

Identification of GBV-D, a Novel GB-like Flavivirus from Old World Frugivorous Bats (*Pteropus giganteus*) in Bangladesh

Jonathan H. Epstein¹, Phenix-Lan Quan², Thomas Briese^{2*}, Craig Street², Omar Jabado², Sean Conlan², Shahneaz Ali Khan^{1,3}, Dawn Verdugo², M. Jahangir Hossain⁴, Stephen K. Hutchison⁵, Michael Egholm⁵, Stephen P. Luby⁴, Peter Daszak^{1*}, W. Ian Lipkin²

1 Conservation Medicine Program, Wildlife Trust, New York, New York, United States of America, **2** Center for Infection and Immunity, Mailman School of Public Health, Columbia University, New York, New York, United States of America, **3** Chittagong Veterinary & Animal Sciences University, Chittagong, Bangladesh, **4** Programme on Infectious Disease and Vaccine Sciences, International Centre for Diarrheal Disease Research, Bangladesh (ICDDR,B), Dhaka, Bangladesh, **5** 454 Life Sciences, Branford, Connecticut, United States of America

Abstract

Bats are reservoirs for a wide range of zoonotic agents including lyssa-, henipah-, SARS-like corona-, Marburg-, Ebola-, and astroviruses. In an effort to survey for the presence of other infectious agents, known and unknown, we screened sera from 16 *Pteropus giganteus* bats from Faridpur, Bangladesh, using high-throughput pyrosequencing. Sequence analyses indicated the presence of a previously undescribed virus that has approximately 50% identity at the amino acid level to GB virus A and C (GBV-A and -C). Viral nucleic acid was present in 5 of 98 sera (5%) from a single colony of free-ranging bats. Infection was not associated with evidence of hepatitis or hepatic dysfunction. Phylogenetic analysis indicates that this first GBV-like flavivirus reported in bats constitutes a distinct species within the *Flaviviridae* family and is ancestral to the GBV-A and -C virus clades.

Citation: Epstein JH, Quan P-L, Briese T, Street C, Jabado O, et al. (2010) Identification of GBV-D, a Novel GB-like Flavivirus from Old World Frugivorous Bats (*Pteropus giganteus*) in Bangladesh. *PLoS Pathog* 6(7): e1000972. doi:10.1371/journal.ppat.1000972

Editor: Colin Parrish, Cornell University, United States of America

Received: January 11, 2010; **Accepted:** May 27, 2010; **Published:** July 1, 2010

Copyright: © 2010 Epstein et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This study was funded by awards from the National Institutes of Health including K08AI067549, AI079231, AI57158 (Northeast Biodefense Center-Lipkin), AI070411, the United States Agency for International Development (USAID) Emerging Pandemic Threats (EPT) Program, PREDICT project, under the terms of Cooperative Agreement Number GHN-A-OO-09-00010-00, the Department of Defense, the Rockefeller Foundation, and Google.org. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: thomas.briese@columbia.edu (TB); daszak@wildlifetrust.org (PD)

Introduction

Bats (order *Chiroptera*), after rodents, comprise the most diverse group of mammals with more than 1,100 species. They are present on six continents, often have substantial habitat overlap with humans [1] and harbor several zoonotic viruses causing significant human morbidity and mortality, including Ebola- and Marburg-virus, Nipah virus (NiV), and SARS-like coronaviruses [2–5]. Proximity of bats to human populations may facilitate the zoonotic transmission of viruses either through direct contact, via amplifying domestic animal hosts, or through food-borne routes [6–8].

The current study was set up as part of a viral discovery effort to target key wildlife reservoirs in emerging disease hotspots. Bangladesh is a ‘hotspot’ for emerging zoonotic diseases [9], with a relatively high diversity of wildlife that likely harbors new zoonotic pathogens, one of the densest human populations on the planet, and a high level of connectivity between people, domestic animals and wildlife. In Bangladesh and India, frugivorous *Pteropus giganteus* bats have been identified as a reservoir for NiV [10,11], which has been recognized as the cause of several outbreaks of encephalitis [12–14]. *Pteropus giganteus* bats are common throughout the Indian subcontinent, living in close association with humans and feeding on cultivated fruit [14]. NiV transmission from bats to humans has been linked with the harvest and consumption of raw date palm sap, which becomes

contaminated with bat feces, urine or saliva overnight when bats such as *P. giganteus* come to feed from the collecting pots [14,15]. Date palm sap or other foods eaten by both bats and people, may also serve as a vehicle for transmission of other bat-borne agents.

Several zoonotic flaviviruses, including Japanese encephalitis virus, West Nile virus, and Kyasanur forest virus have been identified in bats; however, to date, GB viruses have not [1]. GB viruses A and C (GBV-A and -C) represent two recently identified species that are currently unassigned members of the family *Flaviviridae* [16]. GBV-A viruses have been described in New World primates and are not known to infect humans [17–19], while GBV-C (also known as Hepatitis G virus (HGV)) have frequently been isolated from humans in many regions of the World, including India and Bangladesh [19–23], and from wild chimpanzees (*Pan troglodytes*) in Africa [24,25]. Here we describe discovery of a virus in the serum of healthy bats in Bangladesh, tentatively named GB virus D (GBV-D), that is distantly related to GBV-A and -C and represents a new member of the family *Flaviviridae*.

Materials and Methods

Ethics statement

Every effort was made to minimize bat stress and avoid injury during capture, restraint, and sampling procedures. This study was

Author Summary

Bats are important reservoirs for emerging zoonotic viruses with significant impact on human health including lyssaviruses, filoviruses, henipaviruses and coronaviruses. Opportunities for transmission to humans are particularly prominent in countries like Bangladesh, where people live in close association with bats. Whereas previous studies of bats have employed assays that test for known pathogens, we present the first application of an unbiased molecular approach to pathogen discovery in this reservoir for emerging zoonotic disease. Unbiased pyrosequencing of serum from *Pteropus giganteus* bats enabled identification of a novel flavivirus related to Hepatitis C and GB viruses. Viral nucleic acid was present in 5 of 98 (5%) sera, and in the saliva of one animal. Sequence identification of two strains of the virus, tentatively named GBV-D, suggests *P. giganteus* as a natural reservoir. Detection of viral nucleic acid in saliva provides a plausible route for zoonotic transmission. Phylogenetic analysis indicates that GBV-D is ancestral to GBV-A and -C, and separate from the recently classified genus *Hepacivirus*. Our findings provide new insight into the range of known hosts for GB-like viruses and demonstrate the power of unbiased sequencing to characterize the diversity of potentially zoonotic pathogens carried by bats and other reservoirs.

conducted following Wildlife Trust institutional guidelines under IACUC approval G2907 issued by Tufts New England Medical Center, Boston, Massachusetts.

Bat sample collection

As part of a longitudinal surveillance study of Nipah virus in bats, 98 free-ranging *P. giganteus* bats were caught from a colony of approximately 1800 individuals in the Faridpur district of Bangladesh in December 2007 (**Figure 1**). Each bat was anesthetized using isoflurane gas; morphometric measurements (weight, forearm length, head length, and body condition) were taken and bats were aged [10]. Each bat was marked for future identification using an RFID microchip (AVID corp, www.avidid.com) implanted subcutaneously between the scapulae. Three mL of blood were collected and placed into serum separator tubes (vacutainer; Becton Dickinson, Franklin Lakes, NJ, USA). Serum was allowed to separate overnight at 4°C then drawn off without centrifugation and immediately frozen using a liquid nitrogen dry shipper. To inactivate potentially infectious agents, serum samples were heat-treated at 56°C for 30 min and then stored at -70°C. For RNA extraction, 250 µL of serum was added to 750 µL Tri-Reagent LS (Molecular Research Center, Cincinnati, OH, USA). Saliva was collected from the bat's throat using a sterile cotton swab. Urine was collected either by catching urine in a 1.0 mL sterile cryovial while the bat was urinating, or by urethral swab. Urine and saliva swabs were immediately placed into 1 mL Tri-Reagent LS and frozen in liquid nitrogen.

Unbiased high-throughput pyrosequencing (UHTS)

Total RNA from serum was extracted for UHTS analysis to screen for the presence of microorganisms. Five microliters of total RNA from each bat were combined into 4 pools: 4 pregnant bats; 4 non-pregnant female bats, and 2 pools of 4 adult male bats, respectively. Reverse transcription (RT) was performed on DNase I-treated (DNA-free, Ambion Inc., Austin, TX, USA) RNA pools to generate cDNA using Superscript II RT (Invitrogen, Carlsbad, CA, USA) and random octamers linked to a defined arbitrary, 17-mer primer sequence tail (MWG,

Huntsville, AL, USA) [26]. After RNase H treatment cDNA was amplified by the polymerase chain reaction (PCR), applying a 9:1 mixture of the defined 17-mer primer sequence and the random octamer-linked 17-mer primer sequence, respectively [27]. Products of >70 base pairs (bp) were selected by column purification (MinElute, Qiagen, Hilden, Germany) and ligated to specific linkers for sequencing on the 454 Genome Sequencer FLX (454 Life Sciences, Branford, CT, USA) without DNA fragmentation [28,29]. Sequences were analyzed using software applications implemented at the GreenePortal website (<http://tako.cpmc.columbia.edu/Tools/>).

Genome sequencing

Multiple forward and reverse primers for RT-PCR (available upon request) were designed using the sequences obtained by UHTS in order to fill gaps between fragments. Amplifications were performed with Bio-X-act (Biolone, London, UK) according to manufacturer's protocols. Products were size fractionated by electrophoresis and directly sequenced in both directions with ABI PRISM Big Dye Terminator 1.1 Cycle Sequencing kits (Perkin-Elmer Applied Biosystems, Foster City, CA, USA) at a commercial facility (Genewiz, South Plainfield, NJ, USA). Additional methods applied to obtain the genome sequence included touch-down PCR [30], 2-step walking PCR [31], and 3'- and 5'- RACE (Invitrogen).

Quantitative real-time PCR

A real time Taqman PCR assay was developed to screen bat samples for GBV-D. Reactions were performed in a 25 µL volume by using commercial Taqman Universal Master Mix (Applied Biosystems, Foster City, CA, USA). Primers and probe were designed to target a 60 nt region in the NS4A gene region: Fadi-forward, 5'- gCAGCTgCgTgTgCCA; Fadi-reverse, 5'- ACACC-CATgATgTTACCCACgAC; Fadi-probe, 5'- FAM- AggACCCgg-TcGCTCCAgCA-T-BQX (TIB Molbiol, Adelphia, NJ, USA). Cycling conditions were: 50°C for 2 min, and 95°C for 10 min, followed by 45 cycles at 95°C for 15 sec and 60°C for 1 min. Thermal cycling was performed in an ABI 7300 real-time PCR system (Applied Biosystems).

Serum chemistry

A liver function panel was conducted at the International Center for Diarrheal Disease Research (Dhaka, Bangladesh) using non heat-treated bat sera (Automated Chemistry Analyzer AU 640, Olympus Corporation, Tokyo, Japan). The following parameters were analyzed: total protein, albumin, globulin, albumin:globulin ratio, total cholesterol, total bilirubin, alkaline phosphatase, alanine transferase, aspartate aminotransferase, gamma glutamyltransferase, and lactate dehydrogenase.

Phylogenetic and sequence analyses

Sequence alignments were generated with ClustalW software [32] and phylogenetic relationships deduced using Geneious software [33]. Statistical significance was assessed by bootstrap re-sampling of 1000 pseudoreplicate data sets. Sequence relations were determined from p-distance matrices calculated with pairwise deletion for missing data and homogeneous patterns among lineages based on ClustalW alignments as implemented in MEGA software [34]. Sliding window similarity analysis was performed using SimPlot [35]. Potential signalase cleavage sites, glycosylation sites, and phosphorylation sites were analyzed using the respective prediction servers available at the Center for Biological Sequence Analysis (<http://www.cbs.dtu.dk/services/>).



Figure 1. Map showing the location of the bat colony in Faridpur district, Bangladesh from which GBV-D was identified.
doi:10.1371/journal.ppat.1000972.g001

Results

Identification of a GB-like agent from bats

Total RNA from the serum of healthy bats captured at a roost in the Faridpur district of Bangladesh was extracted for UHTS analysis. Extracts of 16 individual bats were combined into 4 pools consisting of 4 pregnant adult bats, 4 non-pregnant adult female bats, or 2×4 adult male bats. Each pool yielded between 1,400 and 2,000 assembled contigs or singleton reads (representing 50,000–75,000 reads ranging in size from 31–328 nt). Two reads of 238 and 215 nucleotides (nt) derived from the pregnant bat pool had distant homology to GBV-A sequences at the deduced amino acid (aa) level in the E2 and NS4A gene regions respectively (BLASTX); no homology was detected by searches at the nt level (BLASTN; local copy of the executables with standard settings except that the reward for a nucleotide match was set to 2 instead of 1). No viral sequences were detected in other pools at the nt or aa levels. Screening of the individual RNA preparations from the pregnant bat pool using primers derived from the UHTS reads confirmed the presence of the GBV-like sequence in the serum of bat 93. A quantitative real time PCR assay indicated a load of approximately 30 000 RNA copies in bat-93 serum extract, and identified an additional 4 positive bat sera from the original 98 samples (5/98; 5%), indicating serum loads ranging from 350 to 70,000 RNA copies per assay. These positive samples came from

male bats that were not included in the initial UHTS pools. Extracts of saliva from the five positive bats indicated a load of approximately 200 RNA copies in bat 93; no signal was obtained with urine extracts from the five positive bats.

Genomic characterization of GBV-D

Near full-length genome sequence was generated from bat-93 and a second positive serum (bat 68), applying primers crossing

Table 1. Percent sequence similarity between GBV-D (bat-68), -A, -C, and hepaciviruses.

nt*	aa				
	GBV-D	GBV-C	GBV-A	GBV-B	HCV-1
GBV-D		48	46	39	41
GBV-C	41		55	41	44
GBV-A	39	47		37	38
GBV-B	25	24	24		41
HVC-1	22	22	22	27	

*Sequence similarity at nt level in upper right, and at aa level in lower left portion of the table.

doi:10.1371/journal.ppat.1000972.t001

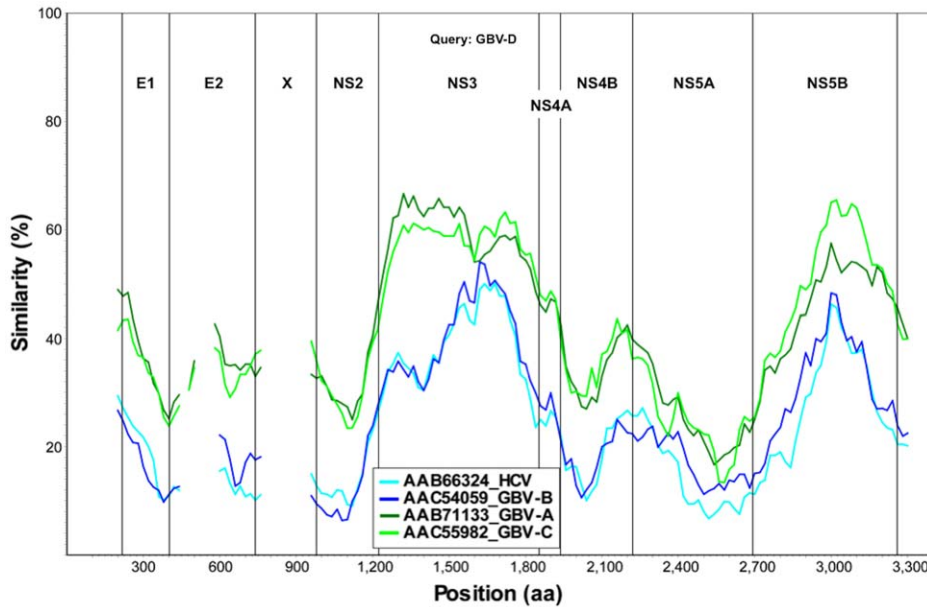


Figure 2. Sliding window similarity analysis between GBV-D and other GBV and hepaciviruses (amino acid sequence; window, 160; step, 20).
doi:10.1371/journal.ppat.1000972.g002

gaps between UHTS reads as well as touch-down PCR [30], 2-step walking PCR [31], and 3'- and 5'-RACE (Invitrogen) protocols. The two genome sequences were 96% identical at the nt level (GenBank Accession nos. GU566734 and GU566735), indicating two strains of the same virus. Comparison of deduced polyprotein sequence to other GBV and hepaciviruses indicated highest nt and aa sequence identities to GBV-A and -C (**Table 1, Figure 2**). The genomic sequence of the GBV-like virus identified in *P. giganteus* bats, tentatively named GBV-D, comprises 9,633 nt with 52 nt of potentially 5'-untranslated region (UTR), one continuous open reading frame (ORF) of 9318 nt (3106 aa) and 265 nt of 3'-UTR (**Figure 3**).

Mature structural proteins in GB viruses, as well as other flaviviruses, are the product of cleavage by host signal peptidase [36]. In GBV-D the first potential signal sequence cleavage site is present after a stretch of 57, largely basic aa (6 kDa, pI = 12), followed by sequence homologous to E1 (pfam 01539, <http://pfam.sanger.ac.uk/>) (**Figure 3**). The single glycosylation site N₁₁₇IT present in that sequence is located in a position comparable to GBV-C, -A, -B and HCV glycosylation sites. Identification of the downstream E2 termini is less apparent as the next 580 aa contain multiple potential signal sequences and 10 potential glycosylation sites that indicate no homology to hepaciviral E2/NS1 (pfam 01560), until the sequence aligns with

N-terminal NS2 motifs (pfam 01538) (**Figure 2, Figure 3**). However, despite similarity to pfam 01538 no signal sequence compatible with cleavage at A₇₅₉/A was found; cleavage may occur at G₈₂₆/R, which combined with potential signalase cleavage at A₅₈₄/F may indicate the existence of a heavily glycosylated potential 26 kDa product instead of the p7 transmembrane protein identified in HCV [37–39] or the 13 kDa variant described in GBV-B [40,41]. Conserved C-terminal motifs of the autocatalytic NS2/NS3 endoprotease domain are compatible with NS2/NS3 cleavage at S₁₀₆₇/A and comparable to other GBV and HCV [42]. **Figure 3** indicates potential cleavage sites for NS3 (peptidase S29, pfam 02907; DEAD box helicase, pfam 07652; helicase C, pfam 00271), NS4A (pfam 01006), NS4B (pfam 01001), NS5A (domain-1a zinc finger, pfam 08300; domain-1b, pfam 08301), and NS5B (pfam 00998).

Conserved aa motifs were recognized in NS proteins. RNA-dependent RNA polymerase (RdRp) motifs in RdRp block III that are conserved with respect to other GBV and hepaciviruses were identified in NS5B (**Figure 3**) [43–46]. Potential phosphorylation sites are present at multiple serine (9), threonine (14) and tyrosine (4) residues in NS5A, compatible with its possible function as a phosphorylation-regulated mediator of viral replication [47]. However, significant conservation of primary sequence is not obvious for phosphorylation sites, proline-rich, or interferon-sensitivity

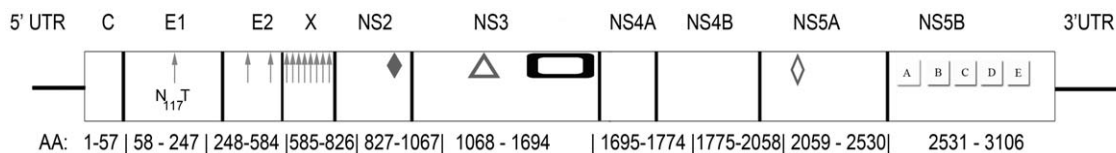


Figure 3. Genomic organization of GBV-D, a novel flavivirus identified in the sera of frugivorous bats in Bangladesh. Arrows, glycosylation sites; solid diamond, active center sites H₉₂₁, E₁₀₁₁, and C₁₀₃₂ in the autocatalytic NS2/NS3 endoprotease domain; triangle, catalytic triad H₁₁₂₃, D₁₁₄₇, S₁₂₀₄ of NS3 serine protease; rectangle, NS3 helicase and DEAD-like helicase motifs; open diamond, zinc finger motif; and NS5 polymerase motifs A (T₂₇₄₄VDAICFDSCIT), B (R₂₈₀₂ASGVLTTSSNCISSFLKVSAAAC), C (F₂₈₃₅LIHGDDVMII), D (L₂₈₇₆DTAQSCSA), and E (H₂₉₀₀YFLSTDFR) motifs.
doi:10.1371/journal.ppat.1000972.g003

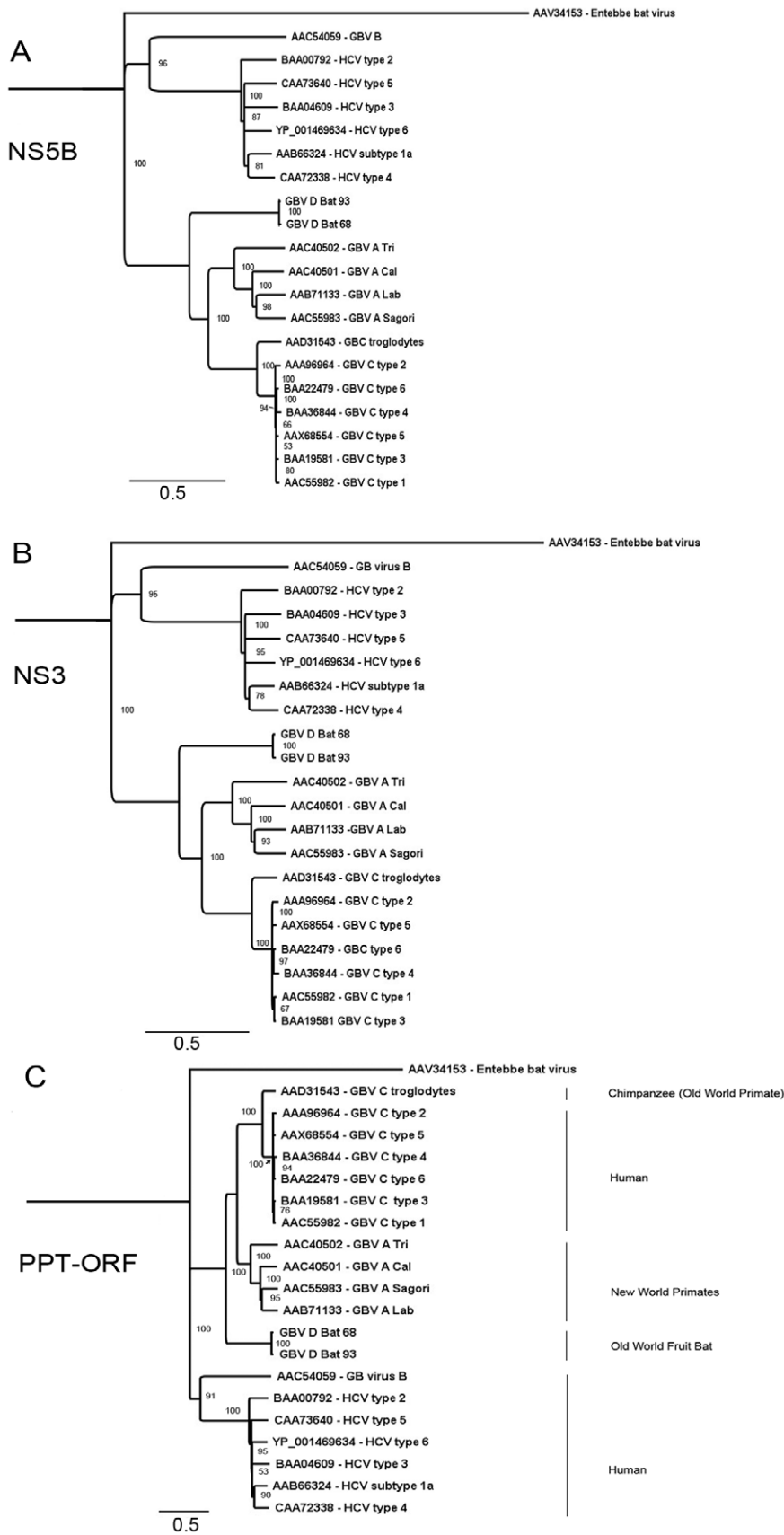


Figure 4. Phylogenetic relationship of GBV-D to other GBV and hepaciviruses. GBV-D amino acid sequences for A: NS5B, B: NS3, and C: the polyprotein (PPT) were analyzed in comparison to representative sequences of GBV-A, -B, -C and hepatitis C viruses. GenBank accession numbers for

the respective sequences are indicated. Entebbe bat virus was used as an outgroup; distance in substitutions per site is indicated by scale bars; percent bootstrap support for values greater than 85% is indicated at respective nodes.
doi:10.1371/journal.ppat.1000972.g004

determining region motifs [48–50]. The C-terminal portion of NS3 has homology to conserved NTPase/helicase motifs [51]; the N-terminal portion includes conserved active triad residues H₁₁₂₃, D₁₁₄₇, S₁₂₀₄ of serine protease [52], the viral protease responsible for cleavage of mature non-structural proteins [53]. Likewise, the active triad H₉₉₁, E₁₀₁₁, C₁₀₃₂ of the *cis*-acting protease activity in the C-terminal portion of NS2 is conserved with respect to other GBV and HCV [42]. The only other discernable motif identified was a well-conserved N₇₅ C/D C motif at the N-terminus of E1 (**Figure 3**) [54].

Phylogenetic analysis

Phylogenetic analysis of GBV-D was performed in comparison to selected representatives of GBV-A, GBV-B, GBV-C and HCV. Analysis of NS5B aa sequence (**Figure 4A**) confirmed a closer relationship of GBV-D to GBV-A and -C than to GBV-B or HCV as also indicated by pairwise sequence comparisons (**Table 1**). The same relationships were also apparent when NS3, or the complete polyprotein sequence were analyzed (**Figure 4B and C**, respectively). All three trees show GBV-D consistently at the root of the GBV-A/-C viruses, indicating an independent phylogenetic clade compatible with a separate species distinct from the recently created genus *Hepacivirus* [16].

Serum chemistries

A liver serum chemistry panel was conducted on sera from 15 bats, the five GBV-D infected and 10 non-infected animals. Standard assays to detect hepatitis and/or impaired liver function were performed [55]. Levels of total protein, alanine transferase, aspartate aminotransferase and total cholesterol were within published ranges reported for *P. giganteus*, except for bat 33 (infected) and bat 73 (uninfected), which had modest elevation in aspartate aminotransferase. Reference values for albumin, globulin, albumin:globulin ratio, total bilirubin, alkaline phosphatase, gamma glutamyltransferase and lactate dehydrogenase are not available for *P. giganteus*, however, values were comparable to those reported for other *Pteropus* species [56]. Mean values did not significantly differ between infected and uninfected bats (**Table 2**).

Discussion

Molecular analyses of sera from *Pteropus giganteus* bats from Faridpur, Bangladesh led to the identification of a 9,633 nt sequence consistent in genomic organization with known GBV and other species within the family *Flaviviridae* [16]. Whereas previous studies of bats have employed assays that test for known pathogens, ours is the

Table 2. Liver function values from *Pteropus giganteus* bats.

	T.Prot (g/dL)	Alb (g/dL)	Glob	A:G	T. Chol (mg/dL)	T. Bili (mg/dL)	Alk Phos	ALT (U/L)	AST (U/L)	GGT (U/L)	LDH (mg/dL)
Ref Range (human)	64.0–82.0	34–50	23–35	1.1–1.8	3.6–6.3	.01–19	30–120	.01–41	.01–38	9–40	135–225
Ref Range (<i>P. giganteus</i>)	65–84	30–41			0–2.3			1.2–84.6	30.2–141.0		
Bat ID											
1	86	50.7	35.3	1.44	1.65	1.8	1572.4	13.9	67.6	87.90	53.4
3	79.7	49.4	30.3	1.63	2.74	2.1	605.7	20.2	67.4	49.20	95
4	70.4	42.9	27.5	1.56	0.5	2.3	532.1	12.3	66.1	75.40	76.7
21	76	48	28	1.71	1.06	1.2	794.9	23.8	83.7	35.70	244.5
23	71.5	46.2	25.3	1.83	1.49	1.7	1650	12	115.1	102.00	96
33	80.5	43.8	36.7	1.19	1.09	2.1	891.6	22.2	146	32.10	752.4
36	89.9	51	38.9	1.31	1.22	2.5	580.4	11.1	109.4	69.00	114.4
37	76.6	47	29.6	1.59	0.85	1.3	205.8	12.6	50.3	72.50	133.9
39	70.4	43.3	27.1	1.6	0.5	1.8	328.7	20.5	77.8	31.00	82.6
68	78.9	44.9	34	1.32	2.54	1.2	813.3	7.9	65.6	46.10	111.8
73	82.4	46.1	36.3	1.27	1.17	1.8	958.8	19.7	185.4	37.00	859.4
82	72.3	43.8	28.5	1.54	0.5	1.4	458.4	22.1	82.1	48.50	350.3
84	82.6	46.2	36.4	1.27	1.43	2.7	923.6	22.9	107.6	46.40	176
93	82.1	47.5	34.6	1.37	1.46	2.6	262	8.6	59.9	47.10	105.2
96	76.5	46.8	29.7	1.58	0.5	1.1	574.6	20.9	79.4	101.80	203.7
Mean uninfected	78.73	47.15	31.58	1.518	1.226	1.81	763.9	18.71	97.55	60.85	227.53
Mean infected	77.7	45.22	32.48	1.406	1.288	1.9	540.96	12.72	77.58	54.64	236
STdv uninfected	6.52	2.63	4.71	0.19	0.68	0.52	391.61	4.62	35.31	27.40	239.64
STdv infected	4.56	1.99	3.80	0.17	0.78	0.62	311.19	5.71	38.77	18.63	289.40

No indication of hepatitis or impaired liver function was observed; no significant differences between mean values for infected (bold) or non-infected bats were apparent. The following parameters were analyzed: Total Protein (T. Prot); Albumin (Alb), Globulin (Glob), Alb:Glob ratio, Total Cholesterol (T. Chol); Total Bilirubin (T. Bili); Alkaline Phosphatase (Alk. Phos); Alanine Transferase (ALT); Aspartate Aminotransferase (AST); Gamma Glutamyltransferase (GGT) and Lactate Dehydrogenase (LDH).
doi:10.1371/journal.ppat.1000972.t002

first report of an unbiased molecular approach to pathogen discovery in this important reservoir of emerging infectious diseases. The modest yield of novel microbial sequences may reflect the choice of sample (e.g., serum vs feces, tissue or another specimen), competition between host and microbial template during unbiased amplification, or both. Efforts to address template competition are under way that include subtraction of host nucleic acids or the use of semi-random primers that do not amplify host sequences. Such efforts will likely enhance the sensitivity and throughput of unbiased sequencing technologies for pathogen discovery.

The discovery of this chiropteran flavivirus broadens both the taxonomical and geographical distribution of GB-like viruses. Three types of GB viruses have been described: GBV-A, -B and -C [18,19,24,25,54,57]. GBV-B, which has never been found in humans and was only reported in captive tamarins after serial passage of the original human GB serum [58], is most closely related to HCV and was recently classified together with HCV into a new genus, *Hepacivirus*, within the family *Flaviviridae* [16]. GBV-A and -C remain unclassified members of the family. GBV-A have been isolated from several New World monkeys. Different genotypes appear to be associated with specific monkey species of the genera *Saguinus*, *Callithrix* (*Callitrichidae* family) and *Aotus* (*Aotidae* family), without any clinical signs associated with infection [24,54,57]. GBV-C have been isolated from humans with non-A-E hepatitis; however, its pathogenicity is unknown and the virus is widespread in the human population [21,59–61]. Population studies showed that GB viruses are enzootic and species-specific within both Old and New World nonhuman primates as well as humans, and have likely co-evolved with their hosts over long periods of time [62]. Previously, the only GBV found in the Old world was GBV-C from chimpanzees (in Africa) and humans. Although GBV-C were found in humans, GB viruses have not been previously reported in primates or other animals on the Indian subcontinent.

GBV-C and -A are remarkable for a truncated or missing capsid (C) protein [18,19]. Due to exhaustion of our samples we were unable to complete assessment of the 5'-terminal sequence; nonetheless, RACE experiments suggest that GBV-D likely codes for a short basic peptide, instead of a full-length C protein. The first methionine (M₁) predicts a peptide of 57 aa (pI = 12); however, the more favorable Kozak context [63] of M₃ indicates a 55 aa peptide. After signalase cleavage from the polyprotein precursor, this peptide may be functional, possibly influencing maturation of, or directly binding to, the E1 and/or E2 glycoproteins.

Phylogenetic analyses of NS5B, NS3 and complete polyprotein sequence place GBV-D at the root of the GBV-A and -C clades and are consistent with a model wherein GBV-D is ancestral to GBV-A and -C clades. Mixed relationships indicative of recombination events [64] were not evident (**Figure 2, Figure 4**). Both pteropid bats and chimpanzees are restricted to the Old World. While the range of chimpanzees (Africa) and *P. giganteus* (the Indian subcontinent) do not overlap, it is possible that other primate species in Bangladesh or India, such as macaques, or other fruit bats in Africa such as *Eidolon spp.*, whose range overlaps that of

chimpanzees, may carry related viruses. While GBV-A is only known from primates of the New World, an African origin has been suggested for GBV-C based on a 12-aa indel sequence in NS5A [65]. Although the NS5A sequence of GBV-D, similar to that of GBV-A, appears elongated in the indel region, compatible with their respective earlier phylogenetic branching compared to GBV-C, little sequence conservation is observed in that region.

The bats in this study, like primates infected with their associated GBV [66], all appeared to be healthy. The lack of chemical evidence of hepatic inflammation or dysfunction suggests that this virus may not target hepatic cells in bats. This is consistent with the behavior of GBV-A in its natural primate hosts [54]. In contrast, elevated alanine transferase levels and mild hepatitis are observed in experimental infections of macaques with GBV-C isolates from humans [67]. Five percent of the bats we studied were infected with one of at least two different strains of GBV-D, which suggests widespread viral circulation within this species. The observation that bats are asymptotically infected with diverse strains that constitute a distinct phylogenetic clade is compatible with a co-evolutionary relationship between GBV and their hosts [57,62], and supports the hypothesis that *P. giganteus* bats may be a natural reservoir for GBV-D. In one case we were able to detect GBV-D nucleic acid in saliva. This suggests a potential route for viral transmission via fighting or grooming behavior, or via food shared by bats.

Pteropus giganteus is a frugivorous bat species that carries NiV, a zoonotic paramyxovirus [10,11]. This species lives in close association with humans in Bangladesh and bats have been observed drinking from (and urinating into) date palm sap collecting pots [14]. Human consumption of contaminated palm juice is proposed to be a major route of NiV transmission [68]. Although it is unclear whether infectious virus was present in bat saliva, the observation that saliva can contain GBV-D nucleic acids provides a biologically plausible mechanism for transmission from infected bats to other hosts. While it is currently unknown whether GBV-D virus occurs in humans, up to 20% of non-A-E hepatitis cases remain unexplained [19].

Acknowledgments

We thank the Forestry Department of Bangladesh for permission to conduct this research; Md. Sheikh Gofur and Md. Pitu Biswas for help in sampling bats; A. Bennett, A. Tashmukhamedova, and R. Tokarz for technical support, and K. Olival for critical comments on the manuscript.

Author Contributions

Conceived and designed the experiments: JHE TB PD WIL. Performed the experiments: JHE PLQ DV SH. Analyzed the data: JHE PLQ TB CS OJ DV SKH ME PD. Contributed reagents/materials/analysis tools: JHE TB CS SC SAK MJH SKH ME SPL WIL. Wrote the paper: JHE PLQ TB WIL. Coordinated field and lab activities and logistics for work in Bangladesh; contributed to paper-writing: MJH. Coordinated permissions, field activities and logistics for work in Bangladesh; contributed to paper-writing: SPL.

References

- Calisher CH, Childs JE, Field HE, Holmes KV, Schountz T (2006) Bats: Important reservoir hosts of emerging viruses. *Clin Microbiol Rev* 19: 531–538.
- Towner JS, Pourrut X, Albarino CG, Nkogue CN, Bird BH, et al. (2007) Marburg virus infection detected in a common African bat. *PLoS ONE* 2: e764, 761–765.
- Epstein JH, Field HE, Luby S, Pulliam JRC, Daszak P (2006) Nipah virus: Impact, Origins, and Causes of Emergence. *Curr Infect Dis Rep* 8: 59–65.
- Li W, Shi Z, Yu M, Ren W, Smith C, et al. (2005) Bats are natural reservoirs of SARS-like coronaviruses. *Science* 310: 676–679.
- Leroy EM, Kumulungui B, Pourrut X, Rouquet P, Hassanin A, et al. (2005) Fruit bats as reservoirs of Ebola virus. *Nature* 438: 575–576.
- Chua K, Bellini W, Rota P, Harcourt B, Tamin A, et al. (2000) Nipah virus: A recently emergent deadly paramyxovirus. *Science* 288: 1432–1435.
- Selvey L, Wells RM, McCormack JG, Ansford AJ, Murray PK, et al. (1995) Infection of humans and horses by a newly described morbillivirus. *Med J Aust* 162: 642–645.
- McColl K, Tordo N, Aguilar-Setien A (2000) Bat lyssavirus infections. *Rev sci tech Off int Epiz* 19: 177–196.
- Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, et al. (2008) Global trends in emerging infectious diseases. *Nature* 451: 990–U994.

10. Epstein JH, Prakash V, Smith CS, Daszak P, McLaughlin AB, et al. (2008) Henipavirus infection in fruit bats (*Pteropus giganteus*), India. *Emerg Infect Dis* 14: 1309–1311.
11. Hsu VP, Hossain MJ, Parashar UD, Ali MM, Ksiazek TG, et al. (2004) Nipah virus encephalitis reemergence, Bangladesh. *Emerg Infect Dis* 10: 2082–2087.
12. Chadha MS, Comer JA, Lowe L, Rota PA, Rollin PE, et al. (2006) Nipah virus-associated encephalitis outbreak, Siliguri, India. *Emerg Infect Dis* 12: 235–240.
13. Gurley ES, Montgomery JM, Hossain MJ, Bell M, Azad AK, et al. (2007) Person-to-person transmission of Nipah virus in a Bangladeshi community. *Emerg Infect Dis* 13: 1031–1037.
14. Luby SP, Hossain MJ, Gurley ES, Ahmed BN, Banu S, et al. (2009) Recurrent Zoonotic Transmission of Nipah Virus into Humans, Bangladesh, 2001–2007. *Emerg Infect Dis* 15: 1229–1235.
15. Luby S, Rahman M, Hossain MJ, Blum LS, Husain NM, et al. (2006) Foodborne Transmission of Nipah Virus, Bangladesh. *Emerg Infect Dis* 12.
16. Theil HJ (2005) Family Flaviviridae. In: Claude M Fauquet, MA Mayo, J Maniloff, U Desselberger, Ball LA, eds. *Virus Taxonomy: VIIIth Report of the International Committee on Taxonomy of Viruses*. 2 ed. San Diego: Academic Press. pp 979–996.
17. Erker JC, Desai SM, Leary TP, Chalmers ML, Montes CC, et al. (1998) Genomic analysis of two GB virus A variants isolated from captive monkeys. *J Gen Virol* 79: 41–45.
18. Leary TP, Desai SM, Erker JC, Mushahwar IK (1997) The sequence and genomic organization of a GB virus A variant isolated from captive tamarins. *J Gen Virol* 78: 2307–2313.
19. Leary TP, Muerhoff AS, Simons JN, PilotMatias TJ, Erker JC, et al. (1996) Sequence and genomic organization of GBV-C: A novel member of the flaviviridae associated with human non-A-E hepatitis. *J Med Virol* 48: 60–67.
20. Linnen J, Wages J, ZhangKeck ZY, Fry KE, Krawczynski KZ, et al. (1996) Molecular cloning and disease association of hepatitis G virus: A transfusion-transmissible agent. *Science* 271: 505–508.
21. Kao JH, Chen PJ, Chen DS (1996) GBV-C in the aetiology of fulminant hepatitis. *Lancet* 347: 120–120.
22. Kar P, Mukhopadhyay S, Gopalkrishna V, Das BC (2000) Infection with hepatitis G-virus and viral hepatitis in India. *Curr Sci* 78: 189–194.
23. Kondo Y, Mizokami M, Nakano T, Kato T, Ohba K, et al. (1997) Genotype of GB virus C hepatitis G virus by molecular evolutionary analysis. *Virus Res* 52: 221–230.
24. Adams NJ, Prescott LE, Jarvis LM, Lewis JCM, McClure MO, et al. (1998) Detection in chimpanzees of a novel flavivirus related to GB virus-C hepatitis G virus. *J Gen Virol* 79: 1871–1877.
25. Birkenmeyer LG, Desai SM, Muerhoff AS, Leary TP, Simons JN, et al. (1998) Isolation of a GB virus-related genome from a chimpanzee. *J Med Virol* 56: 44–51.
26. Palacios G, Quan PL, Jabado OJ, Conlan S, Hirschberg DL, et al. (2007) Panmicrobial oligonucleotide array for diagnosis of infectious diseases. *Emerg Infect Dis* 13: 73–81.
27. Quan PL, Palacios G, Jabado OJ, Conlan S, Hirschberg DL, et al. (2007) Detection of respiratory viruses and subtype identification of influenza A viruses by GreeneChipResp oligonucleotide microarray. *J Clin Microbiol* 45: 2359–2364.
28. Cox-Foster DL, Conlan S, Holmes EC, Palacios G, Evans JD, et al. (2007) A metagenomic survey of microbes in honey bee colony collapse disorder. *Science* 318: 283–287.
29. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, et al. (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437: 376–380.
30. Korbie DJ, Mattick JS (2008) Touchdown PCR for increased specificity and sensitivity in PCR amplification. *Nat Protoc* 3: 1452–1456.
31. Pilhofer M, Bauer AP, Schrällhammer M, Richter L, Ludwig W, et al. (2007) Characterization of bacterial operons consisting of two tubulins and a kinesin-like gene by the novel Two-Step Gene Walking method. *Nucleic Acids Res* 35: 8.
32. Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, et al. (2003) Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res* 31: 3497–3500.
33. Drummond AJ AB, Cheung M, Heled J, Kearse M, Moir R, Stones-Havas S, Thierer T, Wilson A (2009) Gencious v4.6, Available from <http://www.gencious.com/>.
34. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol & Evol* 24: 1596–1599.
35. Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, et al. (1999) Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J Virol* 73: 152–160.
36. Chambers TJ, Hahn CS, Galler R, Rice CM (1990) Flavivirus genome organization, expression, and replication. *Ann Rev Microbiol* 44: 649–688.
37. Lin C, Lindenbach BD, Pragai BM, McCourt DW, Rice CM (1994) Processing in the Hepatitis-C virus E2-N2 region - identification of P7 and 2 Distinct E2-specific products with different C-termini. *J Virol* 68: 5063–5073.
38. Mizushima H, Hijikata M, Asabe SI, Hirota M, Kimura K, et al. (1994) 2 Hepatitis-C virus glycoprotein E2 products with different C-termini. *J Virol* 68: 6215–6222.
39. Griffin S, Clarke D, McCormick C, Rowlands D, Harris M (2005) Signal peptide cleavage and internal targeting signals direct the hepatitis C virus p7 protein to distinct intracellular membranes. *J Virol* 79: 15525–15536.
40. Ghibaud D, Cohen L, Penin F, Martin A (2004) Characterization of GB virus B polyprotein processing reveals the existence of a novel 13-kDa protein with partial homology to hepatitis C virus p7 protein. *J Biol Chem* 279: 24965–24975.
41. Takikawa S, Engle RE, Emerson SU, Purcell RH, St Claire M, et al. (2006) Functional analyses of GB virus B p13 protein: Development of a recombinant GB virus B hepatitis virus with a p7 protein. *PNAS* 103: 3345–3350.
42. Belyaev AS, Chong S, Novikov A, Kongpachith A, Masiarz FR, et al. (1998) Hepatitis G virus encodes protease activities which can effect processing of the virus putative nonstructural proteins. *J Virol* 72: 868–872.
43. Ferron F, Bussetta C, Dutartre H, Canard B (2005) The modeled structure of the RNA dependent RNA polymerase of GBV-C Virus suggests a role for motif E in Flaviviridae RNA polymerases. *Bmc Bioinformatics* 6.
44. Koonin EV (1991) The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. *J Gen Virol* 72: 2197–2206.
45. Poch O, Sauvaget I, Delarue M, Tordo N (1989) Identification of 4 conserved motifs among the RNA-dependent polymerase encoding elements. *Embo Journal* 8: 3867–3874.
46. Muller R, Poch O, Delarue M, Bishop DHL, Bouloy M (1994) Rift-Valley Fever virus L-segment - correction of the sequence and possible functional role of newly identified regions conserved in RNA-dependent polymerases. *J Gen Virol* 75: 1345–1352.
47. Huang Y, Staschke K, De Francesco R, Tan SL (2007) Phosphorylation of hepatitis C virus NS5A nonstructural protein: A new paradigm for phosphorylation-dependent viral RNA replication? *Virology* 364: 1–9.
48. Nanda SK, Herion D, Liang TJ (2006) Src homology 3 domain of hepatitis C virus NS5A protein interacts with Bin1 and is important for apoptosis and infectivity. *vol 130, pg 794, 2006. Gastroenterology* 131: 687–687.
49. Tan SL, Katze MG (2001) How hepatitis C virus counteracts the interferon response: The jury is still out on NS5A. *Virology* 284: 1–12.
50. Tan SL, Nakao H, He YP, Vijaysri S, Neddermann P, et al. (1999) NS5A, a nonstructural protein of hepatitis C virus, binds growth factor receptor-bound protein 2 adaptor protein in a Src homology 3 domain/ligand-dependent manner and perturbs mitogenic signaling. *PNAS* 96: 5533–5538.
51. Dumont S, Cheng W, Serebrov V, Beran RK, Tinoco I, et al. (2006) RNA translocation and unwinding mechanism of HCVNS3 helicase and its coordination by ATP. *Nature* 439: 105–108.
52. Scarselli E, Urbani A, Sbardellati A, Tomei L, DeFrancesco R, et al. (1997) GB virus B and hepatitis C virus NS3 serine proteases share substrate specificity. *J Virol* 71: 4985–4989.
53. Bartenschlager R (1999) The NS3/4A proteinase of the hepatitis C virus: unravelling structure and function of an unusual enzyme and a prime target for antiviral therapy. *J Viral Hepatitis* 6: 165–181.
54. Schlauder GG, Dawson GJ, Simons JN, Pilotmatias TJ, Gutierrez RA, et al. (1995) Molecular and serologic analysis in the transmission of the GB hepatitis agents. *J Med Virol* 46: 81–90.
55. Dufour DR, Lott JA, Nolte FS, Gretch DR, Koff RS, et al. (2000) Diagnosis and monitoring of hepatic injury. I. Performance characteristics of laboratory tests; 2000. *Amer Assoc Clinical Chemistry*. pp 2027–2049.
56. Heard DJ, De Young JL, Goodyear B, Ellis GA (1997) Comparative rectal bacterial flora of four species of flying fox (*Pteropus* sp.). *J Zoo Wild Med* 28: 471–475.
57. Bukh J, Apparg CL (1997) Five new or recently discovered (GBV-A) virus species are indigenous to New World monkeys and may constitute a separate genus of the Flaviviridae. *Virology* 229: 429–436.
58. Deinhard F, Holmes AW, Capps RB, Popper H (1967) Studies on transmission of human viral hepatitis to marmoset monkeys. I. Transmission of disease serial passages and description of liver lesions. *J Exp Med* 125: 673–&.
59. Alter MJ, Gallagher M, Morris TT, Moyer LA, Meeks EL, et al. (1997) Acute non-A-E hepatitis in the United States and the role of hepatitis G virus infection. *New Engl J Med* 336: 741–746.
60. Stapleton JT (2003) GB virus type C/hepatitis G virus. *Seminars in Liver Disease* 23: 137–148.
61. Sheng WH, Hung CC, Wu RJ, Wang JT, Chen PJ, et al. (2007) Clinical impact of GB virus C viremia in patients with HIV type 1 infection in the era of highly active antiretroviral therapy. *Clin Infect Dis* 44: 584–590.
62. Charrel RN, De Micco P, de Lamballerie X (1999) Phylogenetic analysis of GB viruses A and C: evidence for cospeciation between virus isolates and their primate hosts. *J Gen Virol* 80: 2329–2335.
63. Kozak M (1999) Initiation of translation in prokaryotes and eukaryotes. *Gene* 234: 187–208.
64. Worobey M, Holmes EC (2001) Homologous recombination in GB virus C/hepatitis G virus. *Mol Biol & Evol* 18: 254–261.
65. Tanaka Y, Mizokami M, Orito E, Ohba K, Kato T, et al. (1998) African origin of GB virus C hepatitis G virus. *FEBS Lett* 423: 143–148.
66. Leary TP, Desai SM, Yamaguchi J, Chalmers ML, Schaluder GG, et al. (1997) Species-specific variants of GB virus A in captive monkeys (vol 70, pg 9028, 1996). *J Virol* 71: 8953–8953.
67. Cheng Y, Zhang WZ, Li J, Li BA, Zhao JM, et al. (2000) Serological and histological findings in infection and transmission of GBV-C/HGV to macaques. *J Med Virol* 60: 28–33.
68. Luby SP, Gurley ES, Hossain MJ (2009) Transmission of Human Infection with Nipah Virus. *Clin Infect Dis* 49: 1743–1748.