Evolutionary origins of hepatitis A virus in small mammals

Jan Felix Drexler[a,b,1], Victor M. Corman[a,b], Alexander N. Lukashev[c], Judith M. A. van den Brand[d], Anatoly P. Gmyl[c,e], Sebastian Brünink[a], Andrea Rasche[a], Nicole Seggewiß[f], Hui Feng[g], Lonneke M. Leijten[d], Peter Vallo[h], Thijs Kuiken[d], Andreas Dotzauer[f], Rainer G. Ulrich[i], Stanley M. Lemon[g], Christian Drosten[a,b], and the Hepatovirus Ecology Consortium[2]

[a]Institute of Virology, University of Bonn Medical Centre, 53127 Bonn, Germany; [b]German Centre for Infection Research, Bonn-Cologne, Germany; [c]Chumakov Institute of Poliomyelitis and Viral Encephalitides, Moscow 142782, Russia; [d]Department of Viroscience, Erasmus MC, 3000 CA, Rotterdam, The Netherlands; [e]Lomonosov Moscow State University, Moscow 119991, Russia; [f]Laboratory of Virus Research, University of Bremen, 28359 Bremen, Germany; [g]Lineberger Comprehensive Cancer Center, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7292; [h]Institute of Evolutionary Ecology and Conservation Genomics, University of Ulm, 89069 Ulm, Germany; and [i]Friedrich-Loeffler-Institut, Institute for Novel and Emerging Infectious Diseases, 17493 Greifswald–Insel Riems, Germany

Edited by Francis V. Chisari, The Scripps Research Institute, La Jolla, CA, and approved October 2, 2015 (received for review August 26, 2015)

Hepatitis A virus (HAV) is an ancient and ubiquitous human pathogen recovered previously only from primates. The sole species of the genus *Hepatovirus*, existing in both enveloped and nonenveloped forms, and with a capsid structure intermediate between that of insect viruses and mammalian picornaviruses, HAV is enigmatic in its origins. We conducted a targeted search for hepatoviruses in 15,987 specimens collected from 209 small mammal species globally and discovered highly diversified viruses in bats, rodents, hedgehogs, and shrews, which by pairwise sequence distance comprise 13 novel *Hepatovirus* species. Near-complete genomes from nine of these species show conservation of unique hepatovirus features, including predicted internal ribosome entry site structure, a truncated VP4 capsid protein lacking N-terminal myristoylation, a carboxyl-terminal pX extension of VP1, VP2 late domains involved in membrane envelopment, and a *cis*-acting replication element within the 3D^{pol} sequence. Antibodies in some bat sera immunoprecipitated and neutralized human HAV, suggesting conservation of critical antigenic determinants. Limited phylogenetic cosegregation among hepatoviruses and their hosts and recombination patterns are indicative of major hepatovirus host shifts in the past. Ancestral state reconstructions suggest a *Hepatovirus* origin in small insectivorous mammals and a rodent origin of human HAV. Patterns of infection in small mammals mimicked those of human HAV in hepatotropism, fecal shedding, acute nature, and extinction of the virus in a closed host population. The evolutionary conservation of hepatovirus structure and pathogenesis provide novel insight into the origins of HAV and highlight the utility of analyzing animal reservoirs for risk assessment of emerging viruses.

hepatitis A virus | viral evolution | pathogenesis | zoonosis | small mammals

Small mammals such as bats and rodents have been implicated frequently in the evolution and spread of emerging viruses (1). It is uncertain whether this reflects unique aspects of their physiology, immune response to infectious agents, or ecological traits facilitating virus maintenance such as rapid population turnover or tendencies to form large and gregarious social groups (2, 3). The emergence of Ebola virus from bats (4) and hantaviruses from rodents (5) exemplifies the prominent contributions of these taxa to emerging zoonotic threats to human health, but the extent to which such species have contributed to the evolution of well-established human pathogens such as hepatitis A virus (HAV) is less clear.

HAV is unique among the *Picornaviridae*, a large and diverse family of positive-strand RNA viruses (6), not only in its tropism for the liver but also in its structure and life cycle. It infects via the fecal–oral route and is shed in feces as a naked, nonenveloped particle, but circulates in the blood cloaked in an envelope derived from host cell membranes (7). Recent X-ray studies have revealed that HAV possesses a primitive capsid structure related to that of picorna-like viruses infecting insects, hinting at both an ancient evolutionary relationship and a novel mechanism of cell entry (8). The origins of HAV, however, remain shrouded in mystery. Despite evidence of limited replication in guinea pigs (9), only higher primates have been shown to be fully permissive for infection. HAV strains show little variation in nucleotide sequence over time or geography, forming six closely related genotypes comprising only a single serotype (10). Unlike other human hepatitis viruses, HAV infections never persist and uniformly engender lifelong, likely antibody-mediated immunity against reinfection (11). HAV has thus disappeared previously from small, isolated human populations (12, 13), raising questions as to how it could have evolved in early human hunter–gatherer societies.

Results

Identification of Nonprimate Hepatoviruses. To elucidate the evolutionary origins of HAV, we sought the presence of HAV-related viruses in 15,987 specimens (tissue, blood, and feces) collected globally from 209 species of small mammals in five different mammalian orders: Rodentia (rodents), Scandentia (treeshrews), Chiroptera (bats), Eulipotyphla (hedgehogs, shrews), and Afrosoricida (tenrecs; Fig. 1*A*, Fig. S1*A*, and Table S1). Hepatoviruses were

Significance

The origins of human hepatitis A virus (HAV) are unknown. We conducted a targeted search for HAV-related viruses in small mammals sampled globally and discovered highly diversified viruses in bats, rodents, hedgehogs, and shrews. We demonstrate that these viruses share unique biological features with HAV, including structural, genomic, antigenic, and pathogenic properties. We found evidence of major shifts of HAV-related viruses between mammalian hosts in the past, suggesting both an origin of this viral genus in small mammals and a zoonotic origin of human HAV. Our data show that risk assessments for emerging viruses can benefit greatly from the analysis of viral infection patterns that evolved within animal reservoirs.

Author contributions: J.F.D. designed research; J.M.A.v.d.B., A.P.G., S.B., A.R., N.S., H.F., L.M.L., and P.V. performed research; H.E.C. performed and designed field work; R.G.U. contributed new reagents/analytic tools; J.F.D., V.M.C., A.N.L., T.K., A.D., R.G.U., S.M.L., and C.D. analyzed data; and J.F.D., S.M.L., and C.D. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database, www.ncbi.nlm.nih.gov/genbank (accession nos. KT452631–KT452765).

[1]To whom correspondence should be addressed. Email: drexler@virology-bonn.de.

[2]A complete list of the Hepatovirus Ecology Consortium can be found in the *Supporting Information*.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1516992112/-/DCSupplemental.

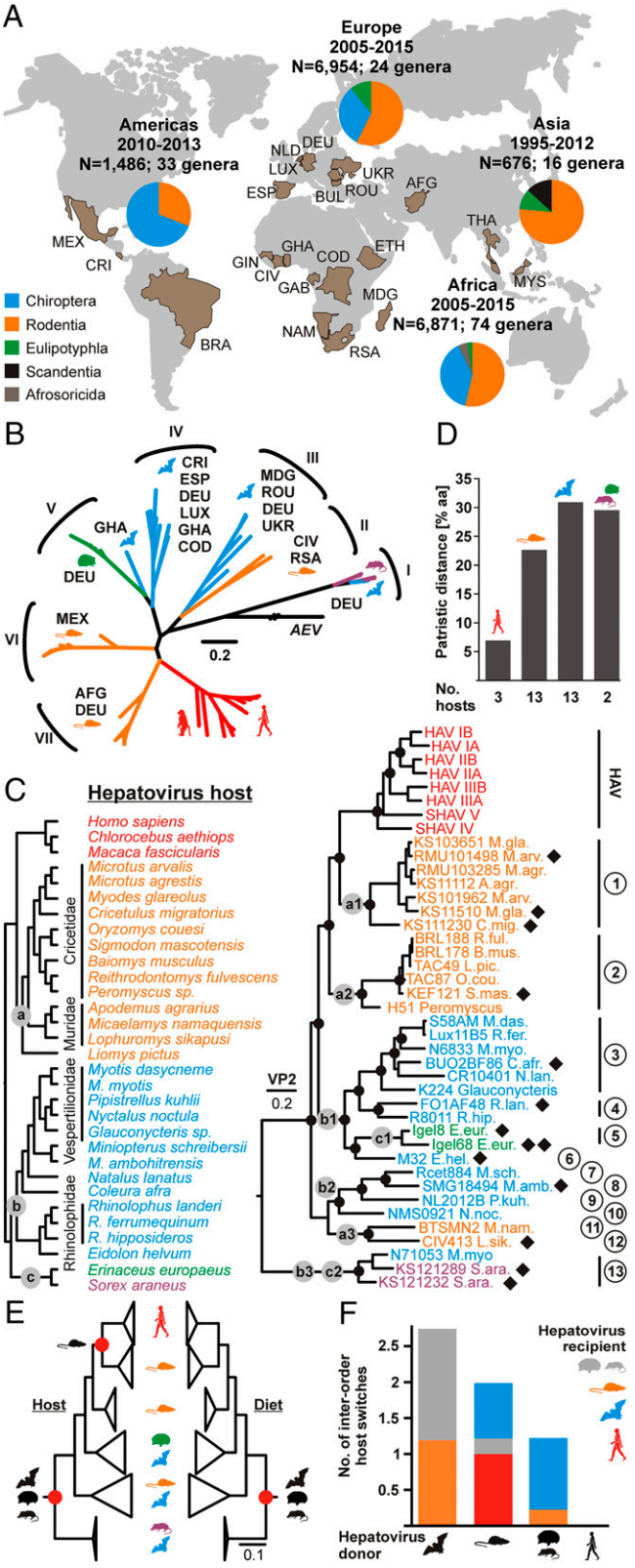

Fig. 1. Hepatovirus evolutionary relationships. (*A*) Number of sampled host genera, specimens, and dates of collection. Country abbreviations, see Table S1. (*B*) Hepatovirus VP2 phylogeny (MrBayes, GTR+G+I nucleotide substitution model). (*C*) Cladogram of hepatovirus hosts (*Left*) and *Hepatovirus* phylogeny (*Right*) as in *B*. Circled numbers, predicted viral species. Diamonds, viral full genome characterizations. Circles at nodes, posterior probabilities >0.9. (*D*) Hepatovirus patristic distance per host order; aa, amino acid. (*E*) Parsimony-based ancestral state reconstructions (ASR) as described previously (3) using 10,000 tree replicates of a VP2 phylogeny (MrBayes, WAG aa substitution model) according to host order (*Left*) or predominantly insectivorous diet (*Right*). (*F*) Averaged number of host switches from ASR shown in *E* originating from and received by each hepatovirus host order.

identified by a broadly reactive nested RT-PCR assay targeting the VP2 domain that is among the most highly conserved segments of the polyprotein-coding RNA. A total of 117 specimens (0.7%) were positive, originating from five continents and a total of 28 different nonprimate hosts, including 13 bat species, 13 rodent species, 1 shrew, and 1 hedgehog species (GenBank accession nos. KT452631–KT452747). Bayesian phylogenetic reconstruction demonstrated considerably more diversity among these novel viruses than primate HAVs, with seven deeply branching clades forming an extended *Hepatovirus* tree (Fig. 1*B* and Fig. S1*B*). Phylogenetically basal clades contained shrew and bat viruses from Central Europe (clade I), African rodent viruses (clade II), Malagasy and European bat viruses (clade III), American, European, and African bat viruses (clade IV), and African and European bat and hedgehog viruses (clade V). Rodent viruses from North America (clade VI), Asia, and Central Europe (clade VII) clustered phylogenetically in an immediate sister relationship to primate hepatoviruses that form an apical monophyletic group. Pairwise sequence distances indicate these novel viruses represent 13 new *Hepatovirus* species (Fig. S1*C*).

Only limited phylogenetic cosegregation was evident for these novel nonprimate hepatoviruses and their mammalian host species. When categorized according to host order (Fig. 1*C*, nodes a–c), viruses from cricetid rodents in Central Europe and Asia (node a1), Northern America (node a2), and murid rodents from Africa (node a3) were not monophyletic. Similarly, viruses from different bat hosts sampled globally occupied three different phylogenetic positions (nodes b1–b3), whereas viruses from Eulipotyphla (hedgehogs and shrews) occupied two different phylogenetic positions (nodes c1–c2). Even at the level of host superorder, viruses from Laurasiatheria (bats, hedgehogs, and shrews), and Euarchontoglires (rodents) were intermixed. In two instances, however, pairs of genetically closely related viruses were identified in the same bat genus in biogeographically distinct regions. One virus pair was obtained from *Rhinolophus* bats from Southern Europe and Africa and a second pair from *Miniopterus* bats in Eastern Europe and Madagascar (Fig. 1*C*). These two virus pairs may suggest some limited degree of ancient host–virus relationships. Finally, closely related viruses were detected in co-occurring rodent hosts belonging to three highly divergent families in Central European and Northern American sampling sites (Fig. 1*C*). Taken together, the phylogenetic relationships evident among nonprimate hepatoviruses strongly suggest multiple host shifts in the past and spillover infections in co-occurring species.

Consistent with the long branches in the hepatovirus phylogenetic tree segregating small mammal viruses, the patristic distance of hepatoviruses was about fourfold greater in small mammals than in primates (Fig. 1*D*). Parsimony-based ancestral state reconstructions (ASRs) (3) consistently projected the primate HAV ancestor to a rodent host (Fig. 1*E*). In contrast, the origin of all mammalian hepatoviruses was consistently projected to a Laurasiatherian host and an insectivorous diet. Bats were most relevant hepatovirus donors for projected host switches, followed by rodents and the Eulipotyphla (Fig. 1*F*). No host switch was projected to involve a primate donor. These results collectively suggest a long and complex evolutionary history of hepatoviruses in small mammals.

Nonprimate Hepatovirus Genome Structure. We determined 14 near-complete genome sequences of nonprimate hepatoviruses representing 9 predicted species (Fig. 1*C*). Each is similar in organization

and length to human HAV, sharing a characteristically low G/C content, avoidance of CpG dinucleotides and a strong codon usage bias (Fig. 2 *A* and *B* and Fig. S1 *D* and *E*) (14). Conserved structural elements in the 5′UTR of most of these viruses include several pyrimidine-rich tracts and a large, predicted cruciform stem-loop resembling the type III HAV internal ribosome entry site (IRES) (Fig. 2*C* and Fig. S2*A*) (15). However, the 5′UTRs of viruses from a Malagasy bat and a West African rodent are predicted to contain a pseudoknot preceding the initiation codon that is characteristic of a type IV IRES (Fig. S2*A*). Because type IV IRES elements typically occur in hepatitis C-related viruses and pestiviruses, this is consistent with recombination events having occurred across viral families, as described previously (16, 17). Irrespective of the type of IRES, each of the genomes contains a series of predicted 5′ terminal pseudoknots similar to those present in human HAV. Notably, each genome also contains a predicted *cis*-acting replication element (*cre*) of variable size but conserved structure in the 3D^{pol} region (18) (Fig. 2*D* and Fig. S2*B*).

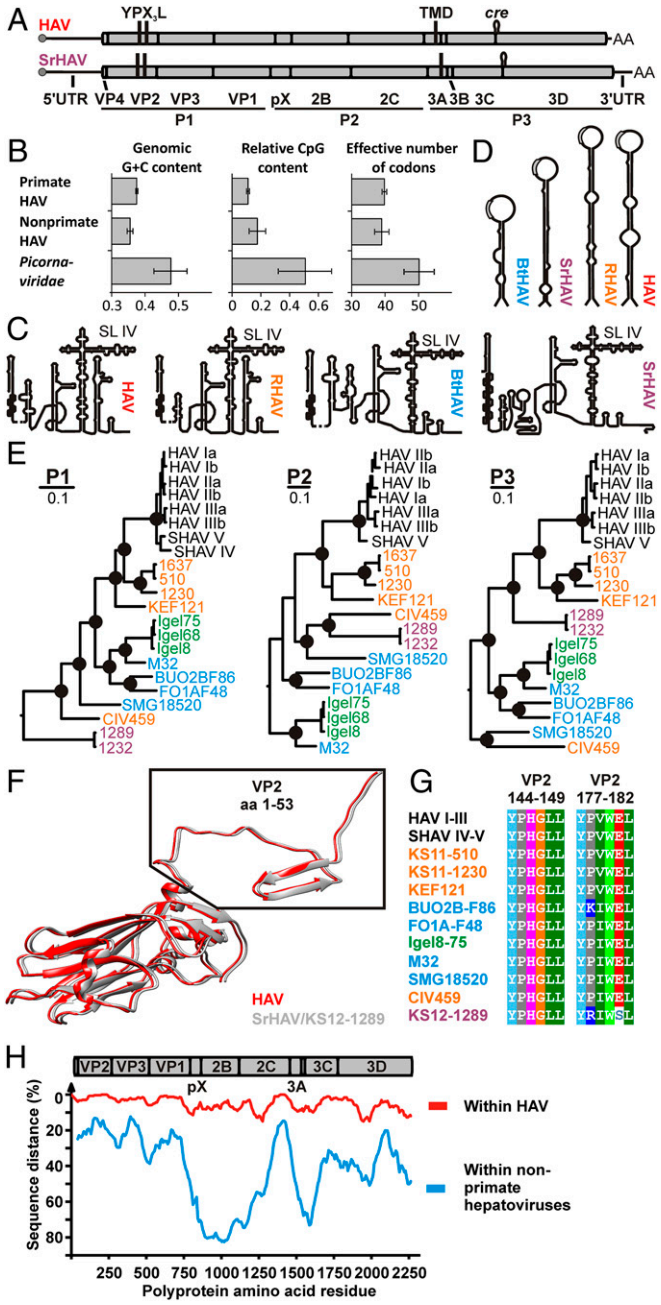

Fig. 2. Properties of the genomes of nonprimate hepatoviruses. (*A*) Hepatovirus genome organization. SrHAV, shrew hepatovirus. (*B*) Genomic features in primate HAV, nonprimate HAV, and the *Picornaviridae*. [Scale bars, mean (SD).] (*C*) IRES folding. RHAV, rodent hepatovirus; BtHAV, bat hepatovirus. (*D*) 3D^{pol} *cre* elements. Gray, conserved AAACG motif. Genomic positions of predicted *cre*: HAV, 5,945–6,055; RHAV KS11-1230, 5,968–6,078; SrHAV KS12-1289, 6,095–6,181; BtHAV, 5,809–5,870. (*E*) Phylogenies of hepatovirus domains P1, P2 (only 2C) and P3 (only 3CD; MrBayes, WAG aa substitution model). Circles at nodes, posterior probabilities >0.9. (*F*) VP2 structure of a SrHAV modeled onto the HAV crystal (8); Box, domain swap. (*G*) VP2 late domains. (*H*) aa sequence distance along hepatovirus polyproteins.

Although these novel viruses were highly diversified (Fig. S2*C*), they demonstrated only 32.4–47.4% distance in the translated polyprotein sequence from human HAV, below the 58% threshold separating *Picornaviridae* genera (6). Similarly, sequence distances in separate comparisons of the P1, P2, and P3 domains were below commonly used thresholds, confirming that all of the novel viruses belong to the genus *Hepatovirus*. Bayesian phylogenies of the P1, P2, and P3 domains [typical picornavirus genomic breakpoints (19)] demonstrated multiple topological incongruences suggestive of past recombination events (Fig. 2*E*). These topological incongruences are exemplified by the clustering of shrew hepatovirus in a basal sister relationship to other hepatoviruses in P1, but in an apical position together with rodent viruses in P3. These results suggest genetic compatibility of diverse structural and nonstructural elements from hepatoviruses infecting different mammalian orders and emphasize the broad host tropism of hepatoviruses suggested by ASR. Consistent with their phylogenetic clustering with human HAV, hepatoviruses from cricetid rodents are predicted to contain the most closely related IRES and *cre* elements, as well as 3A transmembrane domains (Fig. S2*D*).

Conservation of Unique Hepatovirus Features. Features common to each of the nonprimate viruses and characteristic of hepatoviruses include the predicted absence of a leader protein, a small VP4 (16-26 amino acids) lacking an N-terminal myristoylation signal, an approximate 3- to 9-kDa C-terminal pX extension on VP1, and the absence of a 2A protein. Thermodynamic modeling of the VP2 structure suggests it is closely related to that of HAV, which contains a unique domain swap found in insect *Dicistroviridae* but not in mammalian picornaviruses (8) (Fig. 2*F* and Fig. S3 *A* and *B*). YPX$_3$L "late domain" motifs in VP2 (7) that contribute to HAV membrane envelopment by mediating capsid interactions with components of the endosomal sorting complex required for transport (ESCRT) are highly conserved among nonprimate hepatoviruses (Fig. 2*G*).

Further comparisons of the polyprotein sequences of these novel hepatoviruses revealed several conserved domains in VP2 and VP3, suggesting the possibility of antigenic relatedness, whereas there was extreme variation across the pX-2B genomic region (Fig. 2*H*). Protein BLAST comparisons of these highly divergent genomic regions revealed similarities with diverse nonviral elements distantly related to various microbial organisms (Fig. S3*C*). Thus, the hepatovirus pX-2B domain may be a genomic hotspot potentially associated with host adaptation and prone to acquisition of exogenous sequence elements. These hypotheses are consistent with experimental data showing that 2B harbors key mutations associated with cell culture adaptation of human HAV and that the pX domain tolerates large deletions and exogenous sequence insertions (20, 21).

To determine whether the conserved antigenicity evident among primate HAVs extends to the nonprimate viruses, we sought evidence for antibodies recognizing human HAV in sera from 111 bats (10 different species), 114 rodents (7 species, including several samples PCR-positive for the most closely related hepatovirus lineage), 103 shrews (3 species), and 48 hedgehogs

MICROBIOLOGY

(Table S2). Sera from bats, but no other animals, were reactive in immunofluorescence assays (IFAs) of HAV-infected cells at 1:40–1:400 endpoint dilutions, with fluorescence patterns that closely resembled those of a monoclonal antibody control serum (Fig. 3*A*). Of eight positive sera (7.3%), six were from West African *Eidolon helvum*, and one each from Central African *Rousettus aegyptiacus* and *Micropteropus pusillus*. To confirm these results, we assessed the ability of *Eidolon* sera to immunoprecipitate (IP) human HAV. Four of the six IFA-positive sera that were available in sufficient volumes were strongly reactive in this assay, some exceeding the precipitating activity of anti-HAV reference sera (Fig. 3*B*). These four sera also effectively neutralized human HAV infectivity (Table S3). These results hint at significant conservation of capsid antigenicity between the *Eidolon* hepatovirus lineage and human HAV and are consistent with conservation of the sequences of several neutralization epitopes located in the capsid proteins VP3, VP2, and VP1 (Fig. 3*C* and Fig. S4 *A* and *B*). An alternative explanation for the presence of such antibodies in African bats could be exposure to human HAV, perhaps by ingestion of contaminated surface water in highly endemic areas (22). The unique ecologic traits of bats including their longevity and spatial mobility compared with other small mammals may facilitate repeated encounters with human HAV in endemic areas. However, it seems unlikely that such an exposure would be of sufficient magnitude to evoke antibody responses without replication of the virus.

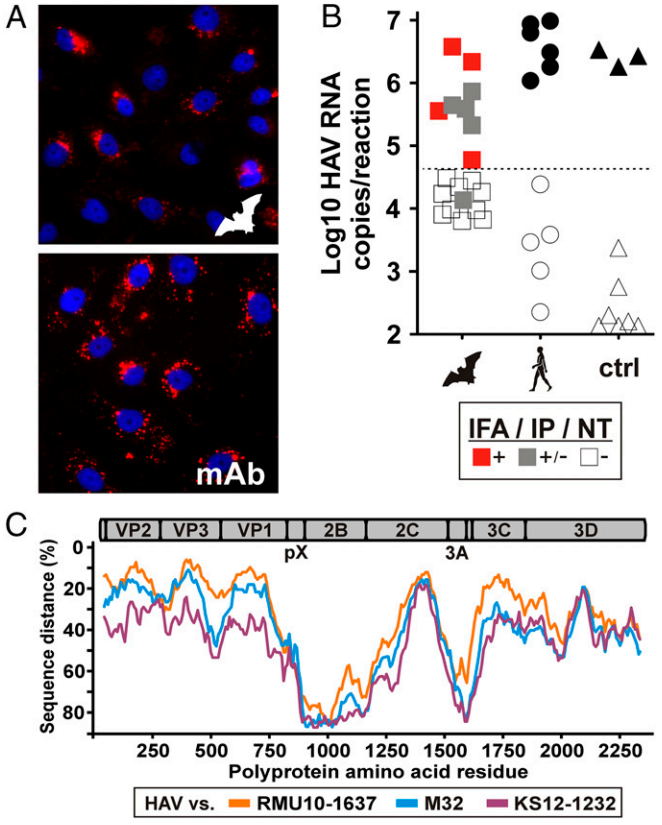

Fig. 3. Antigenic relatedness of human and nonprimate hepatoviruses. (*A*) (*Upper*) immunofluorescence assay (IFA) showing a bat serum reacting with human HAV-infected FRhK-4 cells [red (Cy2); mixed with 50% non-infected cells as internal negative controls]. (*Lower*) Same cells stained with a monoclonal antibody control (mAb 7E7, 100% infected cells). Blue (DAPI), nuclei. (*B*) Immunoprecipitation (IP) of HAV by bat sera [IF, IP, neutralization test (NT); red, positive in all assays; gray, discordant assay results; empty, negative in all assays], human sera, and controls. See *SI Materials and Methods* for details on control sera. Dotted line, threshold precipitation separating positive and negative control sera. (*C*) aa sequence distance between HAV (genotype Ia, GenBank accession no. AB020564), the *Eidolon* BtHAV M32, a RHAV (RMU10-1637), and a SrHAV (KS12-1232).

Because of the genetic relatedness between rodent hepatoviruses and human HAV, the absence of cross-reactive antibodies in rodent sera was surprising. It may be possible that rodents differ from other small mammal hosts in the magnitude and onset of their antibody response after hepatovirus infection, which would be consistent with lack of antibodies in guinea pigs up to 60 d after experimental infection with HAV (9). Alternative explanations may include subtle differences in complex epitopes located in the hepatovirus antigenic sites. Resurrection or isolation of rodent viruses for infection studies and development of hepatovirus-specific serologic assays will aid comparative investigations in further studies.

Pathogenesis and Ecology of Nonprimate Hepatoviruses. Hepatovirus infection patterns in bats, rodents, hedgehogs, and shrews were similar to those of HAV in primates. Viral RNA abundance was highest in liver vs. other solid organs and blood in rodents, shrews, and hedgehogs. In bats, liver, and several extrahepatic tissues, including predominantly spleen, lung, and intestine, showed high viral RNA concentrations (Fig. 4*A*). Viral RNA was identified by in situ hybridization in bat hepatocytes and in large mononuclear cells within the germinal center of splenic lymphatic nodules (Fig. 4*B*). Comparable data are not available for primate hepatoviruses, although HAV antigen has been identified in splenic macrophages (23). Minus-strand RNA (a replicative intermediate) was detected in the liver and spleen of bats, but only in the liver of hedgehogs (Fig. S4*C*). We conclude that these novel viruses are hepatotropic, but may also replicate within the spleen in bats. Liver tissues from 23 PCR-positive and 49 PCR-negative bats, rodents, shrews, and hedgehogs were examined histologically. In some PCR-positive animals we observed a mild increase in periportal inflammatory mononuclear cell infiltrates, but there was no consistent difference with control, PCR-negative animals, and no severe liver pathology.

We next assessed infection outcome in 24 hedgehogs that were serially monitored over a 220-d period in a German animal shelter. Nine of these animals tested positive for hepatoviruses in up to three separate specimens. The viruses identified in these animals were 100% identical in their respective VP2 sequences, consistent with the very low nucleotide substitution rate of human HAV (10) (Fig. S4*D*). Continuous fecal shedding of virus (~10^6 RNA copies per gram feces) was observed over 79–142 d in four animals, resembling HAV shedding in humans (24) (Fig. 4*C*). Importantly, in consecutive specimens taken less than 15 d apart, the second specimen consistently showed a significant decrease in virus concentration, suggesting resolving acute infections (Fig. 4*C*). Four animals completely cleared the infection during the study period. When released, all nine infected hedgehogs were in good physical condition, corroborating low hepatovirus pathogenicity in these small mammals.

Finally, to gain insight into the epizootic nature of nonprimate hepatoviruses, we analyzed specimens obtained over a 3-y period in a German bat maternity roost within which females annually form a colony lasting for approximately 3 mo, giving birth to offspring in a synchronized fashion (25). Bat hepatovirus was detected in fecal droppings immediately on colony formation in early May 2008 (Fig. 4*D* and clade IV in Fig. 1*B*). The detection rate decreased significantly 1 mo later, only to rise again after parturition in mid-July. Shortly before bats left the colony toward the end of that month, the detection rate fell for the second time. These data suggest an absence of persistent infections in the majority of animals, followed by a new wave of acute infections in offspring, and resemble HAV outbreaks in children's camps in the prevaccine era (26). Resampling of the same maternity roost (containing largely the same individuals) 2 y later failed to identify

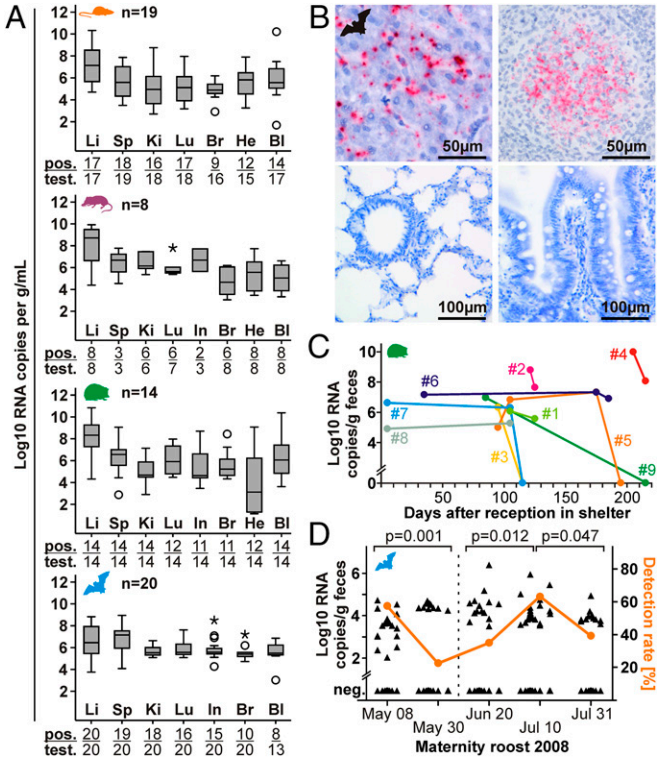

Fig. 4. Hepatovirus infection patterns in small mammals. (*A*) Hepatovirus RNA abundance in solid organs and blood per host taxon determined by qPCR. Li, liver; Sp, spleen; Ki, kidney; Lu, lung; In, intestine; Br, brain; He, heart; Bl, blood. Boxes, median, interquartile range; outliers (circles); extreme values (asterisks). Next to host pictograms, numbers of analyzed animals; Below each organ, ratios of specimens testing positive per specimens available for testing. (*B*) In situ hybridization of BtHAV RNA in *E. helvum* liver (*Upper Left*, 400× magnification, animal M32), spleen (*Upper Right*, 400×, animal GH297), lung (*Lower Left*, 200×, animal M32), and intestine (*Lower Right*, 200×, animal M32). (*C*) Hepatovirus detections and RNA concentrations determined by qPCR in nine hedgehogs sampled longitudinally. (*D*) Hepatovirus amplification in a bat maternity roost. Triangles, viral RNA concentrations in pooled fecal specimens over five sampling points. Secondary *y* axis and orange line, detection rate.

the same hepatovirus lineage, suggesting its extinction in this population. Instead, a highly divergent hepatovirus which was genetically related to viruses from shrews (clade I in Fig. 1*B*) was identified in six specimens (Fig. S4*E*), illustrating that diverse hepatoviruses can be carried by a single host species. The level of amino acid sequence divergence between the two hepatovirus lineages infecting the roost in 2008 and 2010 exceeded 30% in the partial VP2 domain (Fig. S4*E*), which is beyond the distance separating enterovirus serotypes in this genomic region (27), thus hinting at the existence of divergent serotypes in the genus *Hepatovirus*.

Discussion

We identified multiple, highly diversified hepatoviruses in numerous small mammal taxa, extending the host associations of the genus *Hepatovirus* well beyond primates. Our findings render this picornavirus genus exceptionally speciose, comparable only to the genus *Enterovirus* of the family *Picornaviridae* after decades of investigation (6).

The unique properties of human HAV that are shared by these novel nonprimate hepatoviruses and that distinguish it from other mammalian picornaviruses likely reflect those of ancestral viruses infecting small mammals before formation of the primate hepatovirus lineage. Whether the putative hepatovirus introduction took place in the primate stem lineage preceding the split of Hominoidea and Cercopithecoidea about 25 Mya (28) remains unknown because of the scarcity of HAV strains recovered from nonhuman primates. The survival of hepatoviruses before their introduction into primates was likely mediated by large population sizes and/or high population turnover of small mammal hosts (1, 3, 5). On the virus side, an unusually broad host range and genetic plasticity is likely to have contributed further to hepatovirus maintenance and evolution.

The existence of evolutionarily ancestral hepatoviruses in bats and shrews compared with the presence of more closely related viruses in rodents and primates is reminiscent of hantavirus host associations, in which pathogenic human viruses originate from rodents, whereas ancestral viruses occur in bats and Eulipotyphla (29). The relevance of these Laurasiatherian hosts for the evolutionary origins of human hepatitis viruses is demonstrated by the recent detections of ancestral hepatitis B, C, and E viruses in bats (30–32). It remains to be determined whether Laurasiatheria generally harbor a wider genetic diversity of viruses than Euarchontoglires and whether ecological traits such as insectivorous diets influence viral diversity. However, reconstructions that point to a Laurasiatherian host and an insectivorous diet for ancestral hepatoviruses provide a novel link to the structural phylogeny of the HAV capsid and its close relationship to picorna-like viruses of insects (8) and together suggest more distant ancestry in a primordial insect-borne virus. Such a scenario is paralleled by recent suggestions of similar ancestry for other mammalian viruses (33–35) and provides a new perspective on the origins of this ancient human pathogen.

A major barrier to viral host switches is receptor use (36). Human HAV uses the T-cell Ig and mucin domain 1 (TIM-1) for cell entry (37). It will be important to investigate whether receptor use is conserved among the nonprimate hepatoviruses we identified and to determine whether these viruses are capable of infecting primate hosts. Of note, human TIM-1 is used by other zoonotic viruses for cell entry, including the bat-borne Ebola virus, the arthropod-borne dengue virus and several rodent- and bat-borne New World arenaviruses (38–40), illustrating that dependency on this molecule for cell entry may not represent an insuperable barrier to a viral host switch.

Zoonotic infections with emerging viruses have become increasingly relevant for human health due to the invasion of pristine habitats by humans and their livestock, advancing global mobility, and the rapid spread of pathogens within dense human populations (41–43). The antigenic conservation we observed between HAV and nonprimate hepatoviruses from *Eidolon* bats suggests these viruses belong to a common serotype and that the introduction of such viruses into human populations would be limited by herd immunity and vaccine-induced immune responses. However, this may not apply to all of the novel hepatovirus species we identified, some of which on the basis of diversity within the P1 domain likely comprise distinct serotypes. Our study exemplifies the utility of looking beyond phylogenetic criteria alone when conducting risk assessment for emerging RNA viruses and the need to include functional, ecologic, and pathogenic analyses of animal reservoirs.

Materials and Methods

Sampling was done as described previously (3). Hepatovirus detection, quantification, and genomic characterizations were done by PCR-based techniques (see oligonucleotide sequences in Tables S4 and S5). Serologic and histopathologic analyses were done as described previously (17). Details of these and evolutionary analyses are given in *SI Materials and Methods*.

ACKNOWLEDGMENTS. We thank Monika Eschbach-Bludau, Tobias Bleicker, Andres Moreira Soto, Breno Dominguez Souza, Leonardo Calderón Obaldía, Debby van Riel, Marcel A. Müller, Jens P. Teifke and Kevin L. McKnight for assistance and advice; Tabea Binger, Heather J. Baldwin, Augustina Annan, Gael Darren Maganga, Peter H.C. Lina, Markus Dietz, Jacques B. Pir, Heike Philipps, Gertraude Göpner, Manfred Göpner, Lena Grosche, Frauke Meier, Myriam Götz, Ioan Coroiu, Regina Klüppel-Hellmann, Anda Culisier, Mircea-Dan Harlav, Sabrina Stölting, Carsten Dense, Dieudonné Nkoghe, Mathieu Bourgarel, Stoian Yordanov, Tasnim Suleiman, Ndapewa Ithete, Thomas Jäkel, Thomas Kruppa, Stephan Günther, Koussay Dellagi, Hervé Pascalis, Julien Mélade, Pablo Tortosa, Peter Witkowski, Georg Herrler, Jörg Thiel,

Margrit Bemmann, Katarina Kühn, Engelbert Kampling, Mechthild Budde, Dagmar Funck, Jona Freise, Brita Oltmann, Sebastian Guenther, Brigitte Bannert, Matthias Wenk, Henrike Gregersen, Johannes Lang, Anita Plenge-Bönig, Nicole Schauerte, Susanne Modrow, Gerhard Maluck, Richard Kruczewski, Julie Lichière, Maria Nemeth, Anne Balkema-Buschmann, Susanne Jäckel, Horst Schirrmeier, Hanan Sheikh Ali, Christian Kretzschmar, Konrad Wanka, Ute Wessels, Theres Wollny, Kathrin Baumann, Grit Möwert, Franziska Thomas, Bärbel Hammerschmidt, Daniel Windolph, Josephine Schlosser, Marc Mertens, Paul Dremsek, Katja Plifke, Kerstin Tauscher, Angele Breithaupt, Nadja Lorenz, Christian Korthase, André Schütte, Julie Elkins, Sylvia Ferguson, Christiane Denys, Emilie Lecompte, René Ryll, Christin Trapp, Anke Mandelkow, Nicole Schmidt, Annemarie Steiner, Maxi Zumpe, Samuel Bernstein, Stefan Fischer, Sandra Blome, Denny Maaz, Christoph Gertler, Michael Faulde, Jens Lewitzki, Mathias Schlegel, Julia Zeitler, Wolfgang Fiedler, Elisabeth Glatthaar, and Sabine Sauer for assistance with field work and necropsy; and Mediagnost for the mAb 7E7. Samples from small mammal monitoring in Germany were kindly provided by Sabrina Schmidt, Ulrike M. Rosenfeld, Daniela Reil, Christian Imholt, Nastasja Kratzmann, Stephan Drewes, Stefan Fischer, and Jens Jacob from UFOPLAN Projects 370941401 and 371348401. Several samples for this study were taken during field work for projects granted to C.D. by the German Federal Ministry of Education and Research (BMBF) (Projects 01KI1005A and 01KI1016D), the German Research Foundation (DFG) (Projects DR 772/3-1, DR 772/12-1, and DR 772/10-1), and to R.G.U. by the German Armed forces (M/SABX/005 and E/U2AD/CF512/DF557) and the BMBF (Projects 01KI1018 and 01KI1303). Additional funding was obtained by the European Commission through the COMPARE project (to C.D.) (643476), Russian Scientific Foundation Grant 14-15-00619 (to A.N.L.), German Tönjes-Vagt-Stiftung Project XXIX (to A.D.), National Institutes of Health Grants R01-AI103083 and U19-AI109965 (to S.M.L.), and the US Department of Defense (HR0011-13-2-0020) and the DFG (DR 810/1-1) (to J.F.D.).

1. Luis AD, et al. (2013) A comparison of bats and rodents as reservoirs of zoonotic viruses: Are bats special? *Proc R Soc Biol Sci* 280(1756):20122753.
2. Brook CE, Dobson AP (2015) Bats as 'special' reservoirs for emerging zoonotic pathogens. *Trends Microbiol* 23(3):172–180.
3. Drexler JF, et al. (2012) Bats host major mammalian paramyxoviruses. *Nat Commun* 3:796.
4. Leroy EM, et al. (2005) Fruit bats as reservoirs of Ebola virus. *Nature* 438(7068): 575–576.
5. Han BA, Schmidt JP, Bowden SE, Drake JM (2015) Rodent reservoirs of future zoonotic diseases. *Proc Natl Acad Sci USA* 112(22):7039–7044.
6. Knowles NJ, et al. (2012) *Family Picornaviridae. Virus taxonomy. Ninth Report of the International Committee on Taxonomy of Viruses,* eds King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (Elsevier/Academic Press, Amsterdam), pp 855–880.
7. Feng Z, et al. (2013) A pathogenic picornavirus acquires an envelope by hijacking cellular membranes. *Nature* 496(7445):367–371.
8. Wang X, et al. (2015) Hepatitis A virus and the origins of picornaviruses. *Nature* 517(7532):85–88.
9. Hornei B, et al. (2001) Experimental hepatitis A virus infection in guinea pigs. *J Med Virol* 64(4):402–409.
10. Cristina J, Costa-Mattioli M (2007) Genetic variability and molecular evolution of hepatitis A virus. *Virus Res* 127(2):151–157.
11. Walker CM, Feng Z, Lemon SM (2015) Reassessing immune control of hepatitis A virus. *Curr Opin Virol* 11:7–13.
12. Wong DC, Purcell RH, Rosen L (1979) Prevalence of antibody to hepatitis A and hepatitis B viruses in selected populations of the South Pacific. *Am J Epidemiol* 110(3): 227–236.
13. Skinhoj P, Mikkelsen F, Hollinger FB (1977) Hepatitis Ain Greenland: Importance of specific antibody testing in epidemiologic surveillance. *Am J Epidemiol* 105(2):140–147.
14. Aragonès L, Guix S, Ribes E, Bosch A, Pintó RM (2010) Fine-tuning translation kinetics selection as the driving force of codon usage bias in the hepatitis A virus capsid. *PLoS Pathog* 6(3):e1000797.
15. Brown EA, Day SP, Jansen RW, Lemon SM (1991) The 5′ nontranslated region of hepatitis A virus RNA: Secondary structure and elements required for translation in vitro. *J Virol* 65(11):5828–5838.
16. Hellen CU, de Breyne S (2007) A distinct group of hepacivirus/pestivirus-like internal ribosomal entry sites in members of diverse picornavirus genera: Evidence for modular exchange of functional noncoding RNA elements by recombination. *J Virol* 81(11):5850–5863.
17. Drexler JF, et al. (2013) Evidence for novel hepaciviruses in rodents. *PLoS Pathog* 9(6): e1003438.
18. Yang Y, Yi M, Evans DJ, Simmonds P, Lemon SM (2008) Identification of a conserved RNA replication element (cre) within the 3Dpol-coding sequence of hepatoviruses. *J Virol* 82(20):10118–10128.
19. Lukashev AN (2010) Recombination among picornaviruses. *Rev Med Virol* 20(5): 327–337.
20. Beard MR, Cohen L, Lemon SM, Martin A (2001) Characterization of recombinant hepatitis A virus genomes containing exogenous sequences at the 2A/2B junction. *J Virol* 75(3):1414–1426.
21. Emerson SU, Huang YK, McRill C, Lewis M, Purcell RH (1992) Mutations in both the 2B and 2C genes of hepatitis A virus are involved in adaptation to growth in cell culture. *J Virol* 66(2):650–654.
22. Esona MD, et al. (2010) Reassortant group A rotavirus from straw-colored fruit bat (Eidolon helvum). *Emerg Infect Dis* 16(12):1844–1852.
23. Mathiesen LR, et al. (1978) Localization of hepatitis A antigen in marmoset organs during acute infection with hepatitis A virus. *J Infect Dis* 138(3):369–377.
24. Yotsuyanagi H, et al. (1996) Prolonged fecal excretion of hepatitis A virus in adult patients with hepatitis A as determined by polymerase chain reaction. *Hepatology* 24(1):10–13.
25. Drexler JF, et al. (2011) Amplification of emerging viruses in a bat colony. *Emerg Infect Dis* 17(3):449–456.
26. Werzberger A, et al. (1992) A controlled trial of a formalin-inactivated hepatitis A vaccine in healthy children. *N Engl J Med* 327(7):453–457.
27. Nasri D, et al. (2007) Typing of human enterovirus by partial sequencing of VP2. *J Clin Microbiol* 45(8):2370–2379.
28. Springer MS, et al. (2012) Macroevolutionary dynamics and historical biogeography of primate diversification inferred from a species supermatrix. *PLoS One* 7(11):e49521.
29. Guo WP, et al. (2013) Phylogeny and origins of hantaviruses harbored by bats, insectivores, and rodents. *PLoS Pathog* 9(2):e1003159.
30. Drexler JF, et al. (2013) Bats carry pathogenic hepadnaviruses antigenically related to hepatitis B virus and capable of infecting human hepatocytes. *Proc Natl Acad Sci USA* 110(40):16151–16156.
31. Drexler JF, et al. (2012) Bats worldwide carry hepatitis E virus-related viruses that form a putative novel genus within the family Hepeviridae. *J Virol* 86(17):9134–9147.
32. Quan PL, et al. (2013) Bats are a major natural reservoir for hepaciviruses and pegiviruses. *Proc Natl Acad Sci USA* 110(20):8194–8199.
33. Zirkel F, et al. (2011) An insect nidovirus emerging from a primary tropical rainforest. *MBio* 2(3):e00077–e11.
34. Marklewitz M, Zirkel F, Kurth A, Drosten C, Junglen S (2015) Evolutionary and phenotypic analysis of live virus isolates suggests arthropod origin of a pathogenic RNA virus family. *Proc Natl Acad Sci USA* 112(24):7536–7541.
35. Li CX, et al. (2015) Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses. *eLife* 4:4.
36. Drosten C (2013) Virus ecology: A gap between detection and prediction. *Emerg Microbes Infect* 2(5):e31.
37. Feigelstock D, Thompson P, Mattoo P, Zhang Y, Kaplan GG (1998) The human homolog of HAVcr-1 codes for a hepatitis A virus cellular receptor. *J Virol* 72(8):6621–6628.
38. Kondratowicz AS, et al. (2011) T-cell immunoglobulin and mucin domain 1 (TIM-1) is a receptor for Zaire Ebolavirus and Lake Victoria Marburgvirus. *Proc Natl Acad Sci USA* 108(20):8426–8431.
39. Meertens L, et al. (2012) The TIM and TAM families of phosphatidylserine receptors mediate dengue virus entry. *Cell Host Microbe* 12(4):544–557.
40. Jemielity S, et al. (2013) TIM-family proteins promote infection of multiple enveloped viruses through virion-associated phosphatidylserine. *PLoS Pathog* 9(3):e1003232.
41. Brockmann D, Helbing D (2013) The hidden geometry of complex, network-driven contagion phenomena. *Science* 342(6164):1337–1342.
42. Plowright RK, et al. (2015) Ecological dynamics of emerging bat virus spillover. *Proc R Soc Biol Sci* 282(1798):20142124.
43. Pigott DM, et al. (2014) Mapping the zoonotic niche of Ebola virus disease in Africa. *eLife* 3:e04395.
44. Corman VM, et al. (2014) Characterization of a novel betacoronavirus related to middle East respiratory syndrome coronavirus in European hedgehogs. *J Virol* 88(1): 717–724.
45. Alcaide M, et al. (2009) Disentangling vector-borne transmission networks: A universal DNA barcoding method to identify vertebrate hosts from arthropod bloodmeals. *PLoS One* 4(9):e7092.
46. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19(12):1572–1574.
47. Maurer-Stroh S, Eisenhaber B, Eisenhaber F (2002) N-terminal N-myristoylation of proteins: Prediction of substrate proteins from amino acid sequence. *J Mol Biol* 317(4): 541–557.
48. Robert X, Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* 42(Web Server issue):W320–W324.
49. Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 31(13):3406–3415.
50. Simmonds P (2012) SSE: A nucleotide and amino acid sequence analysis platform. *BMC Res Notes* 5:50.
51. Tamura K, et al. (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28(10):2731–2739.
52. Hofmann K, Stoffel W (1993) TMbase: A database of membrane spanning protein segments. *Biol Chem Hoppe Seyler*, 10.1515/bchm3.1993.374.1-6.143.
53. Müller MA, et al. (2007) Coronavirus antibodies in African bat species. *Emerg Infect Dis* 13(9):1367–1370.
54. Houde A, et al. (2007) Comparative evaluation of new TaqMan real-time assays for the detection of hepatitis A virus. *J Virol Methods* 140(1-2):80–89.
55. Dotzauer A, et al. (2000) Hepatitis A virus-specific immunoglobulin A mediates infection of hepatocytes with hepatitis A virus via the asialoglycoprotein receptor. *J Virol* 74(23):10950–10957.
56. Bininda-Emonds OR, et al. (2007) The delayed rise of present-day mammals. *Nature* 446(7135):507–512.

Supporting Information

Drexler et al. 10.1073/pnas.1516992112

The Hepatovirus Ecology Consortium

Chantal G. Akoua-Koffi
(Department of Epidemic Viruses, Institut Pasteur de Cote d'Ivoire, Bouaké, Cote d'Ivoire)

Ramón Seage Ameneiros
(Institute of Evolutionary Ecology and Conservation Genomics, University of Ulm, Ulm, Germany; Group Morcegos de Galicia, Drosera Society)

Steven M. Goodman
(Field Museum of Natural History, Chicago, Illinois, USA)

Elena V. Godlevska
(Schmalhausen Institute of Zoology, Kiev, Ukraine)

Guy-Crispin Gembu Tungaluna
(Université de Kisangani, Faculté des Sciences, Kisangani, Democratic Republic of the Congo)

Tamás Görföl
(Hungarian Natural History Museum, Hungary)

Alexandre Hassanin
(Institut de systématique, évolution, biodiversité, ISYEB–UMR 7205 CNRS, MNHN, Université Paris-6 (UPMC), Sorbonne Universités, Paris, France)

Lamine Koivogui
(Institut National de Santé Publique, Conakry, Guinea)

Detlev H. Krüger
(Institute of Medical Virology (Helmut Ruska Haus), Charité Medical School, Berlin, Germany)

Maklarin B. Lakim
(Sabah Parks, Kota Kinabalu, Sabah, Malaysia)

Leonid Lavrenchenko
(Department of Mammalian Microevolution, A. N. Severtsov Institute of Ecology and Evolution, Russian Academy of Sciences, Moscow, Russia)

Eric M. Leroy
(Centre International de Recherches Médicales de Franceville, Franceville, Gabon)

Sonja Matthee
(Department of Conservation Ecology and Entomology, Stellenbosch University, Stellenbosch, South Africa)

Jan ter Meulen
(Institute of Virology, Philipps University Marburg, Germany)

Yaw Adu-Sarkodie
(Kwame Nkrumah University of Science and Technology, Kumasi, Ghana)

Samuel Oppong
(Kwame Nkrumah University of Science and Technology, Kumasi, Ghana)

Wolfgang Preiser
(Department of Pathology, Faculty of Medicine & Health Sciences Stellenbosch University and NHLS Tygerberg, Tygerberg, South Africa)

Maria João Ramos Pereira
(Department of Zoology Federal University of Rio Grande do Sul, Proto Alegre, Brazil; Centre for Environmental and Marine Studies Wildlife Research Unit University of Aveiro, Aveiro, Portugal)

Chantal B. E. M. Reusken
(Netherlands Center for Infectious Disease Control, Bilthoven, the Netherlands)

Julia Schneider
(Friedrich-Loeffler-Institut, Institute for Novel and Emerging Infectious Diseases, Greifswald–Insel Riems, Germany)

Jonas Schmidt-Chanasit
(Bernhard Nocht Institute for Tropical Medicine, Department of Virology, Hamburg, Germany)

Elisabeth Fichet-Calvet
(Bernhard Nocht Institute for Tropical Medicine, Department of Virology, Hamburg, Germany)

Antje Seebens-Hoyer
(Noctalis, Centre for Bat Protection and Information, Bad Segeberg, Germany)

Florian Gloza-Rausch
(Noctalis, Centre for Bat Protection and Information, Bad Segeberg, Germany)

Alvaro Aguilar Setién
Unidad de Investigación Médica en Inmunología, Hospital de Pediatría Centro Médico Nacional Siglo XXI (IMSS), México DF)

Marco Tschapka
(Institute of Evolutionary Ecology and Conservation Genomics, University of Ulm, Ulm, Germany)

Konstans Wells
(Environmental Futures Research Institute, Griffith University, Brisbane, Australia)

Tom Wegner
(LIMES-Institute Life&Medical Sciences, University of Bonn, Bonn, Germany)

David A. Wilkinson
(Centre de Recherche et de Veille sur les Maladies émergentes dans l'Océan Indien (CRVOI), Plateforme de Recherche CYROI, La Reunion, France).

SI Materials and Methods

Sampling. All animals were handled according to national and European legislation, namely the EU Council directive 86/609/EEC for the protection of animals. For all sampling sites, study protocols including capture, sampling, and testing of animals were approved by the responsible animal ethics committees. All efforts were made to minimize suffering of animals. Any surgical procedure was performed under sodium pentobarbital/ketamine anesthesia. Trapping of rodents and shrews in Germany was conducted in the framework of hantavirus monitoring activities coordinated by the Friedrich-Loeffler-Institute, the Federal Research Institute for Animal Health, as described previously (17); trapping of rodents in other European, African, and Asian countries was described previously (17). Sampling of hedgehogs was done in a German shelter as described previously (44). Sampling of bats was done using mist nets or harp traps as described previously (3). Longitudinal sampling in a bat maternity roost was done as described previously (25).

RNA Purification. Viral RNA was purified from ~30 mg of tissue from solid organs using the RNEasy Kit (Qiagen) as described previously (3) or the MagNA Pure 96 DNA and Viral NA Small Volume Kit (Roche) from blood (input volumes varying between 10 and 50 μL depending on the available volume) or 50 mg feces suspended in an RNAlater (Qiagen) solution as described previously (17).

Hepatovirus Detection, Quantification, and Genomic Characterization. Hepatovirus RNA was detected in a hemi-nested RT-PCR with primers designed under consideration of all divergent HAV sequences available in GenBank and the genetically related tremovirus *Avian encephalomyelitis virus* (AEV). Oligonucleotide primer sequences were HAV-F1089, GAGATAYCAYACWTATGCIAGATTTGG; HAV-R1481, CTRAATTCRTTICTCATCATYTGTG; HAV-R1544, GACATYTTIGCYCTIGCATCYTC. Screening fragments were extended toward the 3′ ends by using consensus reverse oligonucleotides designed to work for all hepatoviruses (HAV-R2009, ACYCTRTANGGTGTRTCWSANATCCANGG; HAV-R4415, CCATANANRTARCANACAACWGGYTCA) in a cDNA synthesis using the SuperScript III reverse transcriptase (Life Technologies) at 45 °C during 1 h (Y = C/T, W = A/T, I = inosine, R = G/A, S = G/C, N = A/C/T/G), with inner gene-specific oligonucleotide forward primers designed after Sanger sequencing of the screening PCR amplicon. The final VP2 dataset used for all phylogenetic analyses comprised 864 nucleotides corresponding to genomic positions 1,123–1,986 in the HAV prototype strain HM175/p16 (GenBank accession no. KP879217). The screening PCR first round used the SuperScript III OneStep RT-PCR kit (Life Technologies) in a 25-μL reaction volume, with 600 nmol of respective forward and reverse primers, 800 μmol of $MgSO_4$, 1× reaction buffer, 1 μg of PCR-grade BSA, 1 μL enzyme mix, and 5 μL RNA extract. The thermal cycling profile involved 20 min at 50 °C for reverse transcription, followed by 3 min at 95 °C, 10 cycles of 15 s at 94 °C, 15 s at 60 °C (with a decrease of 0.5 °C per cycle), and 40 s at 72 °C; another 40 cycles of 15 s at 95 °C, 15 s at 54 °C, and 40 s at 72 °C, with a final elongation step of 2 min at 72 °C. The second round used 1 μL first-round PCR product, 1× Platinum Taq Buffer (Life Technologies), 200 nmol of deoxynucleotide triphosphates each, 2.5 mmol of $MgCl_2$, and 1 U of Platinum Taq polymerase in a total volume of 50 μL. The amplification protocol was as in the first round, without the RT step. PCR fragments were visualized by agarose gel electrophoreses with ethidium bromide staining. Quantification was done using strain-specific real-time RT-PCR assays, using photometrically quantified cRNA controls as standards in vitro transcribed from TA-cloned PCR amplicons containing the respective target regions. Table S4 shows real-time RT-PCR oligonucleotide sequences. Genomic characterizations were done by amplifying genomic sequence islets using broadly reactive PCR primers designed as above, Sanger sequencing of amplicons and amplification of genomic regions between islets using gene-specific bridging forward and reverse primers, followed by primer walking to sequence large PCR amplicons, generated using the Expand High Fidelity kit (Roche). Broadly reactive primers for genomic islet amplification are shown in Table S5. Genome ends were characterized using a rapid amplification of cDNA ends (RACE) strategy using the 5′/3′ RACE Kit (Roche). Strand-specific cDNA priming was done using the SSIII RT (Life Technologies) with a manual hotstart using two separate mixes: one containing enzyme, buffer, primers, and water, and another containing RNA, water, and buffer in a preheated thermocycler for 20 min at 60 °C. After denaturation of the reverse transcriptase at 80 °C for 20 min, cDNA was amplified using the strain-specific real time PCR assays detailed in Table S4. To control for unspecific priming or self-priming, cDNA synthesis was done in parallel without primers. To avoid false-negative results from RNA degradation after repeated freeze-thawing cycles of primary nucleic acid eluates, RNA was newly purified from four bats and four hedgehogs of which sufficient clinical specimens were still available. No adequate material was available from rodents and shrews.

Host Typing. In cases of detection of closely related hepatoviruses in distinct hosts, morphological hosts designations done by specialized field biologists were confirmed by amplification and sequencing of the partial mitochondrial COI gene modified from ref. 45. Morphological and molecular typing was consistent in all hosts.

Bioinformatics. Phylogenies were done with MrBayes V3.1 (46) using a WAG amino acid or a GTR+G+I nucleotide substitution matrix. Two million MCMC iterations were sampled every 100 steps, resulting in 20,000 trees. Burn-in was generally 25% of tree replicates. AEV was generally used as an outgroup. Visualization and annotation was done with FigTree V1.4 from the BEAST package. Myristoylation sites were predicted with the MYR predictor (47). Protein modeling was done using the SWISS-MODEL web server and visualized using the UCSF Chimera package. Comparisons of predicted protein secondary structures across sequence alignments were done using ESPript V3.0 (48). IRESes were folded manually and using mfold (49). Prediction of *cre*-elements was done using mfold (49) in agreement with ref. 18. G+C content, CpG bias, and codon usage were determined using SSE V1.2 (50). Sequence distances within and between hepatovirus polyproteins were calculated using SSE V1.2 (50) with a 250-aa sliding window and a 25-aa step. *Picornaviridae* genera used for comparisons included *Aphthovirus, Aquamavirus, Avihepatovirus, Avisivirus, Cardiovirus, Cosavirus, Dicipivirus, Enterovirus, Erbovirus, Gallivirus, Hunnivirus, Kobuvirus, Kunsagivirus, Megrivirus, Mischivirus, Mosavirus, Oscivirus, Parechovirus, Pasivirus, Passerivirus, Rosavirus, Sakobuvirus, Salivirus, Sapelovirus, Senecavirus, Sicinivirus, Teschovirus,* and *Tremovirus.* Sequence distances and neighbor-joining phylogenetic reconstructions were generated using MEGA5 (51). Transmembrane domains were predicted using TMPRED from the Expasy portal (52). ASRs were done in Mesquite as described previously (3). NT sites were selected as summarized in ref. 8. Genome annotations were done on MAFFT alignments and homology to the HAV 18f reference sequence. Nonrecombinant HAV strains representing the full viral diversity in humans were selected for phylogenetic reconstructions. Virus concentrations in solid organs were plotted, and χ^2 tests for detection rates in the bat maternity roost with one-tailed P values were calculated using SPSS V22 (IBM). The hepatovirus host tree was reconstructed using systematic constraints (available upon request).

Serology. An indirect IFA was done using FRhK-4 cells persistently infected with HAV. Cells were fixed with paraformaldehyde [4% (vol/vol)], permeabilized with 0.5% Triton X-100 in 1× PBS for 5 min, and processed as described previously (53). Bat sera were diluted 1:50. Reactions were detected with goat-anti-bat Ig (Bethyl; 1:1,000) and cyanine 2 (Cy2)–labeled donkey anti-goat Ig (Dianova; 1:100). Rodent sera were tested at 1:10 or 1:40 dilution, depending on the available volumes. For secondary detection, a goat anti-mouse Ig (Dianova; 1:2,000) and a donkey anti-goat cyanine 3-labeled Ig (Dianova; 1:200) were used. Shrew and hedgehog sera were tested at a 1:10 dilution. For secondary detection, a rabbit-anti-shrew Ig (1:200) and a goat anti-rabbit cyanine 3-labeled Ig (Dianova; 1:200) were used. Reactivity of the secondary anti-shrew Ig with hedgehog Ig was confirmed previously (44). Infected cells (100% of cells infected) were mixed with noninfected cells in 1:1 ratio to allow internal negative controls. Additionally, mock-infected cell cultures were used as controls with sera from each animal group under study to exclude nonspecific reactivity with potential cell culture contaminants.

IP was done as follows: 100 μL cell culture supernatant containing the HAV/7 variant of the prototype strain HM175 with a concentration of 3.8×10^9 RNA copies/mL was incubated with 30 μL human/bat serum in a 500-μL reaction containing 220 μL incubation buffer and 150 μL of RNase-free water on a tumbler overnight, followed by overnight incubation with prewashed Dynabeads Protein G for Immunoprecipitation (Life Technologies), followed by five washing steps using a buffer containing TritonX-100, DTT, and protease inhibitor (Roche), followed by RNA purification using the TRIzol LS kit (Life Technologies) and an RNeasy cleanup protocol (Qiagen) with an elution volume of 30 μL, of which 5 μL were tested by real-time RT-PCR as in

ref. 54 with in vitro transcribed cRNA controls for quantification. The HAV-specific TaqMan real time RT-PCR assay oligonucleotide binding sites (54) had several nucleotide mismatches with the *Eidolon* bat virus, making cross-detection highly unlikely. To further exclude cross-detection of coprecipitated bat hepatovirus by the HAV-specific real time RT-PCR assay, the bat-specific real time RT-PCR used for quantification (Table S4) was performed on all bat sera used for IP. Only one serum tested weakly positive. To further rule out falsely high HAV-specific real-time RT-PCR results, a second real-time RT-PCR assay was designed specifically for the HAV inoculum and contained >30 nucleotide mismatches with the bat hepatovirus (Table S4 shows oligonucleotide sequences of the specific assay). All results were consistent with the initial HAV-specific quantification. Negative human control sera were determined using an IgM/IgG ELISA kit on an architect platform (Abbott). Additional negative controls were mAbs against the paramyxovirus SV5 (Abcam), the rhabdovirus Vesicular stomatitis virus (Sigma), the hepatitis C virus (HCV) E1 and E2 proteins (Santa Cruz), and polyclonal rabbit sera raised against the HCV NS3 (Abbiotec) and avian Influenzavirus A H5N1 (Abcam), as well as a human polyclonal serum against the hepatitis B surface antigen (0.2 IU/mL; NIBSC). Bat sera were additionally tested in 1:50 and 1:500 dilutions in PBS, yielding near-identical quantitative relationships.

A virus neutralization test (NT) was performed as described in ref. 55 with minor modifications. HAV [10^4 50% tissue culture infective dose ($TCID_{50}$)/mL; 1.5×10^5 RNA copies/mL as quantified by real-time RT-PCR] was incubated with 50 μL of sera/mL or 20 and 50 mU of monoclonal anti-HAV IgG antibody 7E7 (Mediagnost), respectively, as controls on a tumbler for 3 h at room temperature in DMEM. Neutralization was assayed by infection inhibition, which was determined by titration on FRhK-4 cells ($TCID_{50}$ titer) and compared with that obtained with HAV alone in amounts equal to those in the approaches with serum. Serum samples, which reduced the HAV titer by at least one log (more than 90% neutralization), were considered positive for anti-HAV neutralizing antibodies. As additional controls, HAV (10^4 $TCID_{50}$/mL) was treated with human anti–HAV-positive as well as -negative sera or 1–15 mU/mL of the WHO international standard for anti-HAV Ig (NIBSC). $TCID_{50}$ titer was determined 12 d after inoculation by IFA using the 7E7 monoclonal antibody (1:800; Mediagnost) with a goat anti-mouse Alexa488-coupled secondary antibody (1:1,000; Life Technologies) and calculated by the Kärber method.

In Situ Hybridization. RNAScope RNA probes targeting ~1,000 nucleotides of the VP2 and VP3 genes of *Eidolon helvum* hepatoviruses were custom designed by Advanced Cell Diagnostics. Specimens M32 and GH297 were selected due to best tissue quality and high virus concentration in several organs. In situ hybridization was performed as described by the manufacturer.

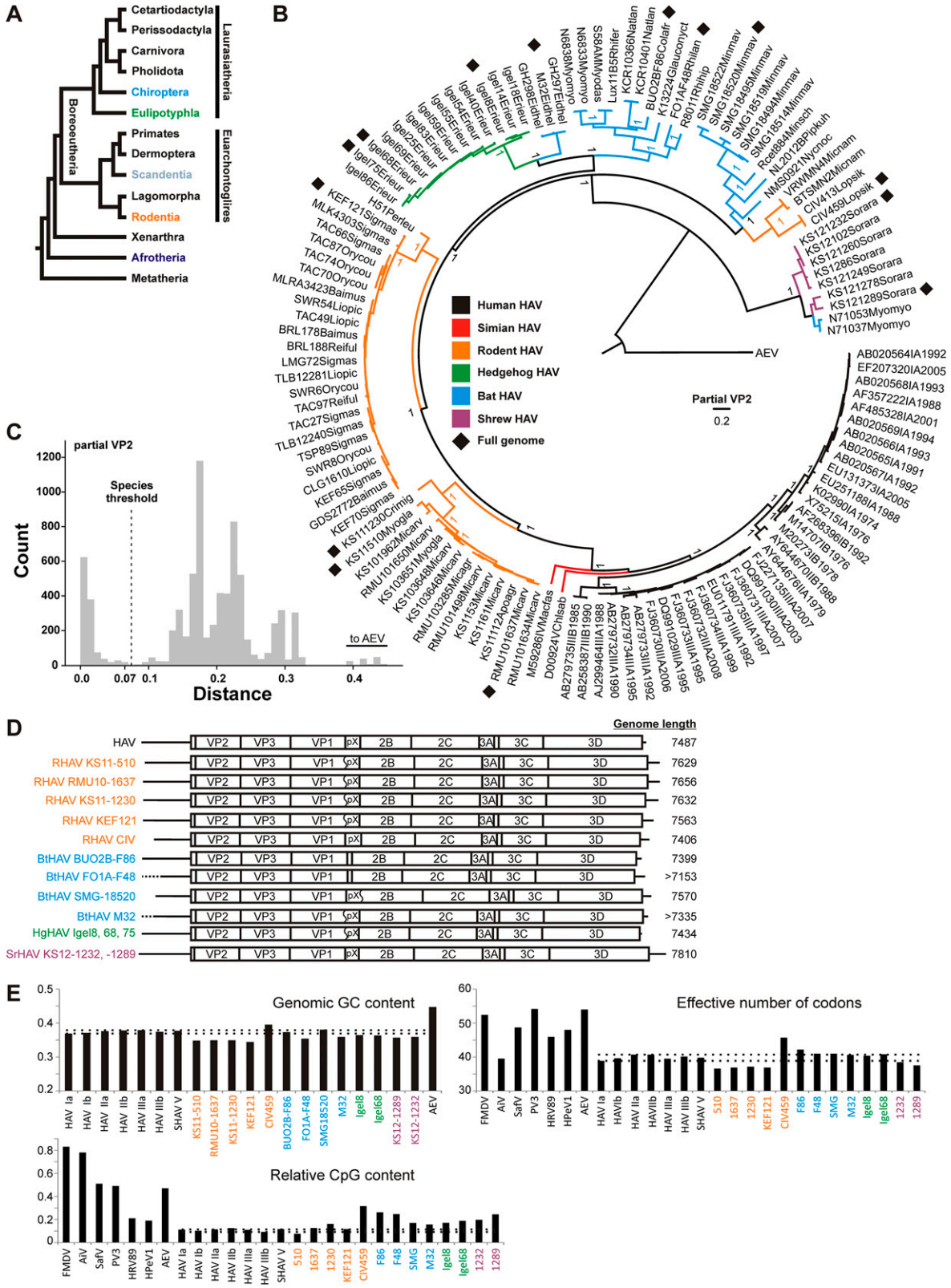

Fig. S1. Host relationships, hepatovirus relationships, viral species prediction, and genomic properties. (*A*) Phylogenetic relationships of mammals included in this study. Relationships of placental mammals according to ref. 56. Sampled taxa are colored. The superorder Afrotheria contains the sampled order Afrosoricida. (*B*) Phylogenetic relationships of hepatoviruses in the partial VP2 encoding domain. Values at nodes show support of grouping from Bayesian posterior probabilities. (*C*) VP2-based species prediction. Pairwise amino acid sequence distances in the translated VP2 were plotted for the complete dataset shown in *B*. The 7% cutoff separating hepatovirus species in this dataset is highlighted by a slashed line. The distance to the next closely related genus *Tremovirus* (represented by AEV) is highlighted to the right. (*D*) Hepatovirus genome architecture. Rodent hepatovirus (RHAV), bat hepatovirus (BtHAV),

Legend continued on following page

hedgehog hepatovirus (HgHAV), and shrew hepatovirus (SrHAV). Near-identical HgHAV and SrHAV genomes were represented by one virus only. Genome lengths are given to the right. Domains whose initiation and termination could not be unambiguously identified are shown with curved lines. Annotations and additional genomic features correspond to those in the main article file. The 5′UTR of BtHAVs FO1A-F48 and M32 could not be completely characterized despite repeated trials. The uncharacterized portion is given with a dotted line and the minimum genome size is identified with >. The HAV reference sequence corresponds to the 18f prototype strain. (*E*) Genomic features of hepatoviruses compared with other picornaviruses. (*Top*) Percentage G+C content of hepatovirus genomes. (*Middle*) Relative CpG content of hepatoviruses compared with representative picornaviruses (FMDV, foot-and-mouth disease virus, genus *Aphthovirus*; AiV, Aichi virus, genus *Kobuvirus*; SafV, Saffold virus, genus *Cardiovirus*; PV3, Polio virus 3, HRV89, human rhinovirus 89, both genus *Enterovirus*; HPeV1, human parechovirus 1, genus *Parechovirus*; AEV, genus *Tremovirus*). (*Bottom*) Effective number of codons. In all panels, dashed lines show minimum and maximum of primate HAV.

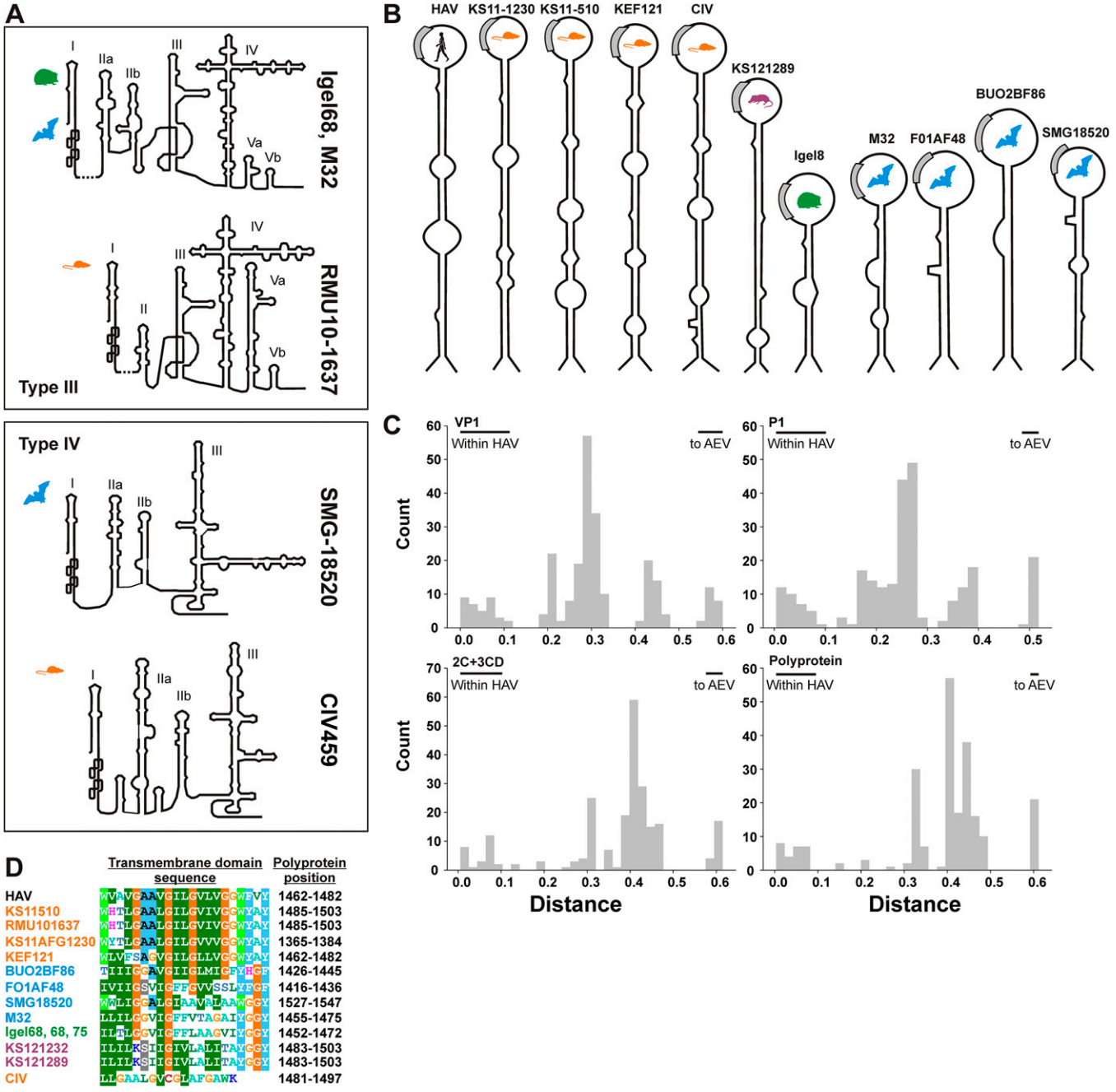

Fig. S2. Hepatovirus secondary structures and diversity in different genome regions. (*A*) Predicted hepatovirus 5′UTR RNA secondary structure of viruses not shown in the main text. Stem-loops (SL) are identified by adjacent numbers. Pseudoknots are given as boxes. Pictograms represent the full 5′UTR from the most 5′ nucleotide to the polyprotein initiation codon at the 3′-end. HgHAV Igel68 represents viruses Igel8 and Igel68; RHAV RMU10-1637 represents the near-identical virus KS11-510 (all these structures were type III IRES-related; SMG-18520 represents Malagasy bat virus; CIV459 represents Ivorian rodent virus). Both viruses contained type IV-like IRES structures including the characteristic pseudoknot preceding the initiation codon. Unpaired sequence between SL I and SL II is represented by dots for graphical reasons. (*B*) Predicted hepatovirus *cis*-acting replication elements (*cre*) in the $3D^{pol}$ genomic region. Hosts are shown with pictograms together with virus strains, HAV is included as a reference (genotype Ia, GenBank accession no. AB020564). The conserved AAACG motif in the upper loop structure is highlighted. Rodent virus RMU10-1637 is represented by the near-identical virus KS11-510. Shrew virus KS12-1289 represents the near-identical virus KS12-1232. Hedgehog virus Igel8 represents the near-identical viruses Igel68 and Igel75. (*C*) Pairwise amino acid sequence distances in the translated VP1, P1, the combined 2C+3CD domains, and the full polyprotein gene plotted for the dataset containing all full genomes used to generate phylogenetic reconstructions shown in the main article. Distances within HAV and those to the next closely related AEV are highlighted. (*D*) Predicted transmembrane domains of hepatoviruses.

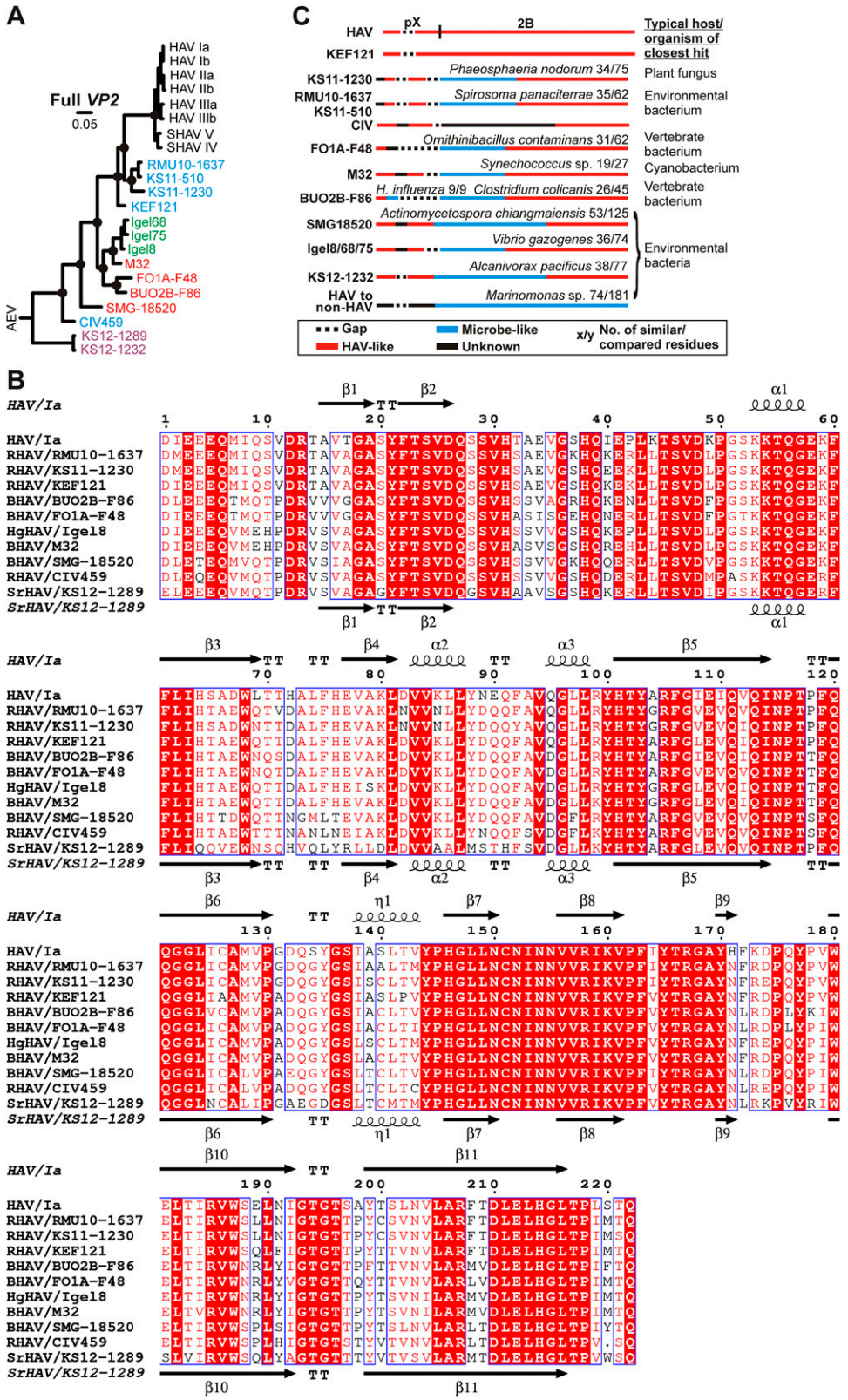

Fig. S3. Conservation of hepatovirus VP2 proteins and divergence of hepatovirus 2B proteins. (*A*) Bayesian phylogeny of the full VP2 using a WAG amino acid substitution model as detailed in *SI Materials and Methods*. Filled circles at nodes represent Bayesian probabilities above O.9. (*B*) Identification of structural elements [α-helices, 3_{10}-helices (η), β-strands, and strict β-turns (TT)] using the structural models built on the HAV crystal structure (8). Conserved alignment domains are highlighted by red color and boxes. The dot symbol at the C terminus of RHAV CIV459 corresponds to a gap compared with the other sequences. (*C*) Protein BLAST comparisons of pX and 2B domains. In comparisons of HAV; hepatoviruses were excluded as hits. For each genome segment, the number of residues with identical or biochemically similar residues is indicated per the total number of compared residues. The BLOSUM62 matrix was used for similarity calculations by BLAST according to default parameters.

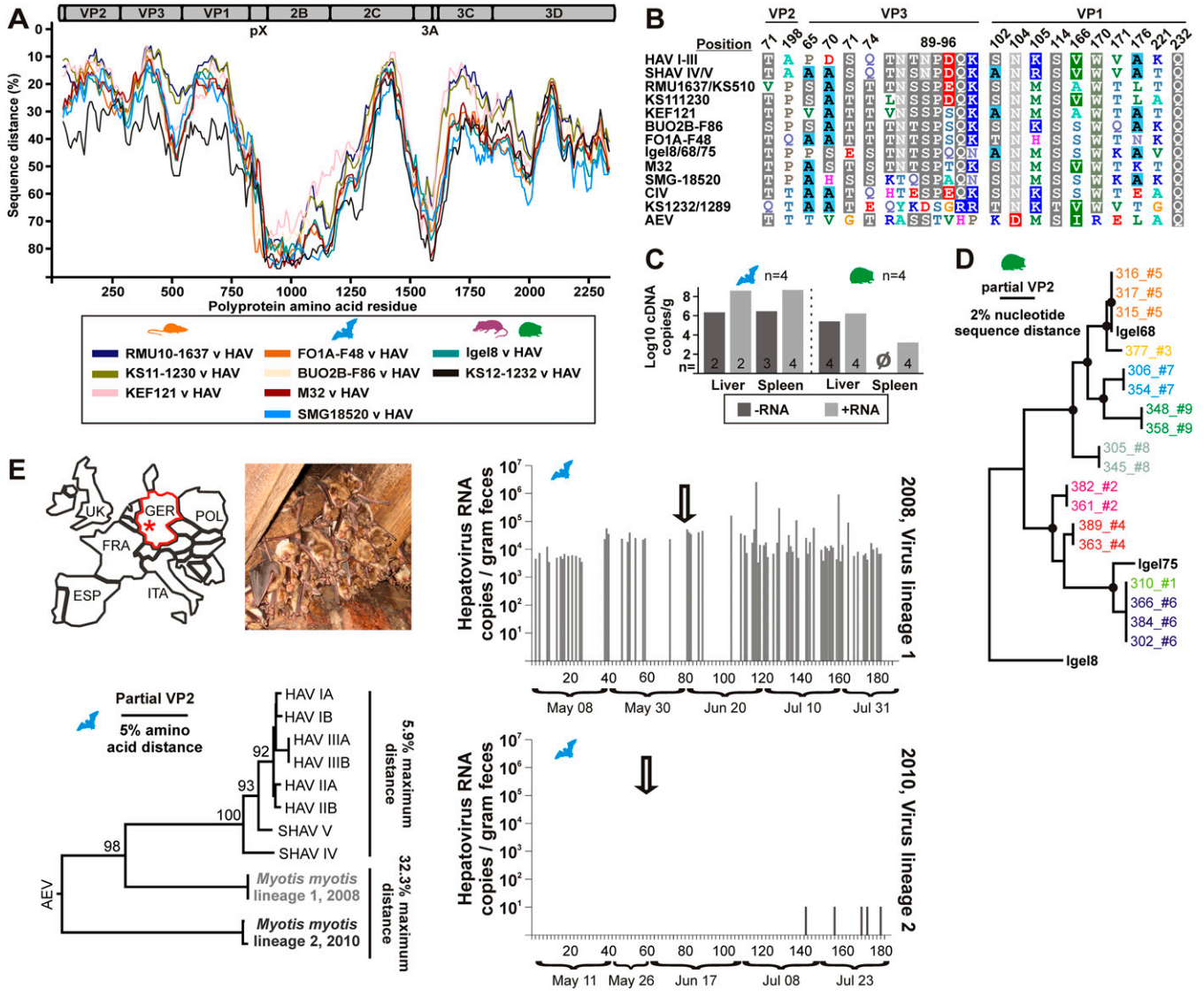

Fig. S4. Hepatovirus epitopes and infection patterns. (*A*) Amino acid sequence distance along the polyproteins between HAV (Gt Ia) and the nonprimate hepatoviruses. A genomic representation of HAV is given on top. Near-identical viruses from rodents, hedgehogs and shrews are represented by one virus only. (*B*) Hepatovirus epitopes associated with neutralization according to ref. 8. Background shading and residue color indicate biochemical properties of residues. AEV was included for comparison. (*C*) Strand-specific cDNA priming in bat and hedgehog liver and spleen. The number of analyzed animals and biological replicates is indicated on top and above organs, respectively. (*D*) Partial VP2 neighbor-joining percentage distance phylogeny of the hedgehog viruses in the animal shelter (complete deletion option, 405 nucleotides, GenBank accession nos. KT452748–KT452765). Individual animals are color-coded, sampling number corresponds to increasing time points. The three hedgehog viruses for which the full genomes were characterized were included for orientation. (*E*) (*Upper Left*) Geographic location of the roost in Western Germany as detailed in ref. 25 and a picture of female *Myotis myotis* bats hanging from the roof of the attic forming the roost. (*Lower Left*) Neighbor-joining amino acid percentage distance phylogeny highlighting the distance between the two hepatovirus lineages detected in the roost in 2008 and 2010 compared with the maximum distance in HAV using the same VP2 fragment. (*Right*) Detections of lineage 1 during 2008 and detections of lineage 2 during 2010. Height of bars correspond to viral RNA copies per gram of feces; empty spaces correspond to negative samples. The arrow symbolizes the birth of the first pup, after which all pups are born in a short time frame. Height of bars in 2010 was set to 10^1 to highlight that these specimens were detectable only by nested RT-PCR but could not be quantified using a strain-specific real-time RT-PCR assay due to low RNA concentrations. Data for the year 2009 is not shown because this sample was incomplete as detailed in ref. 25 and no specimen tested positive for hepatoviruses.

Other Supporting Information Files

Table S1 (DOCX)
Table S2 (DOCX)
Table S3 (DOCX)
Table S4 (DOCX)
Table S5 (DOCX)