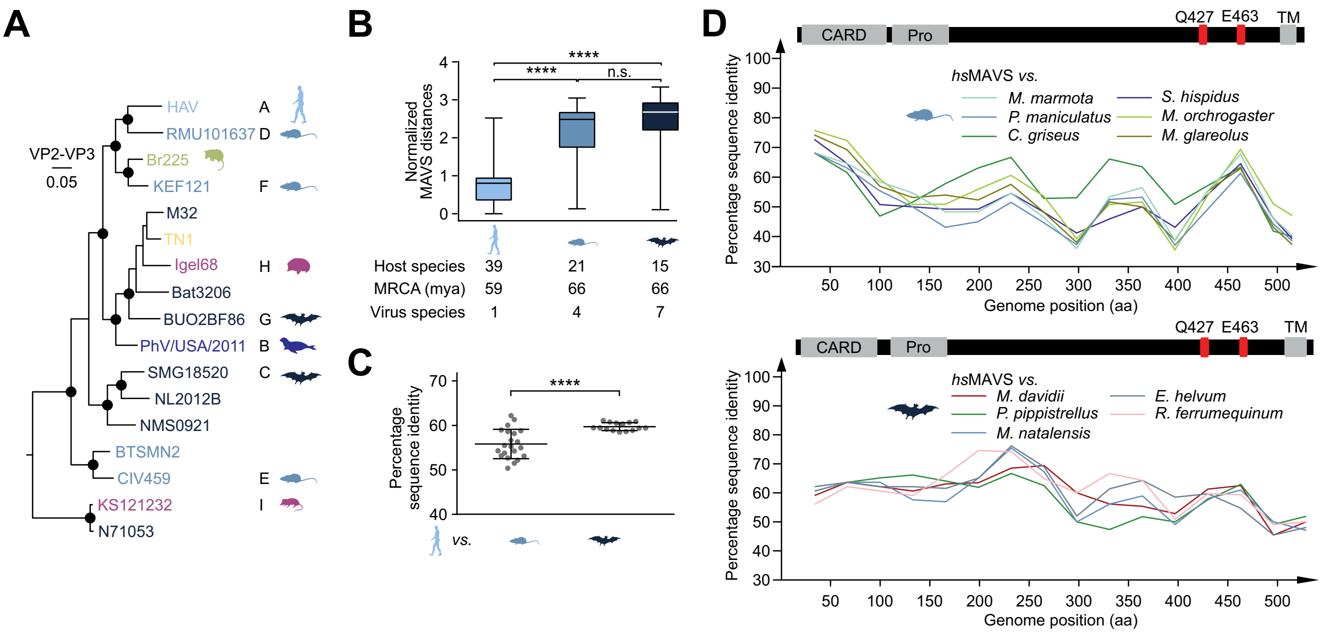


Fig. 4. Hepatovirus diversity and evidence for episodic diversifying selection in MAVS. (A) Partial VP2-VP3 hepatovirus phylogeny. Bayesian posterior probability support above 0.9 at nodes is highlighted by filled circles. Scale bar indicates genetic distance. Avian encephalomyelitis virus is included as an outgroup. (B) Cytochrome b-normalized MAVS mean pairwise distances per host order. (C) Mean pairwise sequence identities ± SD between *hs*MAVS and rodent or bat orthologs. ****p <0.0001 by t test. (D) Amino acid sequence identities along the length of *hs*MAVS with (top) rodent and (bottom) chiropteran MAVS orthlogs. CARD, CARD-domain; *hs*MAVS, *human* MAVS; MRCA (mya), most recent common ancestor of orders (millions of years ago); Pro, proline-rich region; TM, transmembrane domain.

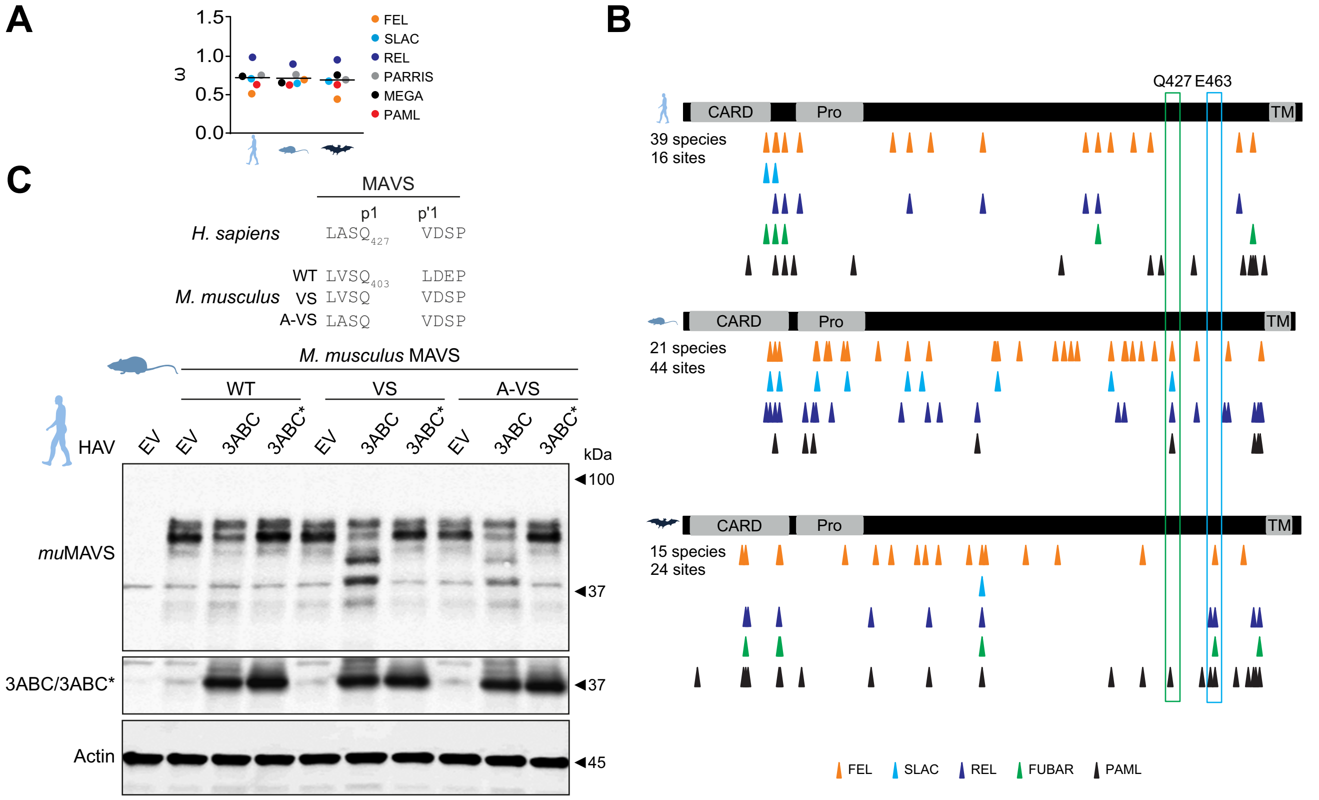


Fig. 5. Episodic diversifying selection in mammalian MAVS. (A) Average dN/dS (ω) ratios of primate, rodent and chiropteran MAVS sequences calculated using indicated methods. Horizontal line, mean dN/dS. (B) Amino acid residues identified under positive selection within primate, chiropteran and rodent MAVS. Triangles mark sites with signals for positive selection in at least 1 of the 5 methods. See Table S4 for details. (C) Immunoblots of Huh-7.5 cells co-expressing *mm*MAVS, with (VS, A-VS) or without (WT) substitutions at p3, p'1, and/or p'3 positions aligning with the site of 3ABC cleavage in *hs*MAVS, and HA-tagged 3ABC or catalytically inactive (*) 3ABC proteases from human HAV. dN/dS, non-synonymous/synonymous; HAV, hepatitis A virus; *hs*MAVS, human MAVS; *mm*MAVS, murine MAVS.

Table 1. Likelihood ratio tests for selected PAML site-specific models.

Selected site models	$\ln L_0$	$\ln L_1$	$2\Delta\ln L$	df	*p* value
Primates					
$M1a_{\text{(nearly neutral)}}$ *vs.* $M2a_{\text{(pos. selection)}}$	−9125.10	−9113.53	23.13	2	<0.0001
$M7_{\text{(beta)}}$ *vs.* $M8_{\text{(beta\&}\omega\text{)}}$	−9131.50	−9113.80	35.40	2	<0.0001
Rodentia					
$M1a_{\text{(nearly neutral)}}$ *vs.* $M2a_{\text{(pos. selection)}}$	−15457.26	−15448.28	17.95	2	<0.001
$M7_{\text{(beta)}}$ *vs.* $M8_{\text{(beta\&}\omega\text{)}}$	−15448.11	−15433.70	28.82	2	<0.0001
Chiroptera					
$M1a_{\text{(nearly neutral)}}$ *vs.* $M2a_{\text{(pos. selection)}}$	−8003.66	−7983.35	40.62	2	<0.0001
$M7_{\text{(beta)}}$ *vs.* $M8_{\text{(beta\&}\omega\text{)}}$	−8009.67	−7981.22	56.90	2	<0.0001

Evidence of positive selection was sought by comparing site models that allow ω >1 (M2a, M8) with models that disallow positive selection (M1a, M7). df, degree of freedom; lnL, log-likelihood scores; PAML, phylogenetic analysis by maximum likelihood.

Using any of the 5 analytic methods, we found no evidence for diversifying selection at the site targeted by the human HAV protease, $\text{Gln}^{427}/\text{Val}^{428}$, among MAVS proteins expressed by primates (Fig. 5B, Table S4). This is consistent with the relatively recent introduction of HAV into primate populations, as discussed above. In contrast, as reported previously,[4] PAML detected a signal for selection at residues aligning with $\text{Cys}^{508}/\text{His}^{509}$, where *hs*MAVS is cleaved by the HCV NS3/4A protease.[33,34] This is consistent with evidence suggesting the existence of hepaciviruses related to HCV among primates for millions of years.[4] We also noted evidence for diversifying selection at residues aligning with this site in MAVS orthologs expressed by rodents and bats (Fig. 5B, Table S4), which also host diverse hepaciviruses.[35,36]

Despite the absence of diversifying selection at the HAV 3ABC cleavage site, $\text{Gln}^{427}/\text{Val}^{428}$, among MAVS proteins expressed by primates, we detected a selection signal at residues aligning with this site in bats using PAML, and a strong signal for selection at this site in rodents using multiple tests for genetic pressure (Fig. 5B, Table S4). These results are consistent with ancestral reconstructions which suggest rodents hosted the immediate ancestor of human HAV.[7] Just as striking, however, was a very strong signal for diversifying selection among chiropteran MAVS othologs at residues aligning with Lys^{461} and Thr^{465} (Fig. 2B), which bracket the site of *hs*MAVS cleavage by the bat viral proteases (Fig. 5B, Table S4). These data are consistent with MAVS having evolved over millenia to a relatively 3ABC-resistant phenotype in these 2 mammalian orders under continued pressure from proteases expressed by ancient hepatoviruses.

Lastly, we asked whether it is possible to convert a 3ABC cleavage-resistant rodent MAVS protein to an active 3ABC substrate by limited substitutions of amino acids at a site aligning with cleavage in human *hs*MAVS. To answer this question, we co-expressed either wild-type or minimally modified murine (*Mus musculus*) *mm*MAVS and 3ABC from human HAV in 293 cells. Our previous work shows that *mm*MAVS is not susceptible to 3ABC cleavage, and that normal mice cannot be infected with HAV, whereas *Mavs*$^{-/-}$ mice lacking expression of *mm*MAVS are highly permissive.[14] *mm*MAVS and *hs*MAVS have only about 50% amino acid identity, and they differ at 3 residues when aligned at the cleavage site in *hs*MAVS targeted by human HAV (Gln^{427}/Val^{428}, Fig. 5C). These differences include the p'1 residue that is a major contributor to substrate activity of peptides[28] (Fig. 2B). Consistent with this, a Leu to Val substitution in *mm*MAVS at the residue aligning with p'1 in *hs*MAVS, coupled with a Glu to Ser substitution at the p'3 position (*mm*MAVS^{VS}) resulted in robust cleavage of the mutated *mm*MAVS by HAV 3ABC, but not by the catalytically inactive 3ABC* (Fig. 5C). Substitutions at all 3 divergent residues (*mu*MAVS$^{A\text{-}VS}$), effectively 'humanizing' *mm*MAVS at the cleavage site, similarly conferred 3ABC cleavage. These data provide evidence that the major constraint on 3ABC cleavage is sequence at the direct site of cleavage, where we find evidence of diversifying selection (Fig. 5B), and not elsewhere within MAVS proteins.

Discussion

HAV is distinct from hepaciviruses like HCV and hepadnaviruses like HBV that commonly cause long-term persistent infection, and thus do not require a continued high rate of transmission between individuals to maintain their presence in human populations. In contrast to HAV, which likely became established among humans only with the development of population densities sufficient to support continuous chains of fecal-oral transmission, these hepatotropic viruses appear to have infected primates for millions of years.[4–6] For example, divergent hepaciviruses have been recovered from both New World tamarins and Old World colobus monkeys,[37] non-human primates that have been separated for millions of years of evolution. Divergent hepaciviruses have evolved over eons in primates, explaining why there are signs of evolutionary pressure at the site of HCV NS3/4A cleavage in primate MAVS orthologs.[4] We confirmed this finding, while finding no pressure at the site of HAV 3ABC cleavage site, consistent with the more recent establishment of HAV among primate hosts (Fig. 5B). Nonetheless, it is noteworthy that divergent genotypes of HAV have also been recovered from non-human primates.[38–40] These viruses are not well studied, and many may have been derived ultimately from a human source. However, some of these viruses, such as the AGM27 strain recovered from an Old World monkey, have biological and even antigenic differences from human HAV,[40] and they deserve further study.

Phylogenetic analyses indicate that the past evolution of hepatoviruses has involved multiple shifts between host species in different mammalian orders.[7] While ancestral state reconstructions suggest that rodents hosted the immediate ancestor of human HAV, hepatoviruses likely have a broader origin in bats and potentially other insectivorous small mammals.[7,41] The factors allowing hepatoviruses to sustain such dramatic host species shifts are not understood, but such events are likely to be favored by the ability of 3ABC proteases to cleave MAVS and disrupt innate immune signaling across broadly divergent mammalian species. X-ray crystallographic studies indicate that the active site of the hepatovirus 3C^{pro} protease is more flexible than that of other, structurally-related chymotrypsin-like proteases.[29] This feature could contribute to its relatively broad substrate specificity. Although not examined in this study, it is intriguing to speculate that the 3C^{pro} protease may also have influenced evolution of its mammalian host species through its interactions with NEMO (*IKBKG*),[42] a crucial adaptor that bridges NF-κB and IRF antiviral signaling,[43] or TRIF (*TICAM1*), an adaptor protein essential for TLR3 signaling.[44]

Previous studies indicate MAVS signaling is a key roadblock that must be overcome for HAV to successfully infect mice.[14] Consistent with this, we found a strong signal for diversifying selection among rodent MAVS orthologs at residues aligning with the site of 3ABC cleavage in *hs*MAVS, Gln^{427}/Val^{428} (Fig. 5B, Table S4). In contrast, diversifying selection among chiropteran MAVS was found to have occurred primarily at an alternative site in the protein that aligns with the site of *hs*MAVS cleavage by bat viral proteases: Glu^{463}/Gly^{464}. This site is highly divergent in bat MAVS (Fig. 2B), consistent with the lack of cleavage by bat hepatoviral proteases. These data are consistent with hepatoviruses having existed in bats for much longer than humans, perhaps millions of years,[7] allowing more time for selective pressure on MAVS and the evolution of 3ABC-resistant MAVS proteins. Hepatovirus infections are rarely lethal in humans, and long-term, persistent HAV infection has never been described. Such limited pathogenicity could also contribute to a lack of diversifying selection at sites of HAV 3ABC cleavage in *hs*MAVS. The consequences of infection could be more serious in Chiroptera, in which our data suggest the possibility of host escape from 3ABC cleavage. It is clear, however, that hepatoviruses have established alternative strategies for their survival in this mammalian order.

Do contemporary bat viruses pose a risk for human populations, given their similar capsid structures, common antigenic determinants, shared tropism for the liver,[7] and, as we show here, capacity to disrupt a critical mediator of innate immunity in humans? None of these viruses has yet been isolated in infectious form and their natural host range and potential pathogenicity in humans remain uncertain. Yet, the history of hepatovirus evolution tells us this is a threat that cannot be discounted. Understanding the capacity of viruses to evade human innate immune responses is crucial for risk assessments and integrated evolutionary and experimental studies such as those described here should help elucidate factors controlling virus emergence and the risks posed to humans.

Financial support

This work was supported by the U.S. National Institutes of Health (NIH), grants U19-AI109965, R01-AI103083 and R01-AI131685 (S.M.L.), and the University of North Carolina Cancer Research Fund.

Conflict of interest

The authors declare no competing financial interests.

Authors' contributions

S.M.L. and H.F. designed the experiments; H.F., A-L.S., and A.M-S. performed experiments; D.Y. provided key reagents; J.F.D. and A.L.S. designed and conducted phylogenetic and evolutionary analyses; S.M.L. supervised the study; H.F., J.F.D. and S.M.L. analyzed the data and wrote the manuscript; all authors reviewed the manuscript.

Acknowledgements

We thank Dr. Joseph Marcotrigiano of the National Institute of Allergy and Infectious Diseases for helpful discussions.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhep.2019.02.020.

References

[1] Lemon SM, Ott JJ, Van Damme P, Shouval D. Type A viral hepatitis: a summary and update on the molecular virology, epidemiology, pathogenesis and prevention. J Hepatol 2018;68:167–184.
[2] Wong DC, Purcell RH, Rosen L. Prevalence of antibody to hepatitis A and hepatitis B viruses in selected populations of the South Pacific. Amer J Epidemiol 1979;110:227–236.
[3] Skinhoj P, Mikkelsen F, Hollinger FB. Hepatitis A in Greenland: importance of specific antibody testing in epidemiologic surveillance. Amer J Epidemiol 1977;105:140–147.
[4] Patel MR, Loo YM, Horner SM, Gale Jr M, Malik HS. Convergent evolution of escape from hepaciviral antagonism in primates. PLoS Biol 2012;10:e1001282.
[5] de Carvalho Dominguez Souza BF, Konig A, Rasche A, de Oliveira Carneiro I, Stephan N, Corman VM, et al. A novel hepatitis B virus species discovered in capuchin monkeys sheds new light on the evolution of primate hepadnaviruses. J Hepatol 2018;68:1114–1122.
[6] Littlejohn M, Locarnini S, Yuen L. Origins and evolution of hepatitis B virus and hepatitis D virus. Cold Spring Harbor Perspect Med 2016;6:a021360.
[7] Drexler JF, Corman VM, Lukashev AN, van den Brand JM, Gmyl AP, Brunink S, et al. Evolutionary origins of hepatitis A virus in small mammals. Proc Natl Acad Sci USA 2015;112:15190–15195.
[8] de Oliveira Carneiro I, Sander AL, Silva N, Moreira-Soto A, Normann A, Flehmig B, et al. A novel marsupial hepatitis A virus corroborates complex evolutionary patterns shaping the genus Hepatovirus. J Virol 2018;92:e00082–18.
[9] Kitchen A, Shackelton LA, Holmes EC. Family level phylogenies reveal modes of macroevolution in RNA viruses. Proc Natl Acad Sci USA 2011;108:238–243.
[10] Woolhouse ME, Haydon DT, Antia R. Emerging pathogens: the epidemiology and evolution of species jumps. Trends Ecol Evol 2005;20:238–244.
[11] Lanford RE, Feng Z, Chavez D, Guerra B, Brasky KM, Zhou Y, et al. Acute hepatitis A virus infection is associated with a limited type I interferon response and persistence of intrahepatic viral RNA. Proc Natl Acad Sci USA 2011;108:11223–11228.
[12] Yang Y, Liang Y, Qu L, Chen Z, Yi M, Li K, et al. Disruption of innate immunity due to mitochondrial targeting of a picornaviral protease precursor. Proc Natl Acad Sci USA 2007;104:7253–7258.
[13] Lanford RE, Walker CM, Lemon SM. Nonhuman primate models of hepatitis A virus and hepatitis E virus infections. Cold Spring Harbor Perspect Med 2018.
[14] Hirai-Yuki A, Hensley L, McGivern DR, Gonzalez-Lopez O, Das A, Feng H, et al. MAVS-dependent host species range and pathogenicity of human hepatitis A virus. Science 2016;353:1541–1545.
[15] Blight KJ, McKeating JA, Rice CM. Highly permissive cell lines for subgenomic and genomic hepatitis C virus RNA replication. J Virol 2002;76:13001–13014.
[16] Ikeda M, Sugiyama K, Mizutani T, Tanaka T, Tanaka K, Sekihara H, et al. Human hepatocyte clonal cell lines that support persistent replication of hepatitis C virus. Virus Res 1998;56:157–167.
[17] Feng H, Lenarcic EM, Yamane D, Wauthier E, Mo J, Guo H, et al. NLRX1 promotes immediate IRF1-directed antiviral responses by limiting dsRNA-activated translational inhibition mediated by PKR. Nat Immunol 2017;18:1299–1309.
[18] Lemon SM, Murphy PC, Shields PA, Ping LH, Feinstone SM, Cromeans T, et al. Antigenic and genetic variation in cytopathic hepatitis A virus variants arising during persistent infection: evidence for genetic recombination. J Virol 1991;65:2056–2065.
[19] Bryksin AV, Matsumura I. Overlap extension PCR cloning: a simple and reliable way to create recombinant plasmids. Biotechniques 2010;48:463–465.
[20] Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 2003;19:1572–1574.
[21] Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 2007;7:214.
[22] Xu B, Yang Z. PAMLX: a graphical user interface for PAML. Mol Biol Evol 2013;30:2723–2724.
[23] Yang Z. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput Appl Biosci 1997;13:555–556.
[24] Kosakovsky Pond SL, Frost SD. Not so different after all: a comparison of methods for detecting amino acid sites under selection. Mol Biol Evol 2005;22:1208–1222.
[25] Murrell B, Moola S, Mabona A, Weighill T, Sheward D, Kosakovsky Pond SL, et al. FUBAR: a fast unconstrained bayesian approximation for inferring selection. Mol Biol Evol 2013;30:1196–1205.
[26] Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 2016;33:1870–1874.
[27] Simmonds P. SSE: a nucleotide and amino acid sequence analysis platform. BMC Res Notes 2012;5:50.
[28] Petithory JR, Masiarz FR, Kirsch JF, Santi DV, Malcolm BA. A rapid method for determination of endoproteinase substrate specificity: specificity of the 3C proteinase from hepatitis A virus. Proc Natl Acad Sci USA 1991;88:11510–11514.
[29] Bergmann EM, Cherney MM, McKendrick J, Frormann S, Luo C, Malcolm BA, et al. Crystal structure of an inhibitor complex of the 3C proteinase from hepatitis A virus (HAV) and implications for the polyprotein processing in HAV. Virology 1999;265:153–163.
[30] Solari S, Baker RJ. Mammal Species of the world: a taxonomic and geographic reference by D.E. Wilson; D.M. Reeder. J Mammal 2007;88:824–830.
[31] Foley NM, Springer MS, Teeling EC. Mammal madness: is the mammal tree of life not yet resolved? Phil Trans R Soc Lond B Biol Sci 2016;371.
[32] Murrell B, Weaver S, Smith MD, Wertheim JO, Murrell S, Aylward A, et al. Gene-wide identification of episodic selection. Mol Biol Evol 2015;32:1365–1371.
[33] Li XD, Sun L, Seth RB, Pineda G, Chen ZJ. Hepatitis C virus protease NS3/4A cleaves mitochondrial antiviral signaling protein off the mitochondria to evade innate immunity. Proc Natl Acad Sci USA 2005;102:17717–17722.
[34] Meylan E, Curran J, Hofmann K, Moradpour D, Binder M, Bartenschlager R, et al. Cardif is an adaptor protein in the RIG-I antiviral pathway and is targeted by hepatitis C virus. Nature 2005;437:1167–1172.
[35] Drexler JF, Corman VM, Muller MA, Lukashev AN, Gmyl A, Coutard B, et al. Evidence for novel hepaciviruses in rodents. PLoS Pathog 2013;9:e1003438.
[36] Quan PL, Firth C, Conte JM, Williams SH, Zambrana-Torrelio CM, Anthony SJ, et al. Bats are a major natural reservoir for hepaciviruses and pegiviruses. Proc Natl Acad Sci USA 2013;110:8194–8199.
[37] Rasche A, Sander A-L, Corman VM, Drexler JF. Evolutionary biology of human hepatitis viruses. J Hepatol 2019;70:501–520.
[38] Lemon SM, LeDuc JW, Binn LN, Escajadillo A, Ishak KG. Transmission of hepatitis A virus among recently captured Panamanian owl monkeys. J Med Virol 1982;10:25–36.
[39] Nainan OV, Margolis HS, Robertson BH, Balayan M, Brinton MA. Sequence analysis of a new hepatitis A virus naturally infecting cynomolgus macaques (*Macaca fascicularis*). J Gen Virol 1991;72:1685–1689.
[40] Tsarev SA, Emerson SU, Balayan MS, Ticehurst J, Purcell RH. Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure and growth in cell culture with other HAV strains. J Gen Virol 1991;72:1677–1683.
[41] Wang X, Ren J, Gao Q, Hu Z, Sun Y, Li X, et al. Hepatitis A virus and the origins of picornaviruses. Nature 2015;517:85–88.
[42] Wang D, Fang L, Wei D, Zhang H, Luo R, Chen H, et al. Hepatitis A virus 3C protease cleaves NEMO to impair induction of beta interferon. J Virol 2014;88:10252–10258.

[43] Zhao T, Yang L, Sun Q, Arguello M, Ballard DW, Hiscott J, et al. The NEMO adaptor bridges the nuclear factor-kappaB and interferon regulatory factor signaling pathways. Nat Immunol 2007;8:592–600.

[44] Qu L, Feng Z, Yamane D, Liang Y, Lanford RE, Li K, et al. Disruption of TLR3 signaling due to cleavage of TRIF by the hepatitis A virus protease-polymerase processing intermediate, 3CD. PLoS Pathog 2011;7:e1002169.

[45] Brubaker SW, Gauthier AE, Mills EW, Ingolia NT, Kagan JC. A bicistronic MAVS transcript highlights a class of truncated variants in antiviral immunity. Cell 2014;156:800–811.

[46] Rebsamen M, Meylan E, Curran J, Tschopp J. The antiviral adaptor proteins Cardif and Trif are processed and inactivated by caspases. Cell Death Differ 2008;15:1804–1811.