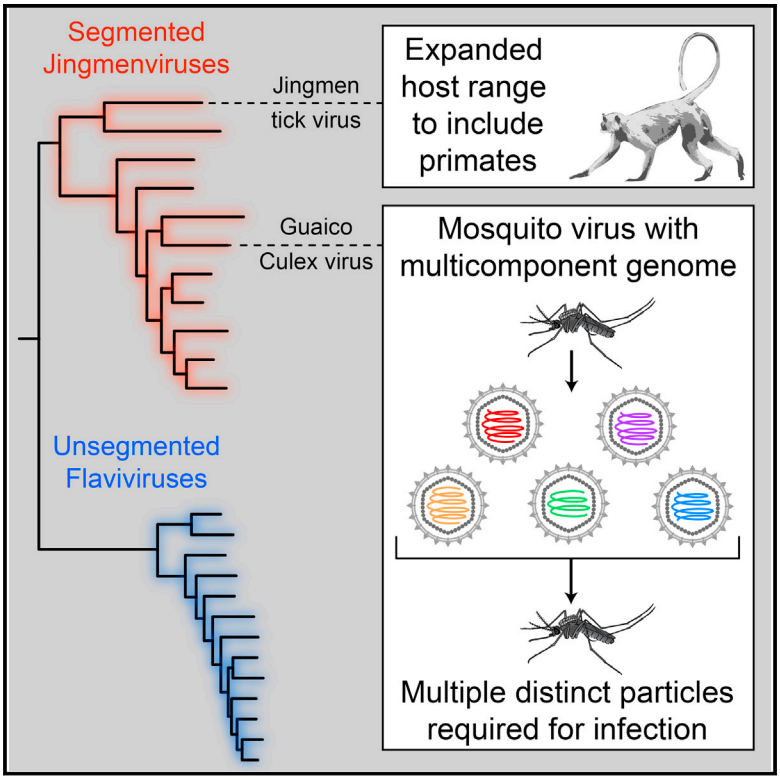
Article

Cell Host & Microbe

A Multicomponent Animal Virus Isolated from Mosquitoes

Graphical Abstract



Highlights

- Multicomponent viruses, formerly seen only in plants/fungi, can infect animal hosts
- GCXV is an enveloped, multicomponent virus that infects mosquitoes
- GCXV has five segments, one optional for replication and variably present in isolates
- An evolutionarily related, segmented virus was observed in a non-human primate host

Authors

Jason T. Ladner, Michael R. Wiley, Brett Beitzel, ..., Laura D. Kramer, Robert B. Tesh, Gustavo Palacios

Correspondence

jason.t.ladner.ctr@mail.mil (J.T.L.), gustavo.f.palacios.ctr@mail.mil (G.P.)

In Brief

Multicomponent viruses, which separately package different genome segments, were thought to be restricted to plant and fungal hosts. Ladner et al. characterize a multicomponent mosquito virus and describe an evolutionarily related, segmented virus in a nonhuman primate. These findings provide evidence for multicomponent animal viruses and suggest relevance to animal health.

Accession Numbers

KM521552–KM521574
KM461666–KM461670
KX377513–KX377516

Ladner et al., 2016, Cell Host & Microbe *20*, 357–367
September 14, 2016
http://dx.doi.org/10.1016/j.chom.2016.07.011

Cell Host & Microbe
Article

A Multicomponent Animal Virus Isolated from Mosquitoes

Jason T. Ladner,[1,*] Michael R. Wiley,[1] Brett Beitzel,[1] Albert J. Auguste,[2] Alan P. Dupuis II,[3] Michael E. Lindquist,[4] Samuel D. Sibley,[7] Krishna P. Kota,[5] David Fetterer,[6] Gillian Eastwood,[3] David Kimmel,[1] Karla Prieto,[1] Hilda Guzman,[2] Matthew T. Aliota,[3,17] Daniel Reyes,[1] Ernst E. Brueggemann,[5] Lena St. John,[1] David Hyeroba,[13] Michael Lauck,[14,16] Thomas C. Friedrich,[7,14] David H. O'Connor,[14,16] Marie C. Gestole,[1,18] Lisa H. Cazares,[5,8,9,10] Vsevolod L. Popov,[2] Fanny Castro-Llanos,[11] Tadeusz J. Kochel,[11,18] Tara Kenny,[5] Bailey White,[1] Michael D. Ward,[5] Jose R. Loaiza,[12] Tony L. Goldberg,[7,13,14] Scott C. Weaver,[2] Laura D. Kramer,[3,15] Robert B. Tesh,[2] and Gustavo Palacios[1,19,*]
[1]Center for Genome Sciences, U.S. Army Medical Research Institute of Infectious Diseases, Fort Detrick, MD 21702, USA
[2]Institute for Human Infections and Immunity, Departments of Pathology, Microbiology & Immunology, and Center for Biodefense and Emerging Infectious Diseases, University of Texas Medical Branch, Galveston, TX 77555, USA
[3]Arbovirus Laboratory, Wadsworth Center, New York State Department of Health, Albany, NY 12159, USA
[4]Virology Division, U.S. Army Medical Research Institute of Infectious Diseases, Fort Detrick, MD 21702, USA
[5]Molecular and Translational Sciences Division, U.S. Army Medical Research Institute of Infectious Diseases, Fort Detrick, MD 21702, USA
[6]Research Support Division, U.S. Army Medical Research Institute of Infectious Diseases, Fort Detrick, MD 21702, USA
[7]Department of Pathobiological Sciences, University of Wisconsin-Madison, Madison, WI 53706, USA
[8]Henry M. Jackson Foundation, Bethesda, MD 20817, USA
[9]DoD Biotechnology High Performance Computing Software Applications Institute, Frederick, MD 21702, USA
[10]Telemedicine and Advanced Technology Research Center, U.S. Army Medical Research and Materiel Command, Fort Detrick, MD 21702, USA
[11]U.S. Naval Medical Research Unit No. 6, Lima, Peru
[12]Centro de Biodiversidad y Descubrimiento de Drogas, Instituto de Investigaciones Científicas y Servicios de Alta Tecnología, Ciudad de Panamá, Panamá
[13]Makerere University, Kampala, Uganda
[14]Wisconsin National Primate Research Center, Madison, WI 53715, USA
[15]School of Public Health, State University of New York at Albany, One University Place Rensselaer, East Greenbush, NY 12144, USA
[16]Department of Pathology and Laboratory Medicine, University of Wisconsin School of Medicine and Public Health, Madison, WI 53706, USA
[17]Present address: Department of Pathobiological Sciences, University of Wisconsin-Madison, Madison, WI 53706, USA
[18]Present address: National Biodefense Analysis and Countermeasures Center, Fort Detrick, MD 21702, USA
[19]Lead Contact
*Correspondence: jason.t.ladner.ctr@mail.mil (J.T.L.), gustavo.f.palacios.ctr@mail.mil (G.P.)
http://dx.doi.org/10.1016/j.chom.2016.07.011

SUMMARY

RNA viruses exhibit a variety of genome organization strategies, including multicomponent genomes in which each segment is packaged separately. Although multicomponent genomes are common among viruses infecting plants and fungi, their prevalence among those infecting animals remains unclear. We characterize a multicomponent RNA virus isolated from mosquitoes, designated Guaico Culex virus (GCXV). GCXV belongs to a diverse clade of segmented viruses (Jingmenvirus) related to the prototypically unsegmented *Flaviviridae*. The GCXV genome comprises five segments, each of which appears to be separately packaged. The smallest segment is not required for replication, and its presence is variable in natural infections. We also describe a variant of Jingmen tick virus, another Jingmenvirus, sequenced from a Ugandan red colobus monkey, thus expanding the host range of this segmented and likely multicomponent virus group. Collectively, this study provides evidence for the existence of multicomponent animal viruses and their potential relevance for animal and human health.

INTRODUCTION

The diversity of genome organizations seen in RNA viruses is truly exceptional, surpassing that of any other group of organisms (Holmes, 2009). Differences are seen in the nature of the genetic material (single or double stranded, positive or negative sense, linear or circular), the number of genome segments (from 1–12), and the manner in which multi-segmented genomes are packaged (together or separately) (King et al., 2011). These differences can have important functional implications for key processes such as gene expression, transmission, and genetic recombination. Genome segmentation, for example, can allow better control over gene expression by creating multiple distinct transcriptional units (Holmes, 2009); however, if these segments are separately packaged, a higher multiplicity of infection will be needed for successful transmission (Goldbach, 1986).

Although great strides have been made in cataloging viral diversity, the evolutionary mechanisms that have generated this

Table 1. Viruses Sequenced in This Study

ID	Species	Source	Locality	Date	GenBank Accession Numbers
TR7094	GCXV	*Culex declarator*	Aripo, Trinidad	11/26/2008	KM521571–KM521574
LO35	GCXV	*Culex coronator*	Loreto, Peru	2/21/2009	KM461666–KM461670
LO47	GCXV	*Culex coronator*	Loreto, Peru	2/22/2009	KM521561–KM521565
GAM204	GCXV	*Culex coronator*	Soberania, Panama	12/7/2012	KM521556–KM521560
PCR18-229	GCXV	*Culex coronator*	Achiotes, Panama	1/2012	KM521566–KM521570
ACH27	GCXV	*Culex interrogator*	Achiotes, Panama	10/2012–1/2013[a]	KM521552–KM521555
RC27	JMTV	*Procolobus rufomitratus*	Kibale National Park, Uganda	2/2/2012	KX377513–KX377516

[a]For ACH27, *C. interrogator* mosquitos were pooled from several collections conducted at different times.

extraordinary variety of genomic organizations are still poorly understood. This is due in part to the extensive genetic divergence that exists between most viruses with different organizations (typically present in distinct families), thus preventing meaningful sequence-based evolutionary comparisons. Occasionally, though, more recent transitions are uncovered, and analysis of these events provides insight into the macroevolution of RNA viruses. Qin et al. (2014) reported just such a connection between segmented and unsegmented RNA viruses by describing a segmented virus (Jingmen tick virus [JMTV]) with sequence homology to the *Flaviviridae*, a large family of vertebrate and invertebrate viruses (including a number of important human pathogens, e.g., Zika, yellow fever, West Nile, dengue, Japanese encephalitis, and hepatitis C viruses) with unsegmented, positive-sense RNA genomes.

Here, we expand upon this finding by describing a genetically distinct, segmented virus isolated from mosquitoes that also exhibits homology to viruses in the family *Flaviviridae* and that appears to be multicomponent (also termed multiparticle or multipartite; Holmes, 2009; Mahy, 2009; Reijnders, 1978), with each genome segment separately packaged into virions. Although multicomponent genomes are relatively common among RNA viruses that infect plants and fungi, this method of genome organization has not previously been seen in animal viruses (Fulton, 1980; King et al., 2011). This virus has tentatively been designated Guaico Culex virus (GCXV), on the basis of the first collection location (near the Guaico community in Trinidad) and the genus of the mosquito that appears to serve as its primary host. Through phylogenetic analysis, we demonstrate that GCXV and JMTV both belong to a highly diverse clade of segmented (and likely multicomponent) viruses, which has recently been termed the Jingmenvirus group (Shi et al., 2015). We also report the detection of a variant of JMTV in a red colobus monkey in Uganda, thus expanding the host range of Jingmenviruses to include primates and highlighting the potential relevance of these viruses to animal and human health.

RESULTS

Genome Characterization of GCXV

We sequenced six isolates of GCXV from *Culex* spp. mosquitoes collected in three countries in Central America and South America (Table 1). Complete genomes were obtained for the isolates from Peru (LO35, LO47) and Trinidad (TR7094); coding-complete (i.e., only missing pieces of non-coding, untranslated regions [UTRs] [Ladner et al., 2014]) genomes were obtained for the isolates from Panama (GAM204, PCR18-229, ACH27). Five genome segments were assembled for four of the isolates (LO35, LO47, GAM204, and PCR18-229), resulting in a total genome size of ~12 kb. RNA extracted from purified GCXV particles confirmed the presence of a segmented genome (Figure 1A).

However, only four segments were assembled for ACH27 and TR7094 (genome size ~10.6 kb). For these two isolates, the four assembled segments corresponded to the four largest segments assembled in the other isolates (Figure 1, segments 1–4). The mosquito pools for both ACH27 and TR7094 contained multiple viruses capable of replicating in mosquito cells, so we were unable to obtain pure cultures for these isolates (Auguste et al., 2014). However, despite the complexity of these pools, very high sequencing depth was obtained for the four assembled segments (ACH27: 16,700–40,500×; TR7094: 970–11,680×). TR7094 contained very low levels of LO47 contamination (all five segments), but for ACH27, no Illumina reads aligned to the segment 5 sequences from the other isolates, and no contigs from ACH27 exhibited significant similarity to the segment 5 sequences from the other isolates. Therefore, segment 5 appears to be absent from ACH27 and TR7094.

Nuclease digestion assays (Figure 1B), along with 5′ and 3′ rapid amplification of cDNA ends (RACE), confirmed that all five genome segments were single-stranded, positive-sense RNA (ssRNA$^+$). Among the GCXV isolates, pairwise nucleotide (nt) divergences, calculated separately for each segment, ranged from 0.2% to 19.9% (Table S4). The phylogenies inferred from segments 2–4 all exhibited essentially identical patterns with isolates clustered by geographic location and with TR7094 as the outlier (Figure 2B). However, on segment 1, TR7094 exhibited lower than expected relative levels of divergence and grouped closely with the Panamanian isolates. This is likely indicative of a segment 1 reassortment event having occurred on the lineage leading to TR7094. The phylogeny inferred from segment 5 is also inconsistent with those of the other segments. In addition to the absence of this segment in TR7094 and ACH27, this segment exhibited very low levels of nt divergence (0.4%–2.3%), with most sequence variations only present in a single isolate.

Rescue through Reverse Genetics

A reverse genetics system was established to test the ability of GCXV to replicate in the absence of segment 5. C6/36 cells were separately transfected with three combinations of in vitro transcribed RNA segments: (1) all five segments, (2) segments

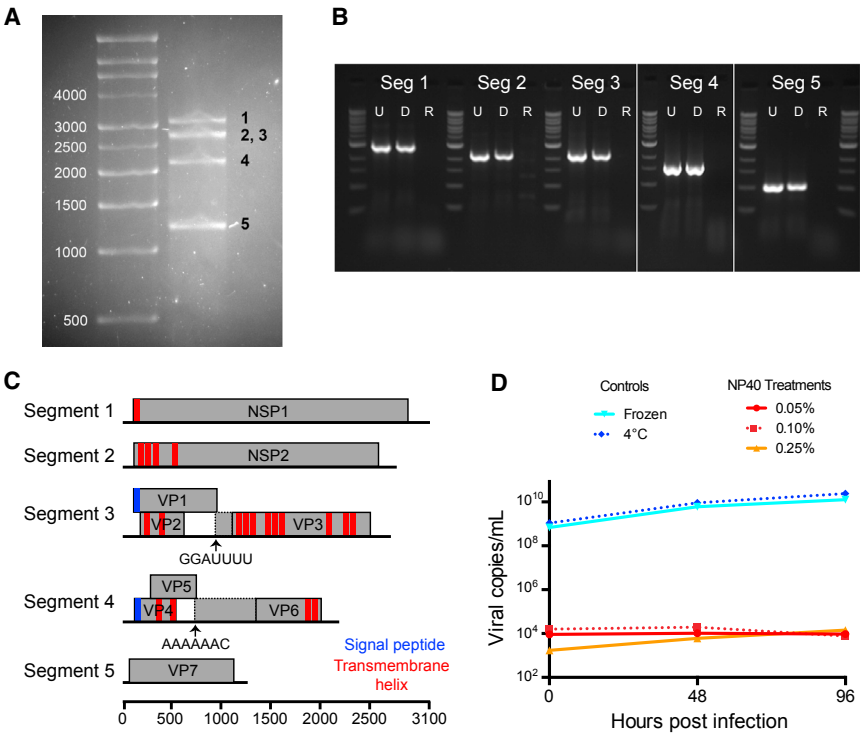


Figure 1. The Genome of GCXV Is Enveloped and Divided into Five Single-Stranded Positive-Sense RNA Segments

(A) Distinct RNA segments of GCXV on an agarose gel. Segment numbers are shown in black; ladder band sizes (nt) are shown in white.

(B) Selective degradation with RNase I demonstrates that GCXV has a single-stranded RNA genome. U, untreated; D, DNase I treated; R, RNase I treated. Segment 4 amplicons were run on a separate gel.

(C) Genome schematic with all ORFs ≥ 400 nt. Dotted lines indicate regions putatively translated through −1 ribosomal frameshifting (arrows indicate slippery heptanucleotides); solid lines indicate the predicted ORFs based on the first conserved AUG.

(D) Treatment with NP40 resulted in highly diminished RNA copy numbers, lack of growth, and absence of CPE in C6/36 cells, indicating that GCXV is enveloped.

See also Figures S2 and S3 and Tables S2 and S3.

1–4, and (3) segments 2–5 (i.e., no RNA-dependent RNA polymerase [RdRp]; used as a negative control). No cytopathic effect (CPE) was detected in the negative control (segments 2–5), and quantitative RT-PCR (qRT-PCR) confirmed the absence of replication. CPE was observed in the other two transfections (segments 1–5 and 1–4), and RNA extracted after post-transfection passages confirmed the establishment of successful infections both with and without segment 5 (Figure 2A).

Segment Packaging

Multicomponent plant viruses were recognized on the basis of deviations from the expected relationship between infectious dose and the number of lesions on infected leaves (i.e., exhibiting multi-hit rather than single-hit kinetics; Flint et al., 2009; Sánchez-Navarro et al., 2013). We used a similar approach to assay the nature of segment packaging for GCXV using cell culture plaques instead of leaf lesions. The dose-response curve for GCXV differed significantly from expectations for a single-component virus (i.e., the number of plaques decreased more quickly than expected with dilution of the inoculant) (Figure 3) (Flint et al., 2009). Assuming the presence of distinct particles (each containing a subset of genome segments) present in similar amounts and with the same likelihood of invading a cell, we used our dose-response curve to estimate the presence of 3.27 ± 0.37 distinct GCXV particles required for plaque formation. Deviations from this assumption (e.g., variation in the abundance and/or probability of cell invasion for different particle types) would lead to a shallower slope and underestimation of the number of distinct particle types. Therefore, this result is consistent with three to five segments being required for plaque formation, assuming that each segment is separately packaged.

Visualizing Transcription/Replication

Segment-specific probe sets were used to visualize GCXV transcription/replication within C6/36 cells. We detected viral RNA using two combinations of probes: (1) segments 1–3 and (2) segments 3–5. At least one genome segment was detected in ≥90% of the assayed cells, and in general, when multiple segments were present, they appeared to be colocalized in the cytoplasm (Figure 4). Segments 1–3 were detected together in all positive cells, whereas segments 4 and 5 were variably present within cells in which segment 3 was detected (Figure 4; Table S1). The detection of different combinations of segments within individual cells is consistent with independent packaging of genome segments. Although most segments are likely required to successfully complete a full cycle of infection, only a subset of segments are likely required for viral transcription/replication. However, we cannot rule out the presence of undetected segments below our limit of detection.

UTRs

None of the segments of GCXV were polyadenylated; however, sequence conservation across segments was seen in the UTRs (Figure S2). The 5′ UTRs exhibited several highly conserved sequence motifs, including the nine most terminal nt, which were strictly conserved across all segments and isolates (5′-AAAUUAAAA-3′), and a larger region just downstream, which is predicted to form a 28- to 31-base stem-loop structure (Figure S2A). The 3′ UTRs also exhibited regions of sequence conservation across segments, particularly at the 3′ terminus. Although some sequence variation was seen on segment 3, the last seven nt were highly conserved across segments and isolates, with a consensus sequence of 5′-CCCAUUU-3′. Notably, the four most terminal nt at the 5′ and 3′ ends were complementary. A second highly conserved motif was seen in the 3′ UTRs of segments 1–4 (5′-AAWUAC-3′). This is predicted to form the distal loop in a conserved stem-loop, although the exact size

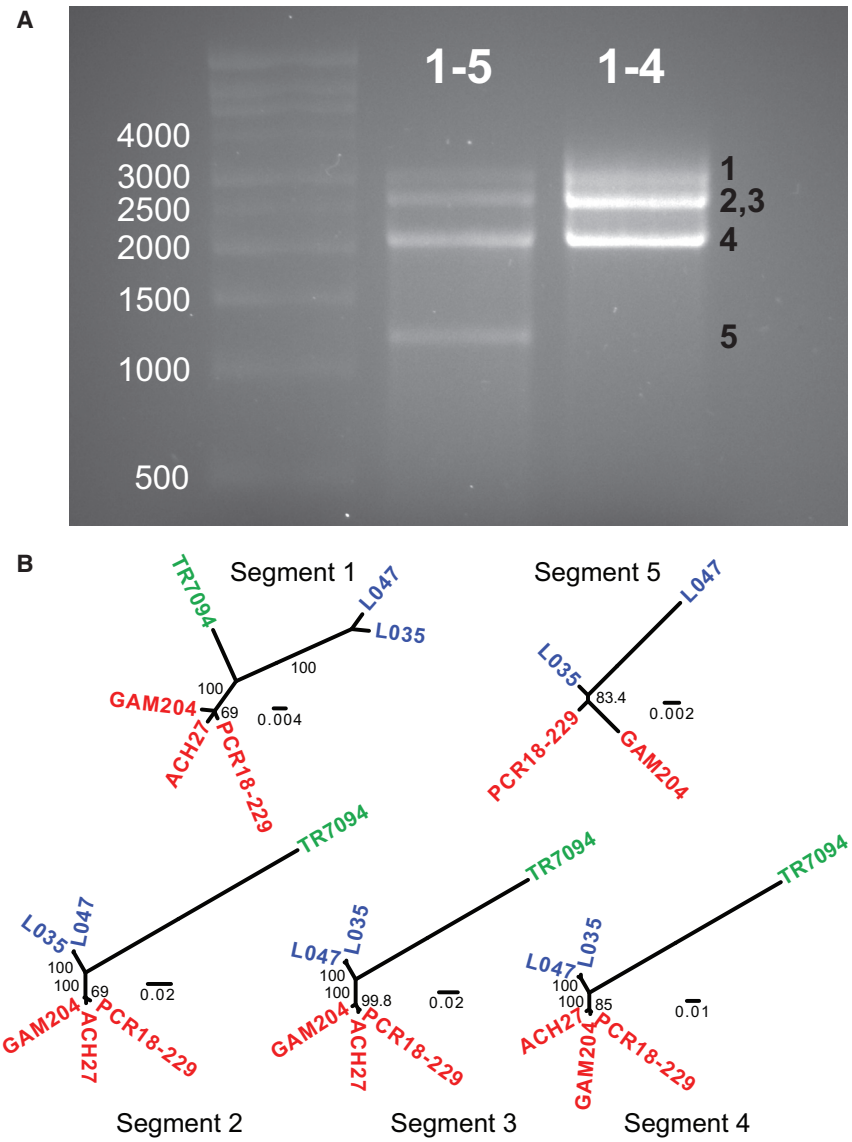


Figure 2. Segment 5 of GCXV Is Not Required for Replication in C6/36 Cells and Is Variably Present in Natural Isolates

(A) Cell culture supernatants resulting from transfection with GCXV segments 1–5 (passage 1) and segments 1–4 (passage 2). Segment numbers are shown in black; ladder band sizes (nt) are shown in white.

(B) Unrooted, nt-level phylogenetic trees including all six isolates of GCXV. ACH27 and TR7094 lack segment 5. Color indicates country of collection: blue for Peru, red for Panama, and green for Trinidad. Branch labels represent bootstrap support values. The scale bar indicates the number of nt changes per site.

See also Table S4.

and structure of the stem-loop varies across segments (Figure S2).

Coding Strategy

Three of the genome segments were monocistronic, while the other two each contained three open reading frames (ORFs) ≥ 400 nt (Figure 1C). The sizes and positions of the predicted ORFs were highly conserved, with the exception of viral protein (VP) 6, for which the position of the first AUG differed substantially across isolates (Table S2). The two largest segments each encoded a single ORF, and both exhibited significant protein-level similarity to non-structural (NS) domains that are conserved within the genus *Flavivirus*. The putative NS protein (NSP) 1 on segment 1 exhibited significant similarity (e-value = $5.02e^{-28}$) to Pfam's Flavi_NS5 family (PF00972), which corresponds to the RdRp, and Pfam's FtsJ-like methyltransferase family ($1.92e^{-4}$; PF01728). The methyltransferase domain in flaviviruses is required for the formation of the 5′ cap (Dong et al., 2008). The putative NSP2 protein on segment 2 exhibited significant similarity to three Pfam families and domains, all of which correspond to the *Flavivirus* NS3 protein: Peptidase_S7 ($5.17e^{-9}$; PF00949), Flavi_DEAD ($2.83e^{-18}$; PF07652) and Helicase_C ($8.01e^{-6}$; PF00271). These results suggest that NSP2 participates in at least two of the roles typically played by the *Flavivirus* NS3: serine protease activity typically used to cleave polypeptides into their mature forms and helicase activity likely involved in viral replication.

The three smallest segments exhibited no significant sequence similarity to known proteins. Five of the six predicted ORFs from segments 3 and 4 were detected in the proteogenomic analysis of viral particles (Figure S3A; Table S3). This suggests that these segments likely encode structural proteins; however, peptides from the putative NSP2 were also detected, so the purified sample may have had some NSP "contamination." Although VP2 was not detected with the proteogenomic data, an analysis of synonymous nt conservation (Firth, 2014), in the reading frame of VP1, supports the existence of a functional VP2 ORF (Figure S3B). Segments 3 and 4 both exhibited evidence for −1 ribosomal frameshifting. In both cases this included (1) overlapping ORFs in the expected orientation, (2) a known slippery heptanucleotide sequence at the end of the first ORF (Figure 1C), (3) predicted secondary structure just after the slippery sequence (Figures S3C and S3D), and (4) the detection of peptide sequences within the second participating ORF but prior to the first conserved AUG (Figure S3A).

Morphology and Host Specificity

Purified GCXV particles (20%–70% sucrose gradient) were 30–35 nm in diameter, spherical, and enveloped (Figure S4A). Treatment with NP40 ablated infectivity, confirming the enveloped nature of the virus (Figure 1D). Multiple attempts to

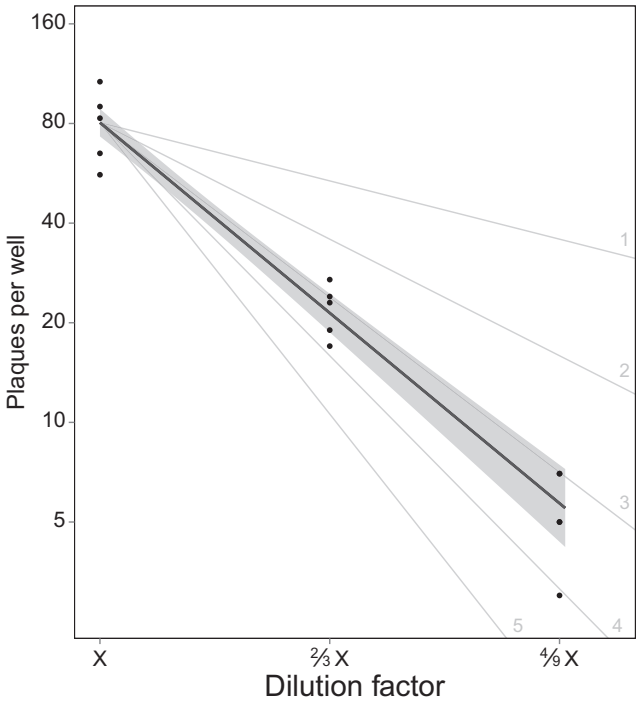


Figure 3. The Dose-Response Curve for GCXV Supports a Multicomponent Genome Organization

Plaque count is shown on the y axis, and the relative dilution of the virus stock is shown on the x axis. The analysis included five replicates per dilution (circles). The black line is the best-fitting line from a Poisson generalized linear model, with 95% confidence intervals (gray band). The gray lines represent expected slopes for multiplicities of infection of 1–5.

definitively identify viral particles within infected C6/36 cells using electron microscopy were unsuccessful; however, infected cells contained intracytosolic vacuoles loaded with vesicles 40–50 nm in diameter and dense particles ~20 nm in diameter (Figure S4B). Vesicles of the same type were also observed at the cell surface. This is similar to the pathology previously observed for flavivirus infected C6/36 cells.

Viral replication was detected in three mosquito cell lines and in intrathoracically inoculated adult female mosquitoes (Figures S5A and S5B). However, no replication was detected in tick-, sandfly-, or vertebrate-derived cells (Figures S5A and S5C), nor did the virus cause any observable illness in intracerebrally inoculated newborn mice, which is often a permissive environment for arbovirus replication. Virus was not detected in the larval progeny of infected mosquitoes, suggesting either the absence or a low occurrence of vertical transmission. No substantial mortality was observed in mosquitoes that survived inoculations during the 14-day viral growth curve, suggesting that GCXV is non-lethal in its mosquito hosts.

Identification of a Jingmenvirus in a Primate Host

We sequenced one variant of JMTV (RC27) directly from plasma collected from a red colobus monkey in Uganda. A draft genome (Ladner et al., 2014) was obtained with 70.6% to 98.1% coverage for each of the four genome segments that have been described for JMTV (Shi et al., 2015). RC27 exhibited high similarity to isolates from ticks: 88%–92.6% nt identity across all four segments compared with SY84 from China (Qin et al., 2014) and Mogiana tick virus (MGTV) from Brazil (Maruyama et al., 2014). In species-level phylogenetic analyses, the Chinese isolates (Qin et al., 2014) formed a distinct clade compared with the isolates from Uganda and Brazil (Figure S6). Of note, RC27 contained a 450 nt deletion relative to SY84 and MGTV within the ORF of segment 2 (numbered according to Shi et al., 2015). This deletion was observed in reads from metagenomic sequencing and confirmed by RT-PCR and Sanger sequencing. Attempts to isolate this virus on a variety of cell lines were unsuccessful.

Phylogenetic Analysis

Despite high levels of divergence, both the NS3 and NS5 phylogenies support the segmented Jingmenviruses as a monophyletic group (83%–99% bootstrap support), which is sister to the Tamana bat virus and the *Flavivirus* genus (Figure 5). The Jingmenviruses collected from insect hosts form a well-supported (69%–100% bootstrap support) sub-clade within this group.

DISCUSSION

Here we describe a segmented virus that is distantly related to the flaviviruses, which we have tentatively designated GCXV. In vitro and in vivo replication experiments suggest GCXV is mosquito specific, but we cannot definitively rule out replication in other organisms. In contrast to the prototypical flavivirus (i.e., unsegmented, single polyprotein), the genome of GCXV is composed of four or five distinct segments, depending on the isolate. The four largest segments were found in all isolates and are therefore assumed to be necessary for infection and transmission, whereas the fifth segment seems to be optional. VP7 peptides were not detected in purified GCXV virions, suggesting that segment 5 likely encodes an NSP. There was no correlation between the presence and absence of this fifth segment and phylogenetic relationships, as inferred from the four "core" segments (Figure 2B), and in vitro experiments confirmed that GCXV can replicate without segment 5. Similarly, natural isolates of Beet Necrotic Yellow Vein Virus (BNYVV; ssRNA$^+$) have been described with both four and five genome segments; the fifth segment is not required for successful infection or transmission but, when present, is thought to play a role in modulating pathogenesis (Tamada et al., 1989). Also similar to BNYVV, GCXV appears to individually package its genome segments (i.e., is multicomponent). Although a multicomponent organization is not required for segment loss, it provides a straightforward mechanism for the establishment of infections with a subset of genome segments.

Multicomponent genome organizations are relatively common among viruses that infect plants and fungi, but no multicomponent animal viruses have been described (Fulton, 1980; King et al., 2011). The dose-response kinetics of GCXV indicate that at least three different particles are required for plaque formation. Given that the vast majority of animal viruses require only a single particle to form a plaque (Flint et al., 2009), this result provides strong evidence for a multicomponent virus. Our results from RNA fluorescence in situ hybridization (FISH) and electron microscopy are also consistent with GCXV's being

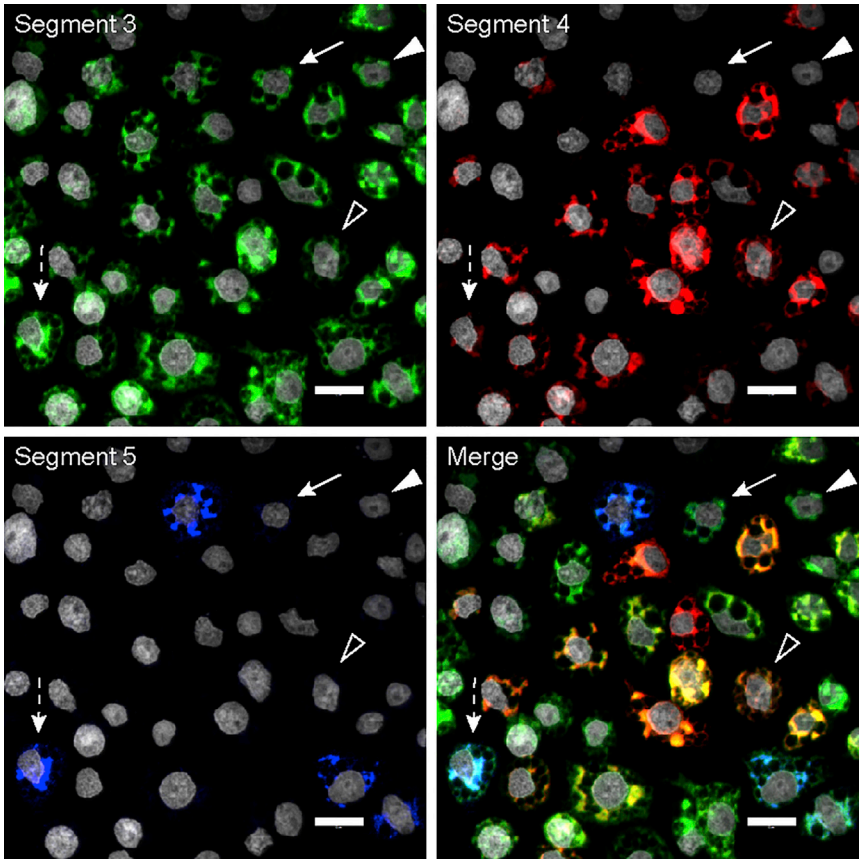


Figure 4. Multiple Combinations of GCXV Genome Segments Were Detected in Infected Cells

A representative 63× magnification image is shown from the co-hybridization of probes for segments 3–5. Four different infected cell types were observed: closed arrowhead, missing segments 4 and 5; open arrowhead, missing segment 5; full arrow, missing segment 4 (note that segment 5 present in low abundance in the indicated cell); and dashed arrow, containing all assayed segments. The scale bar represents 10 μm. See also Figure S1 and Table S1.

multicomponent. Although segments 1–3 were detected in all infected cells, segments 4 and 5 were variably present, and the measured particle size for GCXV, ~30–35 nm, is considerably smaller than that of other flaviviruses. In fact, the spherical volume calculated from this size lies below the 95% confidence limit calculated by Cui et al. (2014) for viruses with 12 kb genomes. Similar to the flaviviruses, however, GCXV particles are enveloped, whereas particles from the multicomponent viruses known to infect plants and fungi do not contain envelopes.

Multiple theories have been proposed regarding potential benefits of segmented genomes; however, it is unclear whether independent packaging could provide additional benefits (over packaging within a single virion; Fulton, 1980; Pressing and Reanney, 1984; Reijnders, 1978) or whether this arrangement may simply be a byproduct of the mechanism of segmentation (most likely the formation of complementary defective viral particles; García-Arriaza et al., 2004). It has been proposed that multicomponent viruses may represent specialized forms of unsegmented genomes, which are facilitated by the existence of efficient mechanisms of transmission. In fact, transmission inefficiency is thought to be the primary reason for the general lack of multicomponent animal viruses (Goldbach, 1986). Vertical transmission has been shown to be the primary mechanism for the insect-specific Culex flavivirus (Bolling et al., 2012); however, vertical transmission was not detected in our experiments with GCXV. Further investigation into the transmission mechanism for GCXV will provide insight into the transmission requirements for multicomponent viruses and the unequal distribution of these viruses among eukaryotic lineages.

In addition to GCXV, several other segmented, flavi-like viruses have been reported. These viruses have tentatively been coined Jingmenviruses (Shi et al., 2015). JMTV and MGTV are two highly similar (88%–90% identical, nt-level), segmented viruses that have been isolated from cattle ticks (*Rhipicephalus microplus*) in China (Qin et al., 2014) and Brazil (Maruyama et al., 2014), respectively. Here we report an additional virus with high similarity to JMTV, which we detected in the blood of a non-human primate in Uganda. Additionally, eight JMTV-like viruses have been sequenced from insects collected in China (Shi et al., 2015), England, and Kenya (Webster et al., 2015) (Table 2). Qin et al. (2014) have also demonstrated that four highly expressed transcripts isolated from a larval roundworm (*Toxocara canis*) (Callister et al., 2008) are actually of viral origin and exhibit homology to JMTV. This virus has tentatively been named *T. canis* larva agent (TCLA) (Qin et al., 2014). Our analysis supports this group of segmented viruses as a monophyletic clade that includes GCXV. Furthermore, we found that GCXV belongs to a well-supported sub-clade within the Jingmenvirus group, which is composed exclusively of the viruses that have been isolated from insects. Along with the extensive divergence of the lineages containing the tick/mammal and roundworm viruses, this finding suggests the presence of sub-groups within the Jingmenviruses that are specialized for different types of hosts.

All of the sequenced Jingmenviruses appear to have at least four genome segments, and homology can be inferred for at least three of these segments across all viruses. Although MGTV was not initially recognized as being segmented, we used the published Illumina data set (SRR525284; Maruyama et al., 2014) to assemble nearly full-length contigs with sequence homology to all four segments of JMTV (see Supplemental Experimental Procedures). Similarly, although Webster et al. (2015) only reported one segment (encoding the putative RdRp) for the three Jingmenviruses they sequenced, we identified several additional contigs in their data set with significant sequence homology to segments 1–4 from GCXV (see

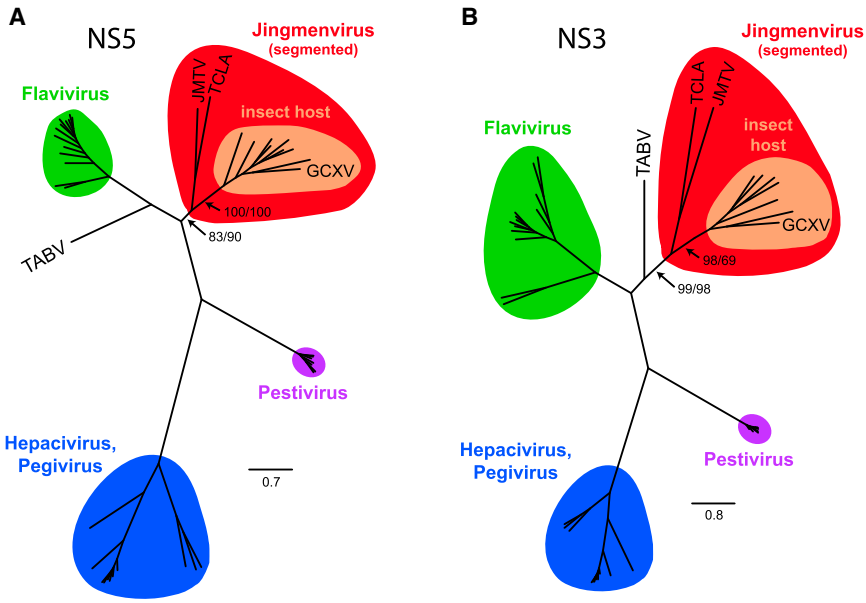


Figure 5. GCXV Belongs to an Insect-Specific Clade within the Segmented Jingmenviruses

Maximum likelihood (ML) trees are shown with bootstrap support values from both the ML and neighbor-joining (NJ) trees: ML/NJ. The scale indicates the number of amino acid changes per site. See also Table S5.

Supplemental Experimental Procedures). All of these viruses contain separate monocistronic segments with sequence homology to the *Flavivirus* NS3 and NS5 proteins. Additionally, they all have one multicistronic segment with two partially overlapping ORFs, consistent with −1 ribosomal frameshifting; the first ORF is predicted to have a signal peptide, and the second contains many predicted transmembrane helices. Protein sequence homology is detectable for this segment between TCLA (ANT-3) and JMTV/MGTV (segment 4) and also among the nine insect-associated viruses, including GCXV (segment 4 in Shi et al., 2015, segment 3 for GCXV). However, no sequence-level similarity is detectable for this segment between these two groups. GCXV, JMTV, and TCLA all exhibit some degree of sequence conservation (across segments) in the non-coding UTRs (Callister et al., 2008; Qin et al., 2014), and these regions may play a role in the initiation of translation and/or replication.

On the basis of the apparent homology of at least three segments, it is likely that the Jingmenviruses shared a segmented common ancestor. The existence of a common, segmented ancestor suggests that the multicomponent organization seen in GCXV is likely also shared by the other Jingmenviruses. However, the method of packaging has not been investigated in any of these other viruses.

Despite a common origin, levels of sequence divergence among the different Jingmenviruses were high. In fact, average levels of amino acid divergence among the Jingmenviruses (NS5: 59.8%; NS3: 70.2%) were higher than the maximum divergence seen among viruses within the *Flavivirus* genus (NS5: 57%; NS3: 69%). Structural differences are also evident among the Jingmenviruses. These differences include genome segments with distinct coding strategies and no measurable homology (e.g., GCXV multicistronic segment 4 versus JMTV monocistronic segment 2) and the presence of polyadenylated 3′ termini only in JMTV/MGTV and TCLA (Table 2). On the basis of the extensive sequence-level and structural genomic diversity, this clade of segmented flavi-like viruses appears to be quite old, and it likely includes substantial viral diversity that has yet to be described.

Because of the high levels of divergence and the lack of an appropriate outgroup, it is not possible to use sequence information to accurately determine the polarity of the transition between segmented and unsegmented genomes. One possibility is that the segmented genome is ancestral and that these pieces were later joined together to form the prototypical, unsegmented flavivirus genome. Such a transition could be mediated by repeated episodes of non-homologous copy-choice recombination (Simon-Loriere and Holmes, 2011); however, to our knowledge no such transitions have been described. The alternative is that these segmented viruses evolved from an unsegmented ancestor. Phylogenetic analysis of the *Tetraviridae* supports a similar transition between the betatetraviruses (monopartite) and the omegatetraviruses (bipartite), which is hypothesized to have been mediated through the formation of subgenomic RNA molecules (Zeddam et al., 2010).

On the basis of the lack of sequence similarity to flaviviruses at two of four segments (those thought to encode the structural proteins), Qin et al. (2014) argued that JMTV is likely of hybrid origin, resulting from the coinfection of a flavivirus and a second, as of yet, uncharacterized virus. Although we cannot rule out this scenario, segmentation of a single flavivirus-like ancestor presents a more parsimonious explanation. The patterns of sequence divergence and similarity seen between members of the genus *Flavivirus* and both JMTV and GCXV are as expected between very divergent flaviviruses, with weak but significant similarity at the highly conserved NS genes, but a lack of detectable homology at the structural genes, which typically exhibit higher rates of evolution (Chambers et al., 1990). In fact, this is exactly the same pattern seen between GCXV and JMTV. Therefore, in the absence of an identified "donor" group for the putative structural genes of JMTV and GCXV, the most parsimonious scenario involves the segmentation of a single, flavivirus-like ancestor.

In summary, we have described a multicomponent virus, which belongs to a diverse clade of segmented viruses related to the family *Flaviviridae*. We have also described a variant of JMTV, the prototype species for this clade, from a non-human primate, thus substantially expanding the host range of the group and indicating potential implications for animal and human health. This clade establishes a strong link between two distinct genome organizations, which may help uncover some of the mysteries associated with the evolution of genome architecture in RNA viruses; however, methods of inference beyond

Table 2. Comparison of Viral Properties, Including All Jingmenviruses and the Three Defined Genera of *Flaviviridae*

Virus/Genus	Documented Hosts	Number of Segments	Genome Size (kb)	Particle Size (nm)	3′ Poly(A)	Reference
GCXV	mosquitoes (*Culex* spp.)	4–5	10.8–12	30–35	no	this study
SAIV7, WHFV, WHCV, WHAV1,2	various insects	4	10.4–11	?	no	Shi et al., 2015
JMTV, MGTV	various ticks, cattle, red colobus monkey	4	11.4[a]	70–80[a]	yes	Maruyama et al., 2014; Qin et al., 2014; this study
TCLA	dog roundworm *(Toxocara canis)*	4	9.7[b]	?	yes	Callister et al., 2008; Qin et al., 2014
Charvil virus	fruit fly (*Drosophila melanogaster*)	?	?	?	?	Webster et al., 2015
Flavivirus-like1,2	unspecified *Drosophilidae* spp.	?	?	?	?	
Flavivirus	various arthropods and vertebrates	1	11	50	no	King et al., 2011
Hepacivirus	*humans*	1	9.6	50	no	
Pestivirus	*pigs and ruminants*	1	12.3	40–60	no	

Question marks indicate a lack of information.
[a]Genome length and particle size are for JMTV.
[b]Estimated from the combined length of the four EST sequences in Callister et al. (2008); it may be an underestimate.

sequence-level homology will be needed to reconstruct the evolutionary history that connects these genome types. The existence of an enveloped, multicomponent animal virus will require us to rethink historical perspectives regarding the advantages and requirements of this type of genome organization.

EXPERIMENTAL PROCEDURES

GCXV Isolation and Sequencing

Mosquitoes were collected in Panama, Peru, and Trinidad during 2008–2013 as a part of several surveillance studies looking for arbo- and insect-specific viruses (Table 1) (Auguste et al., 2010; Eastwood et al., 2016; Vasilakis et al., 2014). In total, six isolates of GCXV were detected through random amplification and high-throughput sequencing. In each case, the virus was obtained from a collection of mosquitoes pooled on the basis of species and collection locality. Each pool was homogenized and clarified supernatant was inoculated onto monolayers of an *Aedes albopictus* mosquito-derived cell line, C6/36. The homogenate from five of these pools resulted in CPE. In these cases, the cell culture supernatant was preserved in TRIzol LS, and RNA was extracted using Direct-zol RNA MiniPrep (Zymo). For the pool that did not exhibit CPE (ACH27), RNA was extracted directly from the mosquito homogenate using the same protocol. The RNA was then amplified using sequence-independent single primer amplification, as described previously (Djikeng et al., 2008). Amplicons were sheared to ~400 bp and used as starting material for Illumina TRUseq DNA libraries. Sequencing was performed on a MiSeq using either 300 or 500 cycle kits (2 × 150, 2 × 250). Cutadapt (Martin, 2011) and Prinseq-lite (Schmieder and Edwards, 2011) were used to trim primers and remove poor quality reads, respectively, and then genomes were assembled using Ray Meta (Boisvert et al., 2012) in combination with custom scripts. For three of the isolated viruses, the terminal ends were determined using the SMARTer RACE cDNA Amplification kit (Clontech). In order to ensure detection of the most 3′ terminal base and to test for the presence of a polyadenylated 3′ terminus, 3′ RACE was conducted using both poly-A and poly-U polymerases. For one isolate (LO35), the genome sequence was verified with Sanger sequencing.

Sequence Analysis

TMHMM version 2.0 (Krogh et al., 2001) was used to predict transmembrane helical segments, and SignalP version 4.1 (Petersen et al., 2011) was used to identify signal peptide sequences. MFOLD/Quickfold (Zuker, 2003) and RNAstructure (Reuter and Mathews, 2010) were used to predict RNA secondary structures.

MEGA version 5 (Tamura et al., 2011) was used to conduct species-level nt alignments using MUSCLE, to calculate pairwise divergences (p-distance with pairwise deletion), and to construct maximum likelihood phylogenetic trees with 1,000 bootstrap replicates (Tamura-Nei model, uniform rates, complete deletion). The Pfam database (Punta et al., 2012) was searched using CD-search (Marchler-Bauer and Bryant, 2004) to identify conserved protein domains. Family-wide amino acid alignments (see Table S5 for list of sequences included) were conducted using the E-INS-I algorithm in MAFFT version 7 (Katoh and Standley, 2013), and trimAl was used to trim ambiguously aligned positions (Capella-Gutiérrez et al., 2009). Trees were built using both neighbor joining (MEGA, JTT substitution model, pairwise deletion, uniform rates) and maximum likelihood (PhyML version 3.1 [Guindon et al., 2010], LG substitution model Γ-distributed rate variation) methods with 100 bootstrap replicates.

Nuclease Digestion

To test the nature of the genomic material, nucleic acids extracted from GCXV-LO35 were separately incubated in the presence of DNase I (Ambion), RNase I (Thermo Fisher), and nuclease-free water at 37°C for 20 min. After incubation, the sample was cleaned (Direct-zol), and cDNA was synthesized using SuperScript III (Invitrogen) with random hexamers. Segment-specific amplicons were generated for each treatment to test for the presence of intact genomic segments (see Table S6 for primers).

Rescue by Reverse Genetics

All five genome segments of GCXV-LO35 were amplified by RT-PCR with forward primers that included the T7 RNA polymerase promoter at the 5′ end. Three combinations (segments 1–5, 1–4, and 2–5) of the full-length amplicons were mixed in equimolar amounts, and the mixtures were used to produce RNA by in vitro transcription using the MessageMAX T7 kit (CellScript), which produces a mixture of capped and uncapped RNAs. The in vitro transcribed RNAs were transfected into C6/36 cells (Mirus TransIT mRNA), and the cells were monitored for changes in morphology. After changes were detected, a small amount of supernatant from the transfected cells was passaged onto fresh C6/36 cells for one or two passages before viral RNA was isolated from the supernatants of the infected cells.

Plaque Assay

C6/36 cells were seeded in six-well plates so that the cells would be 90%–100% confluent the following day. A 1.5-fold serial dilution series was generated with GCXV-LO35 stocks using complete medium. Diluted virus (500 μl) was added to each of five wells of the plate; the sixth well served as a no-virus control. Virus was allowed to bind to the C6/36 cells for 1 hr at room temperature with gentle rocking every 15 min. The virus inoculum was then removed from the wells, the wells were washed with PBS to remove any unbound viruses and 2 ml of overlay (1% tragacanth in complete medium) was added to each well. After a 4-day incubation at 28°C, the overlay was gently removed,

and cells were fixed with 4% paraformaldehyde. The fixed cells were stained with crystal violet to visualize plaques. A Poisson generalized linear model was used to estimate the multiplicity of infection.

RNA FISH

Probe sets for RNA FISH were designed using the Stellaris Probe Designer and synthesized by Biosearch Technologies. One probe set (40–48 distinct probes) was designed for each segment using the GCXV-LO35 genome sequence; each set was labeled with one of three fluorophores: Quasar 570, Quasar 670, or FAM (Table S7). Two sets of probes (segments 1–3 and 3–5) were independently hybridized to C6/36 cell cultures 24 hr after inoculation with GCXV-LO35 (see Supplemental Experimental Procedures for details). DAPI (300 nM in PBS) was used for nuclei detection. Ten images of infected cells (20× magnification) and one of mock-infected cells were obtained for each probe set. Individual cells were identified and fluorescent intensities were quantified using the Columbus software package (PerkinElmer). To be conservative regarding segment presence or absence, cells were counted only if all assayed segments had intensities that were either (1) greater than the maximum intensity observed in the mock-infected cells (considered present) or (2) lower than the average intensity observed in the mock-infected cells (considered absent).

Virus Purification

To isolate virions, cell culture supernatant was purified on a 20%–70% continuous sucrose gradient (see Supplemental Experimental Procedures for details). To test for the presence of an envelope, virion infectivity was assayed using the sucrose-purified particles following treatment with the anionic detergent, NP40, at concentrations of 0.05%, 0.1%, and 0.25% for 30 min. Both treated and untreated samples were assayed for infectivity on C6/36 cells using CPE assays as well as qRT-PCR to assess growth kinetics. Transmission electron microscopy was used to visualize individual viral particles as well as infected C6/36 cells (see Supplemental Experimental Procedures for details).

Proteogenomics

Mass spectrometry-based proteomics was used to aid in the annotation of the GCXV genome, including the identification of structural proteins (Jaffe et al., 2004). Briefly, purified virions were fractionated through PAGE and divided into 20 samples. Each individual fraction was trypsin digested, and the peptides were analyzed using a single-segment data-dependent 30,000 resolution MS1 scan followed by ms/ms rapid scans of the 15 highest intensity ions. Peptides were searched against a database that contained all potential protein ORFs in the GCXV-LO35 genome plus 8,873 *Aedes albopictus* sequences from the NCBI protein non-redundant database. For detailed methodology, see Supplemental Experimental Procedures.

In Vitro Culture of GCXV

A variety of vertebrate and invertebrate cell lines were used to determine infectivity and viral growth of GCXV-LO35: three derived from mammals (Vero, *Chlorocebus sabaeus*; HeLa, *Homo sapiens*; BHK-21 *Mesocricetus auratus*), one avian (DF-1, *Gallus gallus*), three from mosquitoes (C6/36, *Aedes albopictus*; *Culex tarsalis*; Aag2, *Aedes aegypti*), one from a sandfly (LL-5, *Lutzomyia longipalpis*), and one from a tick (ISE6, *Ixodes scapularis*). Using supernatant from GCXV-infected C6/36 cells as the inoculum, one-step growth curves were conducted with one time point per day; qRT-PCR was used to quantify viral concentration.

Mosquito Infections

Viral replication was also assayed in adult mosquitoes (*Aedes albopictus* and *Culex quinquefasciatus*). Briefly, mosquitoes were intrathoracically inoculated with 0.5 μl of GCXV-LO35 at 10^8 genome copies/ml, and viral copy number was estimated using qRT-PCR at four time points (1, 4, 7, and 14 days post inoculation [dpi]). For vertical transmission studies, intrathoracically inoculated *Culex quinquefasciatus* mosquitoes were offered blood meals at 7 dpi to induce the production of eggs, which were reared to fourth instar larvae and then assayed for the presence of GCXV-LO35. See Supplemental Experimental Procedures for details.

Inoculation of mice

Culture media from C6/36 cells infected with GCXV-LO47, LO35, and TR7094 were used to inoculate a litter (n = 10) of 2-day-old CD1 mice intracerebrally. Mice were observed for signs of illness for 14 days. Animal experiments were carried out under a protocol approved by the University of Texas Medical Branch (UTMB) Institutional Animal Care and Use Committee.

JMTV Sequencing

With the approval of the Uganda Wildlife Authority, the Uganda National Council for Science and Technology, and the University of Wisconsin Animal Care and Use Committee, plasma was collected from a red colobus monkey (*Procolobus rufomitratus*) on February 2, 2012, in Kibale National Park, Uganda. This animal showed no signs of disease at the time of sampling. This sample was processed as described previously (Sibley et al., 2014). Briefly, viral RNA was isolated from 1 ml of plasma (Qiagen QIAamp MinElute virus kit), treated with DNase I (DNA-free; Ambion), and converted to double-stranded cDNA using random hexamers (Superscript double-stranded cDNA Synthesis kit; Invitrogen). Purified cDNA was sequenced on an Illumina MiSeq using the Nextera XT DNA sample preparation kit (Illumina). Sequence data were analyzed using CLC Genomics Workbench version 8.5 (CLC bio) (see Supplemental Experimental Procedures for details). JMTV contigs and unassembled reads were identified on the basis of nt-level (blastn) similarity to published JMTV sequences (GenBank). Gaps in the de novo assembly were filled using PCR amplification and Sanger sequencing with primers designed using the de novo contigs/reads as well as the complete genome from the original JMTV isolate (SY84, GenBank: KJ001579–KJ001582).

ACCESSION NUMBERS

The accession numbers for the viruses sequenced and reported in this paper are GenBank: KM521571–KM521574, KM461666–KM461670, KM521561–KM521565, KM521556–KM521560, KM521566–KM521570, KM521552–KM521555, and KX377513–KX377516.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and seven tables and can be found with this article online at http://dx.doi.org/10.1016/j.chom.2016.07.011.

AUTHOR CONTRIBUTIONS

Conceptualization, J.T.L., M.R.W., B.B., A.J.A., and G.P.; Investigation, M.R.W., B.B., A.J.A., A.P.D., M.E.L., S.D.S., D.K., K.P., H.G., M.T.A., D.R., E.E.B., L.S.J., M.C.G., L.H.C., T.K., and B.W.; Formal Analysis, J.T.L., M.R.W., K.P.K., D.F., V.L.P., and M.D.W.; Writing – Original Draft, J.T.L.; Writing – Review & Editing, J.T.L., M.R.W., B.B., A.J.A., S.D.S., T.L.G., T.C.F., D.H.O., M.L., D.H., S.C.W., L.D.K., R.B.T., and G.P.; Resources, G.E., F.C.-L., T.J.K., J.R.L., and R.B.T.; Supervision, T.L.G., S.C.W., L.D.K., and G.P.

ACKNOWLEDGMENTS

Work at U.S. Army Medical Research Institute of Infectious Diseases was funded by the Defense Threat Reduction Agency, project 1881290. Work at UTMB was supported by NIH contract HHSN272201000040I/HHSN27200004/D04 to R.B.T. A.J.A. was supported by the James W. McLaughlin Endowment fund. Mosquito collection in Panama was funded by Smithsonian Tropical Research Institute – Environmental Protection Agency grant DW33-92296801-0 to Montira J. Pongsiri. Work in Panama was supported by a Robert E. Shope fellowship awarded to G.E.

We thank Biosearch Technologies for providing Stellaris probes; Anne Payne and Greta VanSlyke for tick cell culture and in vivo mosquito help, respectively; Thomas Wood, Steven Widen, and Jill Thompson for sequencing support; Frederick Murphy for assistance interpreting electron micrographs; Roberto Fernandez and Anibal Huayanay for help with mosquito identification and field work, respectively; Jose R. Rovira, Larissa C. Dutari, Mauricio

Quintero, Anel J. Duncan, Denis Lezcano, Gaspar Ho, and Apolonio Valdez for field assistance in Panama; Carolina Guevara and Patricia Aguilar for help preparing mosquito pools and shipping genetic materials; Edward Holmes and Scott Adkins for insightful discussions; and Andrew Firth for help running the synonymous site conservation analysis. We are grateful to Panama's Environmental Authority for supporting collections in Achiote and Gamboa and to LT Kirk Mundal and the Peruvian Ministry of Health for facilitation of collections in Loreto.

Received: March 4, 2016
Revised: June 9, 2016
Accepted: July 26, 2016
Published: August 25, 2016

REFERENCES

Auguste, A.J., Adams, A.P., Arrigo, N.C., Martinez, R., Travassos da Rosa, A.P., Adesiyun, A.A., Chadee, D.D., Tesh, R.B., Carrington, C.V., and Weaver, S.C. (2010). Isolation and characterization of sylvatic mosquito-borne viruses in Trinidad: enzootic transmission and a new potential vector of Mucambo virus. Am. J. Trop. Med. Hyg. *83*, 1262–1265.

Auguste, A.J., Carrington, C.V., Forrester, N.L., Popov, V.L., Guzman, H., Widen, S.G., Wood, T.G., Weaver, S.C., and Tesh, R.B. (2014). Characterization of a novel Negevirus and a novel Bunyavirus isolated from Culex (Culex) declarator mosquitoes in Trinidad. J. Gen. Virol. *95*, 481–485.

Boisvert, S., Raymond, F., Godzaridis, E., Laviolette, F., and Corbeil, J. (2012). Ray Meta: scalable de novo metagenome assembly and profiling. Genome Biol. *13*, R122.

Bolling, B.G., Olea-Popelka, F.J., Eisen, L., Moore, C.G., and Blair, C.D. (2012). Transmission dynamics of an insect-specific flavivirus in a naturally infected Culex pipiens laboratory colony and effects of co-infection on vector competence for West Nile virus. Virology *427*, 90–97.

Callister, D.M., Winter, A.D., Page, A.P., and Maizels, R.M. (2008). Four abundant novel transcript genes from Toxocara canis with unrelated coding sequences share untranslated region tracts implicated in the control of gene expression. Mol. Biochem. Parasitol. *162*, 60–70.

Capella-Gutiérrez, S., Silla-Martínez, J.M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics *25*, 1972–1973.

Chambers, T.J., Hahn, C.S., Galler, R., and Rice, C.M. (1990). Flavivirus genome organization, expression, and replication. Annu. Rev. Microbiol. *44*, 649–688.

Cui, J., Schlub, T.E., and Holmes, E.C. (2014). An allometric relationship between the genome length and virion volume of viruses. J. Virol. *88*, 6403–6410.

Djikeng, A., Halpin, R., Kuzmickas, R., Depasse, J., Feldblyum, J., Sengamalay, N., Afonso, C., Zhang, X., Anderson, N.G., Ghedin, E., and Spiro, D.J. (2008). Viral genome sequencing by random priming methods. BMC Genomics *9*, 5.

Dong, H., Zhang, B., and Shi, P.-Y. (2008). Flavivirus methyltransferase: a novel antiviral target. Antiviral Res. *80*, 1–10.

Eastwood, G., Loaiza, J.R., Pongsiri, M.J., Sanjur, O.I., Pecor, J.E., Auguste, A.J., and Kramer, L.D. (2016). Enzootic arbovirus surveillance in forest habitat and phylogenetic characterization of novel isolates of Gamboa virus in Panama. Am. J. Trop. Med. Hyg. *94*, 786–793.

Firth, A.E. (2014). Mapping overlapping functional elements embedded within the protein-coding regions of RNA viruses. Nucleic Acids Res. *42*, 12425–12439.

Flint, S.J., Enquist, L.W., Racaniello, V.R., Skalka, A.M., Barnum, D.R., and de Evaluación, E. (2009). Principles of Virology Volume I: Molecular Biology (ASM Press).

Fulton, R.W. (1980). Biological significance of multicomponent viruses. Annu. Rev. Phytopathol. *18*, 131–146.

García-Arriaza, J., Manrubia, S.C., Toja, M., Domingo, E., and Escarmís, C. (2004). Evolutionary transition toward defective RNAs that are infectious by complementation. J. Virol. *78*, 11678–11685.

Goldbach, R. (1986). Molecular evolution of plant RNA viruses. Annu. Rev. Phytopathol. *24*, 289–310.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. *59*, 307–321.

Holmes, E.C. (2009). The Evolution and Emergence of RNA Viruses (Oxford University Press).

Jaffe, J.D., Berg, H.C., and Church, G.M. (2004). Proteogenomic mapping as a complementary method to perform genome annotation. Proteomics *4*, 59–77.

Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. *30*, 772–780.

King, A.M., Adams, M.J., and Lefkowitz, E.J. (2011). Virus Taxonomy: Classification and Nomenclature of Viruses: Ninth Report of the International Committee on Taxonomy of Viruses, *Volume 9* (Elsevier).

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E.L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J. Mol. Biol. *305*, 567–580.

Ladner, J.T., Beitzel, B., Chain, P.S., Davenport, M.G., Donaldson, E.F., Frieman, M., Kugelman, J.R., Kuhn, J.H., O'Rear, J., Sabeti, P.C., et al.; Threat Characterization Consortium (2014). Standards for sequencing viral genomes in the era of high-throughput sequencing. MBio *5*, e01360–e14.

Mahy, B.W. (2009). The Dictionary of Virology (Academic Press).

Marchler-Bauer, A., and Bryant, S.H. (2004). CD-Search: protein domain annotations on the fly. Nucleic Acids Res. *32*, W327–W331.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnetjournal *17*, 10–12.

Maruyama, S.R., Castro-Jorge, L.A., Ribeiro, J.M.C., Gardinassi, L.G., Garcia, G.R., Brandão, L.G., Rodrigues, A.R., Okada, M.I., Abrão, E.P., Ferreira, B.R., et al. (2014). Characterisation of divergent flavivirus NS3 and NS5 protein sequences detected in Rhipicephalus microplus ticks from Brazil. Mem. Inst. Oswaldo Cruz *109*, 38–50.

Petersen, T.N., Brunak, S., von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods *8*, 785–786.

Pressing, J., and Reanney, D.C. (1984). Divided genomes and intrinsic noise. J. Mol. Evol. *20*, 135–146.

Punta, M., Coggill, P.C., Eberhardt, R.Y., Mistry, J., Tate, J., Boursnell, C., Pang, N., Forslund, K., Ceric, G., Clements, J., et al. (2012). The Pfam protein families database. Nucleic Acids Res. *40*, D290–D301.

Qin, X.-C., Shi, M., Tian, J.-H., Lin, X.-D., Gao, D.-Y., He, J.-R., Wang, J.-B., Li, C.-X., Kang, Y.-J., Yu, B., et al. (2014). A tick-borne segmented RNA virus contains genome segments derived from unsegmented viral ancestors. Proc. Natl. Acad. Sci. U S A *111*, 6744–6749.

Reijnders, L. (1978). The origin of multicomponent small ribonucleoprotein viruses. Adv. Virus Res. *23*, 79–102.

Reuter, J.S., and Mathews, D.H. (2010). RNAstructure: software for RNA secondary structure prediction and analysis. BMC Bioinformatics *11*, 129.

Sánchez-Navarro, J.A., Zwart, M.P., and Elena, S.F. (2013). Effects of the number of genome segments on primary and systemic infections with a multipartite plant RNA virus. J. Virol. *87*, 10805–10815.

Schmieder, R., and Edwards, R. (2011). Quality control and preprocessing of metagenomic datasets. Bioinformatics *27*, 863–864.

Shi, M., Lin, X.D., Vasilakis, N., Tian, J.H., Li, C.X., Chen, L.J., Eastwood, G., Diao, X.N., Chen, M.H., Chen, X., et al. (2015). Divergent viruses discovered in arthropods and vertebrates revise the evolutionary history of the Flaviviridae and related viruses. J. Virol. *90*, 659–669.

Sibley, S.D., Lauck, M., Bailey, A.L., Hyeroba, D., Tumukunde, A., Weny, G., Chapman, C.A., O'Connor, D.H., Goldberg, T.L., and Friedrich, T.C. (2014). Discovery and characterization of distinct simian pegiviruses in three wild African Old World monkey species. PLoS ONE *9*, e98569.

Simon-Loriere, E., and Holmes, E.C. (2011). Why do RNA viruses recombine? Nat. Rev. Microbiol. *9*, 617–626.

Tamada, T., Shirako, Y., Abe, H., Saito, M., Kiguchi, T., and Harada, T. (1989). Production and pathogenicity of isolates of beet necrotic yellow vein virus with different numbers of RNA components. J. Gen. Virol. *70*, 3399–3409.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. *28*, 2731–2739.

Vasilakis, N., Castro-Llanos, F., Widen, S.G., Aguilar, P.V., Guzman, H., Guevara, C., Fernandez, R., Auguste, A.J., Wood, T.G., Popov, V., et al. (2014). Arboretum and Puerto Almendras viruses: two novel rhabdoviruses isolated from mosquitoes in Peru. J. Gen. Virol. *95*, 787–792.

Webster, C.L., Waldron, F.M., Robertson, S., Crowson, D., Ferrari, G., Quintana, J.F., Brouqui, J.M., Bayne, E.H., Longdon, B., Buck, A.H., et al. (2015). The discovery, distribution, and evolution of viruses associated with Drosophila melanogaster. PLoS Biol. *13*, e1002210.

Zeddam, J.L., Gordon, K.H., Lauber, C., Alves, C.A., Luke, B.T., Hanzlik, T.N., Ward, V.K., and Gorbalenya, A.E. (2010). Euprosterna elaeasa virus genome sequence and evolution of the Tetraviridae family: emergence of bipartite genomes and conservation of the VPg signal with the dsRNA Birnaviridae family. Virology *397*, 145–154.

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. *31*, 3406–3415.