Published in final edited form as:

J Hepatol. 2019 July ; 71(1): 25–34. doi:10.1016/j.jhep.2019.02.020.

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

Hui Feng^{1,2}, Anna-Lena Sander³, Andrés Moreira-Soto³, Daisuke Yamane^{1,2,*}, Jan Felix Drexler^{3,4}, and Stanley M. Lemon^{1,2,5}

¹Lineberger Comprehensive Cancer Center, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA;

²Department of Medicine, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA;

³Charité-Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Institute of Virology, 10117 Berlin, Germany;

⁴German Centre for Infection Research (Deutsches Zentrum für Infektionsforschung), Berlin, Germany;

⁵Department of Microbiology & Immunology, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

SUMMARY

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 infecting rodents and bats suggest 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 hepatitis virus 3ABC proteases to cleave MAVS and disrupt antiviral signaling in HEK293 and human hepatocyte-derived cell lines, and applied maximum-likelihood and Bayesian algorithms to identify sites of diversifying selection in MAVS orthologs from 75 chiropteran, rodent and primate species.

Corresponding author: Stanley M. Lemon, 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, smlemon@med.unc.edu.

*Current address: Department of Microbiology and Cell Biology, Tokyo Metropolitan Institute of Medical Science, Setagaya-ku, Tokyo, Japan

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.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflict of interest: The authors declare no competing financial interests.

Results: 3ABC proteases from bat, but not rodent hepatoviruses efficiently cleaved human MAVS at Glu⁴⁶³/Gly⁴⁶⁴, disrupting virus activation of the interferon- β promoter, whereas human HAV 3ABC cleaved at Gln⁴²⁷/Val⁴²⁸. 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 three 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

HAV, 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 infecting 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.

Keywords

hepatitis A virus; chymotrypsin-like protease; innate immunity; virus evolution; cross-species transmission

INTRODUCTION

Despite effective vaccines, hepatitis A virus (HAV) remains a common cause of acute viral hepatitis in many regions of the world¹. 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 longterm persistent 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 HAV⁷.

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 liver¹¹. This is due in part to cleavage of mitochondrial antiviral signaling protein (MAVS), a critical innate immune adaptor protein, by a precursor of the mature HAV 3C^{pro} protease, 3ABC¹². Although the natural host range of HAV appears limited to primates¹³, the virus robustly infects *Mavs*^{-/-} knockout mice¹⁴. Unlike the human MAVS protein (*hsMAVS*), murine MAVS (*mmMAVS*) 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 *hsMAVS*, and whether this host-pathogen interaction has shaped the past evolution of hepatoviruses and their hosts.

METHODS

Cells.

HEK293FT cells and human hepatocyte-derived cell lines including Huh-7.5¹⁵, PH5CH8¹⁶, 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 HAV¹⁸ 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 previously¹⁷. Sendai virus (SeV, Cantell strain) was purchased from Charles River Laboratory.

Virus Infection and Transduction.

Cultures were infected with HAV (10⁶ genome equivalents/culture) or SeV (600 HA units/mL) at 37°C for 4 and 6 hrs, 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)FLAG¹² and pcDNA3-(N)FLAG/eGFP were used to express MAVS sequences. Lentivector pLOC GFP¹⁷ 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 strategy¹⁹ that generated a nonrelevant 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-*h*MAVS (100 ng) and either empty vector (EV) or nonprimate hepatoviral 3ABC proteases (HA-3ABC, 100 ng). GFP-fused nonprimate 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 *IFNB1* promoter-driven firefly luciferase plasmid (40 ng IFN β -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 hrs (Huh-7.5 and PH5CH8 cells) post transfection, cells were lysed and immunoblotting carried out as described¹⁷. For detailed primary antibodies information, please check 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 hrs post-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 Sendai virus versus 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 V3.1²⁰ and a WAG substitution model. Trees were run for two million generations, sampled every 100 steps. After excluding 5000 of the total 20,000 trees as burn-in, final trees were annotated with TreeAnnotator from the BEAST package²¹ 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 three 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 (LRT) on two pairs of site-specific models, M1a versus M2a and M7 versus M8, were performed to determine whether MAVS genes from each the three orders are evolving under positive selection using the chi-square calculator implemented in the pamlX package. 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 (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 MEGA7²⁶ 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 V1.2²⁷. 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 two-sided *t*-test using Prism 6.0 (GraphPad Software, Inc.).

RESULTS

3ABC proteases from bat hepatoviruses cleave human *hsMAVS*.

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 (Figure 1A). We were surprised therefore to find that N-terminally HA-tagged 3ABC proteases from two bat hepatoviruses, M32 (*Eidolon helvum*) or FO1A-F48 ('F48', *Rhinolophus landeri*), cleaved *hsMAVS* fused at its amino terminus to GFP and co-expressed in HEK293 cells (Figure 1B). Replacing the active-site Cys residues in these proteases with Ala (M32* and F48*, respectively) abolished *hsMAVS* cleavage (Figure 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 *hsMAVS* or did so less efficiently (Figure 1B), although these rodent virus proteases are phylogenetically closer to human HAV 3ABC than the bat hepatovirus proteases (Figure

1A). Ablating the active site Cys residues in the rodent proteases (KS11* and KEF121*, respectively) enhanced their levels of expression (Figure 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 Sendai virus (SeV)-induced activation of the interferon- β (*IFNB1*) promoter, whereas the catalytically-inactive mutants did not (Figure 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 IRF3 (Figure 1D). Consistent with these data, HAV replication was significantly enhanced in lentivirus-transduced PH5CH8 cells expressing the wild-type versus mutant bat proteases, despite the fact that the human viral 3ABC protease was also expressed by the HAV replicating in these cells (Figure 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 *hsMAVS* at Glu⁴⁶³/Gly⁴⁶⁴, rather than Gln⁴²⁷/Val⁴²⁸, where *hsMAVS* is targeted by the human virus¹² (Figure 2A). Both sites in *hsMAVS* 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 Gln⁴²⁷/Val⁴²⁸ site²⁸ (Figure 2B). However, peptides with Gly at the p'1 position (as in the Glu⁴⁶³/Gly⁴⁶⁴ site) are cleaved significantly more efficiently than those with p'1 Val. Overall, the preference for Glu⁴⁶³/Gly⁴⁶⁴ 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 protease²⁹ (Figure 2C).

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* (*rsMAVS*). Similarly, we expressed MAVS from the hispid cotton rat (*S. hispidus*, *shMAVS*) and the Chinese hamster (*C. griseus*, *cgMAVS*) 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 host, *E. helvum* (*ehMAVS*), in PH5CH8 cells knocked out for endogenous MAVS expression by CRISPR/Cas9 gene editing (Figure 3A). The synthesis of IFIT1, an interferon-stimulated gene product, was also stimulated (Figure 3A). Moreover, RIG-I signaling could be efficiently reconstituted in these cells by representative bat (*rsMAVS*) and rodent (*cgMAVS*) MAVS proteins, resulting in SeV-induced IRF3 dimerization (Figure 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 *rsMAVS*, and no cleavage of *ehMAVS* by ectopically expressed human HAV 3ABC (Figure 3C, lanes 1–2 and 7–8). Consistent with this, *IFNBI* promoter activation induced by expression of *rsMAVS*, but not *ehMAVS*, was reduced by co-expressing HAV 3ABC versus the catalytically-inactive mutant 3ABC* (Figure 3D). Importantly, the p1 Gln residue is preserved in *rsMAVS* at the position aligning with Gln⁴²⁷ in *hsMAVS* (Gln⁴²¹), whereas it is replaced by Cys in *ehMAVS* (Figure 2B). By contrast, there is no conservation of the p1 and p'1 residues aligning with Glu⁴⁶³/Gly⁴⁶⁴ in either bat MAVS ortholog. Surprisingly, HAV 3ABC robustly cleaved *shMAVS* (cotton rat), and less efficiently, *cgMAVS* (hamster) (Figure 3C, lanes 13–14 and 17–18). These cleavages were functionally important, as co-expressing the active protease reduced *IFNBI* promoter activation induced by either rodent MAVS ortholog (Figure 3D). In line with these results, the amino acid sequence around the Gln⁴²⁷/Val⁴²⁸ cleavage site in *hsMAVS* is well preserved in alignments of the rodent orthologs, whereas the Glu⁴⁶³/Gly⁴⁶⁴ site targeted by bat hepatovirus 3ABC proteases is ablated by deletion of 3 residues (Figure 2B). Further experiments confirmed that the HAV 3ABC protease cleaves each of these MAVS orthologs at residues aligning with Gln⁴²⁷/Val⁴²⁸ in *hsMAVS* (Figures 3E).

Despite cleaving human *hsMAVS* (Figure 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 (Figure 3C, lanes 3–6 and 9–12). This is consistent with the Glu⁴⁶³/Gly⁴⁶⁴ site targeted by these proteases in *hsMAVS* being poorly conserved in bat MAVS (Figure 2B). Similarly, the cotton rat *shMAVS* was only cleaved minimally, and the hamster *cgMAVS* cleaved not at all, by the cognate rodent 3ABC proteases (Figure 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 nonprimate viruses was capable of significantly disrupting *IFNBI* promoter signaling transduced through either cognate or noncognate rodent or bat MAVS orthologs (Figure 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 *hsMAVS* (Figure 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 years⁷. Two-thirds of ~5400 extant mammalian species belong to the orders Rodentia and Chiroptera³⁰, which together host the greatest diversity of hepatoviruses (Figure 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 four bat (*E. helvum*, *Hypsignathus monstrosus*, *Pipistrellus pipistrellus*, *R. ferrumequinum*) and two 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 (Figure 4B). Surprisingly, MAVS diversity was not related to overall host genetic relationships. *hsMAVS* shared less identity with rodent than bat MAVS (Figure 4C), although rodents and primates are genetically closer than bats and primates³¹. Sequence distances were unevenly distributed across the MAVS reading frame (Figure 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 three mammalian orders, irregardless of the method used for calculation (Figure 5A). However, significant episodic diversifying selection was evident in primate, rodent and chiropteran MAVS ($p < 0.001$) according to both likelihood ratio tests (Table 1) and a branch-site unrestricted statistical test (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 (Figure 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 two different orders, suggesting similar evolutionary pressures shaping MAVS across different orders.

Using any of the five analytic methods, we found no evidence for diversifying selection at the site targeted by the human HAV protease, Gln⁴²⁷/Val⁴²⁸, among MAVS proteins expressed by primates (Figure 5B, Table S4). This is consistent with the relatively recent introduction of HAV into primate populations, as discussed above. In contrast, as reported previously⁴, PAML detected a signal for selection at residues aligning with Cys⁵⁰⁸/His⁵⁰⁹, where *hsMAVS* is cleaved by the HCV NS³/A protease^{33,34}. This is consistent with evidence suggesting the existence of hepaciviruses related to HCV among primates for millions of years⁴. We also noted evidence for diversifying selection at residues aligning with this site in MAVS orthologs expressed by rodents and bats (Figure 5B, Table S4), which also host diverse hepaciviruses^{35,36}.

Despite the absence of diversifying selection at the HAV 3ABC cleavage site, Gln⁴²⁷/Val⁴²⁸, 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 (Figure 5B, Table S4). These results are consistent with ancestral reconstructions which suggest rodents hosted the immediate ancestor of human HAV⁷. Just as striking, however, was a very strong signal for diversifying selection among chiropteran MAVS orthologs at residues aligning with Lys⁴⁶¹ and Thr⁴⁶⁵ (Figure 2B), which bracket the site of *hsMAVS* cleavage by the bat viral proteases (Figure 5B, Table S4). These data are consistent with MAVS having evolved over millenia to a relatively 3ABC-resistant phenotype in these two 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 *hsMAVS*. To answer this question, we co-expressed either wild-type or minimally modified murine (*Mus musculus*) *mmMAVS* and 3ABC from human HAV in 293 cells. Our previous work shows that *mmMAVS* is not susceptible to 3ABC cleavage, and that normal mice cannot be infected with HAV, whereas *Mavs*^{-/-} mice lacking expression of *mmMAVS* are highly permissive¹⁴. *mmMAVS* and *hsMAVS* have only about 50% amino acid identity, and they differ at 3 residues when aligned at the cleavage site in *hsMAVS* targeted by human HAV (Gln⁴²⁷/Val⁴²⁸, Figure 5C). These differences include the p'1 residue that is a major contributor to substrate activity of peptides²⁸ (Figure 2B). Consistent with this, a Leu to Val substitution in *mmMAVS* at the residue aligning with p'1 in *hsMAVS*, coupled with a Glu to Ser substitution at the p'3 position (*mmMAVS*^{VS}) resulted in robust cleavage of the mutated *mmMAVS* by HAV 3ABC, but not by the catalytically-inactive 3ABC* (Figure 5C). Substitutions at all three divergent residues (*mmMAVS*^{A-VS}), effectively 'humanizing' *mmMAVS* 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 (Figure 5B), and not elsewhere within MAVS proteins.

DISCUSSION

HAV is distinct from hepaciviruses like HCV and hepadnaviruses like HBV that commonly cause longterm 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 years⁴⁻⁶. For example, divergent hepaciviruses have been recovered from both New World tamarins and Old World colobus monkeys³⁷, nonhuman 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/A cleavage site in primate MAVS orthologs⁴. 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 (Figure 5B). On the other hand, it is noteworthy that divergent genotypes of HAV have also been recovered from nonhuman primates³⁸⁻⁴⁰. 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 HAV⁴⁰, and they deserve further study.

Phylogenetic analyses indicate that the past evolution of hepatoviruses has involved multiple shifts between host species in different mammalian orders⁷. 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 hepatitis C virus 3C^{pro} protease is more flexible than that of other, structurally-related chymotrypsin-like proteases²⁹. 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 NEMO⁴², a crucial adaptor that bridges NF- κ B and IRF antiviral signaling⁴³, or TRIF, an adaptor protein essential for TLR3 signaling⁴⁴.

Previous studies indicate MAVS signaling is a key roadblock that must be overcome for HAV to successfully infect mice¹⁴. 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 *hsMAVS*, Gln⁴²⁷/Val⁴²⁸ (Figure 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 *hsMAVS* cleavage by bat viral proteases: Glu⁴⁶³/Gly⁴⁶⁴. This site is highly divergent in bat MAVS (Figure 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 years⁷, allowing more time for selective pressure on MAVS and the evolution of 3ABC-resistant MAVS proteins. Hepatovirus infections are rarely lethal in humans, and longterm, 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 *hsMAVS*. 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 liver⁷, 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 it poses for humans.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

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

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

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* (2017).
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* 110, 227–236 (1979). [PubMed: 224696]
3. Skinhoj P, Mikkelsen F & Hollinger FB Hepatitis A in Greenland: importance of specific antibody testing in epidemiologic surveillance. *Amer J Epidemiol* 105, 140–147 (1977). [PubMed: 189600]
4. Patel MR, Loo YM, Horner SM, Gale M Jr. & Malik HS Convergent evolution of escape from hepaciviral antagonism in primates. *PLoS Biol* 10, e1001282 (2012). [PubMed: 22427742]
5. de Carvalho Dominguez Souza BF et al. A novel hepatitis B virus species discovered in capuchin monkeys sheds new light on the evolution of primate hepadnaviruses. *J Hepatol* 68, 1114–1122 (2018). [PubMed: 29428874]
6. Littlejohn M, Locarnini S & Yuen L Origins and evolution of hepatitis B virus and hepatitis D virus. *Cold Spring Harb Perspect Med* 6, a021360 (2016). [PubMed: 26729756]
7. Drexler JF et al. Evolutionary origins of hepatitis A virus in small mammals. *Proc Natl Acad Sci U S A* 112, 15190–15195 (2015). [PubMed: 26575627]
8. de Oliveira Carneiro I et al. A novel marsupial hepatitis A virus corroborates complex evolutionary patterns shaping the genus Hepatovirus. *J Virol* (2018).
9. Kitchen A, Shackelton LA & Holmes EC Family level phylogenies reveal modes of macroevolution in RNA viruses. *Proc Natl Acad Sci U S A* 108, 238–243 (2011). [PubMed: 21173251]
10. Woolhouse ME, Haydon DT & Antia R Emerging pathogens: the epidemiology and evolution of species jumps. *Trends Ecol Evol* 20, 238–244 (2005). [PubMed: 16701375]
11. Lanford RE et al. Acute hepatitis A virus infection is associated with a limited type I interferon response and persistence of intrahepatic viral RNA. *Proc Nat'l Acad Sci USA* 108, 11223–11228 (2011). [PubMed: 21690403]
12. Yang Y et al. Disruption of innate immunity due to mitochondrial targeting of a picornaviral protease precursor. *Proc Natl Acad Sci U S A* 104, 7253–7258 (2007). [PubMed: 17438296]
13. Lanford RE, Walker CM & Lemon SM Nonhuman primate models of hepatitis A virus and hepatitis E virus infections. *Cold Spring Harb Perspect Med* doi: 10.1101/cshperspect.a031815 (2018).
14. Hirai-Yuki A et al. MAVS-dependent host species range and pathogenicity of human hepatitis A virus. *Science* 353, 1541–1545 (2016). [PubMed: 27633528]
15. Blight KJ, McKeating JA & Rice CM Highly permissive cell lines for subgenomic and genomic hepatitis C virus RNA replication. *J Virol* 76, 13001–13014 (2002). [PubMed: 12438626]
16. Ikeda M et al. Human hepatocyte clonal cell lines that support persistent replication of hepatitis C virus. *Virus Res* 56, 157–167 (1998). [PubMed: 9783464]
17. Feng H et al. NLRX1 promotes immediate IRF1-directed antiviral responses by limiting dsRNA-activated translational inhibition mediated by PKR. *Nat Immunol* 18, 1299–1309 (2017). [PubMed: 28967880]
18. Lemon SM et al. Antigenic and genetic variation in cytopathic hepatitis A virus variants arising during persistent infection: evidence for genetic recombination. *J Virol* 65, 2056–2065 (1991). [PubMed: 1705995]
19. Bryksin AV & Matsumura I Overlap extension PCR cloning: a simple and reliable way to create recombinant plasmids. *Biotechniques* 48, 463–465 (2010). [PubMed: 20569222]
20. Ronquist F & Huelsenbeck JP MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572–1574 (2003). [PubMed: 12912839]
21. Drummond AJ & Rambaut A BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7, 214 (2007). [PubMed: 17996036]
22. Xu B & Yang Z PAMLX: a graphical user interface for PAML. *Mol Biol Evol* 30, 2723–2724 (2013). [PubMed: 24105918]

23. Yang Z PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput Appl Biosci* 13, 555–556 (1997). [PubMed: 9367129]
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* 22, 1208–1222 (2005). [PubMed: 15703242]
25. Murrell B et al. FUBAR: a fast, unconstrained bayesian approximation for inferring selection. *Mol Biol Evol* 30, 1196–1205 (2013). [PubMed: 23420840]
26. Kumar S, Stecher G & Tamura K MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* 33, 1870–1874 (2016). [PubMed: 27004904]
27. Simmonds P SSE: a nucleotide and amino acid sequence analysis platform. *BMC Res Notes* 5, 50 (2012). [PubMed: 22264264]
28. Petithory JR, Masiarz FR, Kirsch JF, Santi DV & Malcolm BA A rapid method for determination of endoprotease substrate specificity: specificity of the 3C proteinase from hepatitis A virus. *Proc Natl Acad Sci U S A* 88, 11510–11514 (1991). [PubMed: 1662396]
29. Bergmann EM 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*. 265, 153–163 (1999). [PubMed: 10603326]
30. Solari S & Baker RJ Mammal Species of the World: A Taxonomic and Geographic Reference by D. E. Wilson; D. M. Reeder. *J Mammal* 88, 824–830 (2007).
31. Foley NM, Springer MS & Teeling EC Mammal madness: is the mammal tree of life not yet resolved? *Phil Trans Royal Soc Lond. Series B, Biological sciences* 371 (2016).
32. Murrell B et al. Gene-wide identification of episodic selection. *Mol Biol Evol* 32, 1365–1371 (2015). [PubMed: 25701167]
33. Li XD, Sun L, Seth RB, Pineda G & Chen ZJ Hepatitis C virus protease NS₃A cleaves mitochondrial antiviral signaling protein off the mitochondria to evade innate immunity. *Proc.Natl.Acad.Sci U.S.A.* 102, 17717–17722 (2005). [PubMed: 16301520]
34. Meylan E et al. Cardif is an adaptor protein in the RIG-I antiviral pathway and is targeted by hepatitis C virus. *Nature* 437, 1167–1172 (2005). [PubMed: 16177806]
35. Drexler JF et al. Evidence for novel hepaciviruses in rodents. *PLoS Pathog* 9, e1003438 (2013). [PubMed: 23818848]
36. Quan PL et al. Bats are a major natural reservoir for hepaciviruses and pegiviruses. *Proc Natl Acad Sci U S A* 110, 8194–8199 (2013). [PubMed: 23610427]
37. Rasche A, Sander A-L, Corman VM & Drexler JF Evolutionary biology of human hepatitis viruses. *J Hepatol* (2018).
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* 10, 25–36 (1982). [PubMed: 6290600]
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* 72, 1685–1689 (1991). [PubMed: 1649902]
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* 72, 1677–1683 (1991). [PubMed: 1649901]
41. Wang X et al. Hepatitis A virus and the origins of picornaviruses. *Nature* 517, 85–88 (2015). [PubMed: 25327248]
42. Wang D et al. Hepatitis A virus 3C protease cleaves NEMO to impair induction of beta interferon. *J Virol* 88, 10252–10258 (2014). [PubMed: 24920812]
43. Zhao T et al. The NEMO adaptor bridges the nuclear factor-kappaB and interferon regulatory factor signaling pathways. *Nat Immunol* 8, 592–600 (2007). [PubMed: 17468758]
44. Qu L et al. Disruption of TLR3 signaling due to cleavage of TRIF by the hepatitis A virus protease-polymerase processing intermediate, 3CD. *PLoS Pathog* 7, e1002169 (2011). [PubMed: 21931545]

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* 156, 800–811 (2014). [PubMed: 24529381]
46. Rebsamen M, Meylan E, Curran J & Tschoop J The antiviral adaptor proteins Cardif and Trif are processed and inactivated by caspases. *Cell Death Differ* 15, 1804–1811 (2008). [PubMed: 18756281]

HIGHLIGHTS

- 3ABC proteases expressed by bat hepatoviruses cleave human MAVS at Glu⁴⁶³/Gly⁴⁶⁴.
- 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.

Figure 1. Bat hepatovirus 3ABC proteases cleave human *hsMAVS*.

(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 *hsMAVS* 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. EV, empty vector; *, catalytically-inactive 3ABC proteases; FL, full length GFP-tagged MAVS and endogenous MAVS; C, cleavage product; Mini, a truncated MAVS variant identified previously⁴⁵. The indicated section of the central panel is shown with high sensitivity detection on the right. (C) SeV activation of the *IFNβ* 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 one 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.

Figure 2. Substrate specificities of bat and human hepatoviral 3ABC proteases.

(A) Bat hepatovirus 3ABC cleavage of *hsMAVS* mutants with amino acid substitutions at p1 positions of possible cleavage sites. The D429E mutation ablates caspase cleavage in apoptotic cells⁴⁶. WT=wild-type *hsMAVS*. See legend to Fig. 1B. (B) Alignments of human, rodent and bat MAVS amino acid sequences at 3ABC cleavage sites in *hsMAVS*. Cleavage sites targeted by 3C^{pro} in the HAV polypeptide (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 (Lys¹⁴⁷, which abuts the substrate cleft near the S3 specificity pocket). Active-site His⁴⁴ and Cys¹⁷² residues are colored red.

Figure 4. Hepatovirus diversity and evidence for episodic diversifying selection in MAVS. (A) Partial VP2-VP3 hepatovirus (HepV) 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. MRCA (mya), most recent common ancestor of orders (millions of years ago). (C) Mean pairwise sequence identities $\pm$ SD between *hsMAVS* and rodent or bat orthologs. **** $p < 0.0001$ by *t*-test. (D) Amino acid sequence identities along the length of *hsMAVS* with (top) rodent and (bottom) chiropteran MAVS orthologs. CARD, CARD-domain; Pro, proline-rich region; TM, transmembrane domain.

Figure 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 one of the five methods. See Supplementary Table 4 for details. (C) Immunoblots of Huh-7.5 cells co-expressing murine *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.

Table 1.

Likelihood ratio tests (LRTs) for selected PAML site-specific models.

Selected site models	$\ln L_0$	$\ln L_1$	2 $\ln L$	df	p-value
Primates					
M1a _(nearly neutral) vs. M2a _(pos. selection)	-9125.10	-9113.53	23.13	2	p<0.0001
M7 _(beta) vs. M8 _(beta&ω)	-9131.50	-9113.80	35.40	2	p<0.0001
Rodentia					
M1a _(nearly neutral) vs. M2a _(pos. selection)	-15457.26	-15448.28	17.95	2	p<0.001
M7 _(beta) vs. M8 _(beta&ω)	-15448.11	-15433.70	28.82	2	p<0.0001
Chiroptera					
M1a _(nearly neutral) vs. M2a _(pos. selection)	-8003.66	-7983.35	40.62	2	p<0.0001
M7 _(beta) vs. M8 _(beta&ω)	-8009.67	-7981.22	56.90	2	p<0.0001

Evidence of positive selection was sought by comparing site models that allow $\omega > 1$ (M2a, M8) with models that disallow positive selection (M1a, M7). df, degree of freedom; $\ln L$, log-likelihood scores.