Virus Evolution, 2016, 2(2): vew033

doi: 10.1093/ve/vew033
Research article

Transmission bottlenecks and RNAi collectively influence tick-borne flavivirus evolution

Nathan D. Grubaugh,[1,†,‡] Claudia Rückert,[1,‡] Philip M. Armstrong,[2] Angela Bransfield,[2] John F. Anderson,[2] Gregory D. Ebel,[1] and Doug E. Brackney[2,*]

[1]Department of Microbiology Immunology and Pathology, College of Veterinary Medicine and Biomedical Sciences, Fort Collins, CO, USA and [2]The Connecticut Agricultural Experiment Station, Center for Vector Biology and Zoonotic Diseases, New Haven, CT, USA

*Corresponding author: E-mail: doug.brackney@ct.gov

[†]Present address: Department of Immunology and Microbial Science, The Scripps Research Institute, La Jolla, CA, USA.

[‡]These authors contributed equally to this work.

Abstract

Arthropod-borne RNA viruses exist within hosts as heterogeneous populations of viral variants and, as a result, possess great genetic plasticity. Understanding the micro-evolutionary forces shaping these viruses can provide insights into how they emerge, adapt, and persist in new and changing ecological niches. While considerable attention has been directed toward studying the population dynamics of mosquito-borne viruses, little is known about tick-borne virus populations. Therefore, using a mouse and *Ixodes scapularis* tick transmission model, we examined Powassan virus (POWV; *Flaviviridae, Flavivirus*) populations in and between both the vertebrate host and arthropod vector. We found that genetic bottlenecks, RNAi-mediated diversification, and selective constraints collectively influence POWV evolution. Together, our data provide a mechanistic explanation for the slow, long-term evolutionary trends of POWV, and suggest that all arthropod-borne viruses encounter similar selective pressures at the molecular level (i.e. RNAi), yet evolve much differently due to their unique rates and modes of transmission.

Key words: flavivirus; evolution; Powassan virus; RNAi; Ixodes; tick.

1. Introduction

Arthropod-borne RNA viruses (arboviruses) pose a continuous threat to human health worldwide. This is highlighted by the rapid global emergence of Zika virus (Fauci and Morens 2016), West Nile virus (WNV) (Kramer et al. 2008), and chikungunya virus (Weaver and Lecuit 2015), and the local emergences of several tick-borne viruses (Mansfield et al. 2009; Ebel 2010; McMullan et al. 2012; Liu et al. 2014). While environmental disturbances and animal behavior present new ecological opportunities, arbovirus emergence is driven by their inherent genetic plasticity. RNA viruses replicate quickly using error-prone polymerases which produce large, genetically diverse intrahost populations that facilitate their adaptation to novel environments (Holmes 2009). The virus–host interactions that influence the genetic diversity of many single-host viruses are well studied, but arbovirus evolution is comparatively complex. For instance, influenza A virus and HIV-1 have substitution rates of $>10^{-3}$ substitutions per site per year (Fusaro et al. 2011; Alizon and Fraser 2013), while the rates of dengue virus and WNV evolution are ~tenfold slower (Jenkins et al. 2002; Snapinn et al.

2007). This difference is attributed to the requirements of arboviruses to replicate within divergent hosts to maintain transmission (Woelk and Holmes 2002). Despite these constraints, there is still evidence for positive selection that can increase transmission by both arthropod vectors (Moudy et al. 2007; Tsetsarkin et al. 2007, 2014) and vertebrate hosts (Brault et al. 2007) by mechanisms that are not completely understood.

Our current understanding of the micro-evolutionary forces shaping arboviral populations is almost exclusively informed by mosquito-driven systems. Of these, the WNV (*Flaviviridae*, *Flavivirus*) transmission model is the best studied. WNV is rapidly transmitted between mosquitoes and birds during short periods of intense transmission (i.e. months). During mosquito infection, WNV populations are shaped by two primary forces; RNAi-mediated diversification and genetic drift associated with anatomical barriers to infection (i.e. bottlenecks) (Brackney et al. 2009, 2015; Grubaugh et al. 2016). Populations experience further bottlenecks during horizontal transmission and are subject to strong purifying selection in their avian host (Grubaugh et al. 2015, 2016). Collectively, purifying and diversifying selection, bottlenecks, and the rate and mode of transmission facilitate WNV evolution (Brackney et al. 2009; Ciota et al. 2012; Grubaugh et al. 2016).

In contrast, comparatively little is known about the selective pressures that viruses encounter during tick transmission. Like mosquitoes, ticks have an intact and fully functional RNAi pathway that targets viral genomes in cell culture (Garcia et al. 2005, 2006; de la Fuente et al. 2007; Barry et al. 2013; Schnettler et al. 2014); however, its role in tick-borne virus evolution remains to be determined. Beyond this similarity, several physiological and life history traits differentiate ticks and mosquitoes. For example, ixodid tick blood feeding episodes last several days, occur once per life stage, and bloodmeals are digested intracellularly over the course of weeks or months while molting (Evans 1992). Consequently, ticks feed once every couple of months or years thereby limiting the number of horizontal transmission events. In addition, ticks regularly acquire viral infections through alternative transmission pathways such as vertical transmission (mother to offspring) (Costero and Grayson 1996), co-feeding (a naïve tick becomes infected while feeding in close proximity to an infected tick) (Costero and Grayson 1996; Jones et al. 1997), and transstadially (viruses being passed from one life stage to the next) (Costero and Grayson 1996; Ebel and Kramer 2004). Elucidating the relative importance of these transmission events and RNAi on tick-borne virus evolution is critical to understanding the adaptive potential of these viruses.

Powassan virus (POWV; *Flaviviridae*, *Flavivirus*), the sole North American member of the tick-borne encephalitis virus (TBEV) sero-complex, consists of two genetic lineages that are maintained in distinct enzootic transmission cycles (Costero and Grayson 1996; Telford et al. 1997; Ebel et al. 2000). Maximum-likelihood and Bayesian phylogenetic inferences indicate that tick-borne flaviviruses, including POWV, evolve gradually and are extremely stable in enzootic foci (Pesko et al. 2010; Heinze et al. 2012; Kovalev and Mukhacheva 2014). This can in part be explained by the extremely low intrahost population diversity observed in field-collected adult *Ixodes scapularis* ticks (Brackney et al. 2010). While these findings are indicative of gradual evolution, evidence also suggests that tick-borne flaviviruses occasionally experience periods of rapid diversification during adaptation to new vector species (Simpson 1944; Kovalev and Mukhacheva 2014). An example of this has been proposed for the radiation of TBEV subtypes in Europe and Asia (Kovalev and Mukhacheva 2014). Such events could be facilitated by a combination of hybrid tick species and co-feeding transmission (Jones et al. 1997; Balashov Iu et al. 1998). Similarities between the evolution and ecology of TBEV subtypes and POWV lineages suggest that these two systems may have experienced analogous evolutionary paths. While these studies highlight the macro-evolutionary trends of tick-borne flaviviruses, the forces shaping their intrahost populations (i.e. micro-evolution) and the concomitant effects on their adaptive potential represent a significant knowledge gap.

In this study, we defined the extent to which horizontal transmission from mice to ticks, transstadial transmission between tick life stages, and the antiviral RNAi response during long-term tick infection impact POWV genetic diversity. Using next-generation sequencing (NGS), we characterized the intrahost POWV populations and the virus-derived small RNA (vsRNA) in ticks and discovered divergent evolutionary processes between the two transmission routes. We observed that POWV populations quickly diversify within mice following intra-peritoneal inoculation. Subsequently, POWV populations encounter a severe genetic bottleneck during horizontal transmission similar to arbovirus transmission by mosquitoes (Ciota et al. 2012; Forrester et al. 2012; Gutierrez et al. 2015; Grubaugh et al. 2016; Lequime et al. 2016). Once in ticks the viral populations stabilize as a result of relaxed bottlenecks associated with transstadial transmission and selective constraint against RNAi-mediated diversification. Together, these data provide an experimental explanation for the slow, long-term rates of POWV evolution, and may be indicative of the evolutionary trajectories of other tick-borne viruses.

2. Methods

2.1 Ethics statement

Experiments involving ticks and mice (protocol P25-15) were approved by the Connecticut Agricultural Experiment Station (CAES) Institutional Animal Care and Use Committee. The protocol adhered to the guidelines outlined in the Animal Welfare Act Regulations (CAES IACUC: OLAW Animal Welfare Assurance Number: A4050-01).

2.2 Mouse and tick infections

POWV lineage II deer tick virus (strain P0375; GenBank accession number KU886216) was originally recovered from *Ix. scapularis* ticks in 2010 from Bridgeport, CT. The virus was isolated after one passage on BHK-21 cells. *Ix. scapularis* ticks from the CAES colony were used for the experiments. Six-week old, female BALB/C mice were ordered from Charles River Laboratories (Wilmington, MA) and housed in A-BSL3 for 1 week prior to beginning the experiments. Mice were intraperitoneally injected with $\sim 1 \times 10^4$ plaque forming unit (PFU) POWV diluted to 0.1 ml in endotoxin-free PBS. Following infection, mice were allowed to recover for 1 h prior to tick infestation. Due to the difficulty in collecting sera from tick-infested mice, sera from three additional un-infested mice were collected 3 days post inoculation (dpi) in order to characterize POWV populations from the vertebrate host. In nature, *Ix. scapularis* larvae and nymphs can become infected with POWV, therefore in order to model both scenarios mice were either infested with ~200 *Ix. scapularis* larvae (transmission pathway #1) or ~40 *Ix. scapularis* nymphs (transmission pathway #2). Ticks were allowed to feed to repletion, at which point they dropped off the mice and were collected. Larvae and nymphs were held at RT in desiccators

containing 10% potassium sulfate solution and subjected to a 16/8 h light/dark photoperiod. For transmission pathway #1, a subset of ten engorged ticks were collected 3 weeks post feed, placed in PBS-G [phosphate-buffered saline with 0.5% gelatin, 30% rabbit serum, and 1% 100× antibiotic–antimycotic (10,000 μg/ml of streptomycin and 25 μg/ml of amphotericin B)], and macerated with a mixer mill. The remaining larvae were allowed to molt and upon emergence, the nymphs from transmission pathway #1 (i.e. infected as larvae) were cold anesthetized and the third right leg from forty-one ticks was removed in order to determine the infection status for each individual by qRT-PCR. Subsequently, infected nymphs were individually placed on naïve mice. Because others have shown that members of the TBEV sero-complex can be transmitted by co-feeding (Randolph 2011), we used one mouse per infected nymph. After feeding to repletion, engorged nymphs were collected and housed as before and three were killed 3 weeks post bloodfeeding (pbf). The remaining ticks were allowed to molt and adults 1 week post molting were collected. Similarly, adults from transmission pathway #2 (i.e. infected as nymphs) were collected 1 week post molting and screened for POWV infection by qRT-PCR (Fig. 1).

2.3 RNA extractions and qRT-PCR

Throughout the experiment, ticks from all life stages were harvested in 300 μl PBS-G, macerated on a mixer mill and RNA extracted from 50 μl of the homogenate. Total RNA was recovered using the Mag-Bind Viral DNA/RNA 96 Kit (Omega Bio-tek Inc., Norcross, GA) on a Kingfisher Flex (Thermo Fisher Scientific, Waltham, MA) automated nucleic acid extraction device. Carrier RNA was not used in the extractions and samples were eluted in 50 μl ddH_2O.

POWV from tick and mouse samples were quantified by qRT-PCR using a newly designed assay. In brief, an ~1,000 bp region of the POWV nonstructural protein 5 was amplified with a forward primer containing the T7 promoter and a nonmodified reverse primer (Supplementary Table S3). This amplicon was purified and used as template to produce RNA transcripts using the T7 Megascript Kit (Thermo Fisher Scientific) according to the manufacturer's instructions (Life Technologies, Carlsbad, CA). RNA quality and purity was assessed on a Bioanalyzer 2100 (Agilent, Santa Clara, CA) and quantified on a Qubit Fluorometer (Thermo Fisher Scientific). Using POWV-specific primers and probe (Supplementary Table S3), the sensitivity of the assay was assessed on a CFX96 Touch Real-Time PCR Detection System (Bio-Rad, Hercules, CA) using the following parameters; RT—50°C for 20 min, 95°C for 2 min, PCR—95°C for 15 s, 60°C for 30 s followed by a plate read. The PCR was repeated for 40 cycles. 20 μl reactions with either 5 μl of standard or sample RNA were used with the TaqMan Universal Master Mix II reagents (Thermo Fisher Scientific) and the data were analyzed using the Bio-Rad CFX Manager 3.1. We detected 10^2–10^8 POWV genome equivalents (GE) GE/reaction with a primer efficiency of 93% with an R^2 value of 0.999, a slope of −3.502, and y-intercept = 43.27.

2.4 Total RNA library preparation and data processing

Total RNA from tick, mouse, and input viral supernatant were used for library preparation. In brief, cDNA was generated and amplified using the Ovation RNA-Seq System V2 (NuGEN, San Carlos, CA). The majority of the samples total RNA concentrations were too low to quantify so the maximum allowable volume was used for input (15 μl). For those with quantifiable amounts of RNA ~75 ng of total RNA was used. Samples were quantified after the cDNA amplification and 1 ng of cDNA was used as template for library production using the Nextera XT Kit (Illumina, San Diego, CA). Each of the 24 samples was uniquely barcoded. Finished libraries were analyzed for correct size distribution using Bioanalyzer 2100 (Agilent) and concentration using the Library Quantification qPCR kit for Illumina platforms (Kapa Biosystems, Wilmington, MA). 100 nt paired-end reads were generated using the Illumina HiSeq 2500 platform (GENEWIZ, South Plainfield, NJ).

Demultiplexed reads from the three input virus replicates were *de novo* assembled using Trinity (Grabherr et al. 2011) to build the consensus POWV genome, which was deposited in GenBank (accession number KU886216). All samples were then aligned to the POWV consensus sequence using MOSAIK (Lee et al. 2014) and duplicate reads were removed using the MarkDuplicates tool within Picard to limit the influence of PCR artifacts. iSNVs and iLVs (includes both single and double insertions and deletions) were called using VPhaser2 (Yang et al. 2013) and sites with significant strand bias were removed. All raw NGS data are available on SRA BioProject PRJNA342831.

2.5 Viral population genetics

Analysis was limited to the protein coding sequences. Due to low coverage of the mouse POWV populations (111–521× coverage per nucleotide per site), only variants >0.02 frequency were used for population genetic analysis, but all called intra-tick variants were used for comparisons to the vsRNA profiles. Richness was calculated by the sum of the variant sites >0.02 frequency detected in each intrahost population. Genetic distance was calculated by the sum of the nucleotide and amino acid variant frequencies from each population and reported as the variants per coding sequence. Complexity, measured by the

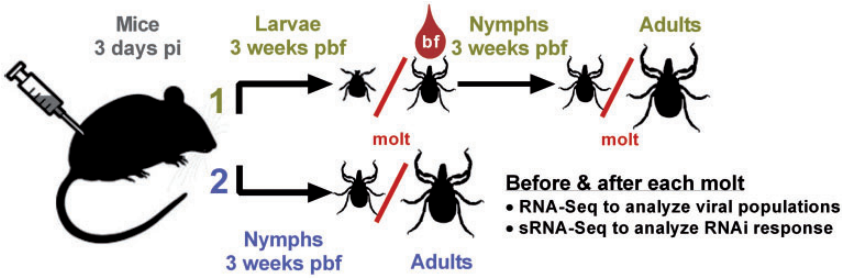

Figure 1. Experimental POWV transmission. Immature ticks were allowed to feed on mice from 0 to 3 dpi to initiate two transstadial (from one life stage, or stadium, to the next) transmission pathways. In pathway 1, engorged larvae molted into nymphs, were offered a non-infectious bloodmeal, and then molted into adults. In pathway 2, engorged nymphs molted into adults. In both pathways, whole ticks were collected at 3 weeks pbf and within 1 week post-molt (at each tick symbol). The POWV genetic population structures were determined from the input virus, mouse serum, and tick bodies. vsRNA was profiled from the matching tick bodies.

Shannon entropy (S), was calculated for each population (i) using the variant frequency (p) at each nucleotide position (s):

$$S_{i.s} = -p_s(\ln p_s) + (1-p_s) \times \ln(1-p_s) \quad (1)$$

The mean S from all sites s is used to estimate the mutant spectra complexity. F_{ST} (fixation index) was used to estimate genetic divergence between two viral populations as described (Fumagalli et al. 2013):

$$a_s = \frac{4n_i(p_{i.s} - p_s)^2 + 4n_j(p_{j.s} - p_s)^2 - b_s}{2(2n_in_j/(n_i+n_j))} \quad (2)$$

and

$$b_s = \frac{n_i(2p_{i.s}(1-p_{i.s}) + n_j(2p_{j.s}(1-p_{j.s})}{n_i+n_j-1}, \quad (3)$$

where $p_{i,s}$, $p_{j,s}$, and p_s are the frequencies of the input POWV consensus nucleotide at site s from populations i, j, and the consensus site (i.e. 1), respectively. Only VPhaser2-called variants >0.02 frequency were used, all other sites $P=1$. The number of individuals sampled, n, was set to the lowest coverage depth accepted (100 nt) to normalize for sequencing variations. The estimate of F_{ST} for the protein coding *locus* of m sites (10,245 nt) is

$$F_{ST}^{(locus)} = \frac{\sum_{s=1}^{m} a_s}{\sum_{s=1}^{m}(a_s+b_s)} \quad (4)$$

The effective population size (N_e) initiating infection in the mice or during horizontal transmission to larvae and nymphs (i.e. by blood feeding on mice) was determined by the genetic variance caused by drift within and between populations using F_{ST} (Monsion et al. 2008):

$$N_e = \frac{1 - F_{ST}}{F'_{ST} - F_{ST}}, \quad (5)$$

where F_{ST} is the genetic variance between the founding populations (i.e. input virus) and F'_{ST} is the genetic variance between the final populations (i.e. mice 3 days pi or ticks at all stages). For N_e, F_{ST} was calculated using only synonymous variants sites (to minimize the influence of selection) and from all replicate combinations per group (i.e. replicate A vs B, A vs C, and B vs C) using Equations (2)–(4).

Intrahost selection was estimated by the ratio of nonsynonymous (d_N) to synonymous (d_S) SNVs per site (d_N/d_S) using the Jukes-Cantor formula (Jukes and Cantor 1969)

$$d_N = \frac{-3 \times \mathrm{Ln}(1-((4p_n)/3)}{4} \quad (6)$$

and

$$d_S = \frac{-3 \times \mathrm{Ln}(1-((4p_s)/3)}{4}, \quad (7)$$

where p_n equals *Nd* (sum of the nonsynonymous variant frequencies >0.02, as compared with the intrahost consensus sequence) divided by the number of nonsynonymous sites and p_s equals *Sd* (sum of the synonymous variants >0.02) divided by the number of synonymous sites. DnaSP (Librado and Rozas 2009) was used to determine the number of nonsynonymous (7,784.33) and synonymous (2,460.67) sites from the intrahost consensus sequence using the Nei-Gojorori method (Nei and Gojobori 1986). d_N/d_S values >1 for divergent lineages are the hallmark of positive selection, but the power of d_N/d_S to detect positive selection within hosts is very low considering that the majority of the nonsynonymous mutations are under strong purifying selection (Kryazhimskiy and Plotkin 2008). Therefore, we only used the d_N/d_S ratio to cautiously estimate the overall strength of purifying selection.

UpA and CpG dinucleotide frequencies were calculated as described (Blitvich and Firth 2015). We substituted each variant nucleotide into the consensus sequence for each viral population analyzed. Then the numbers of UpA and CpG dinucleotides and A, C, G, and U mononucleotides were counted. Dinucleotide frequencies were expressed relative to their expected frequencies

$$\frac{f_{XpY}}{f_X \times f_Y}. \quad (8)$$

2.6 Small RNA library preparation and data analysis

Small RNA from the tick samples was isolated using a slightly modified protocol with the Mag-Bind Viral DNA/RNA 96 Kit (Omega Bio-tek Inc.). Specifically, 2 µl of linear acrylamide was added to the lysis solution and after lysing the samples in lysis buffer/isopropanol for 10 min an additional aliquot of isopropanol was added to each sample prior to performing the extractions. Libraries were prepared using the NEBNext Multiplex Small RNA Library Prep Set for Illumina (New England Biolabs, Ipswich, MA) which generates strand-specific sRNA libraries. The libraries were first size selected using magnetic beads and then on a BluePippin (Sage Science, Beverly, MA) with a 3% gel without ethidium bromide using a range of 120–170 bp. The size-selected libraries were quantified on a Qubit Fluorometer (Thermo Fisher Scientific) and analyzed on the Bioanalyzer 2100 (Agilent). Four of the tick libraries did not have products within the desirable 145–155 bp size class and thus, were not pooled with the other samples; these included one post-molt nymph from transmission pathway #1, two 3 weeks pbf nymphs from transmission pathway #1, and one 3 weeks pbf nymph from transmission pathway #2. Pooled libraries were sequenced on the Illumina NextSeq500 platform at the Colorado State University IDRC Genomics Core using 75 cycles of single read high output sequencing.

The generated sRNA sequencing data (FASTQ files) were then analyzed using a pipeline established in our laboratory. Using the pipeline, FASTQ files were trimmed of the 3′ adapter using FASTX Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/), size selected (initially all 19–32 nt reads; then 19–23 nt, and 24–32 nt) and aligned to the POWV consensus sequence of the respective individual samples using Bowtie 0.12.8 (Langmead et al. 2009) allowing for 1-mismatch. The -a –best –strata mode was used, which instructs Bowtie to report only those alignments in the best alignment stratum. SAM output files produced by Bowtie were used as the input for processing through SAMtools (Li et al. 2009). From the individual read output, histograms were generated to show overall size and polarity distribution of vsRNA reads. Nucleotide targeting of the viral genome (i.e. the number of vsRNA reads covering each nt position of the WNV genome) was determined using the mpileup function of SAMtools for 19–23 nt reads and 24–32 nt reads separately. GC

content, nucleotide bias of specific positions, and correlations of targeting intensity among experimental replicates and tick life stages using Microsoft Excel and Graphpad Prism 6.

3 Results

3.1 POWV population demographics during transmission

We initiated POWV transmission cycles similar to a previously described study (Costero and Grayson 1996) to determine how tick-borne viruses evolve within and between hosts. First, BALB/C mice were inoculated with POWV and immature *Ix. scapularis* ticks (larvae and nymphs) were allowed to feed on the mice starting on the day they were inoculated through 3 dpi. Engorged larvae and nymphs were used to start two transmission pathways (Fig. 1). In the first transmission pathway, engorged larvae molted into nymphs, which were allowed to feed on uninfected mice to molt into adults. In the second transmission pathway, engorged nymphs were allowed to molt into adults. Three biological replicate ticks were collected 3 weeks pbf and within 1 week after molting in each pathway to be used for total RNA and small RNA (sRNA) NGS.

Intrahost POWV population sizes (POWV GE $\geq$ per tick) continuously increased after molting and during the 3-week extrinsic incubation (EI) period (Fig. 2A), but to a lesser extent than what was previously demonstrated (Ebel and Kramer 2004). RNA from these samples was then used for NGS. On average, 1.9% of reads from ticks aligned to the POWV genome, resulting in >500× coverage per sample (Supplementary Fig. S1A and Table S1). However, only 0.2% of the reads from mouse serum aligned to the POWV genome and we only obtained 111–

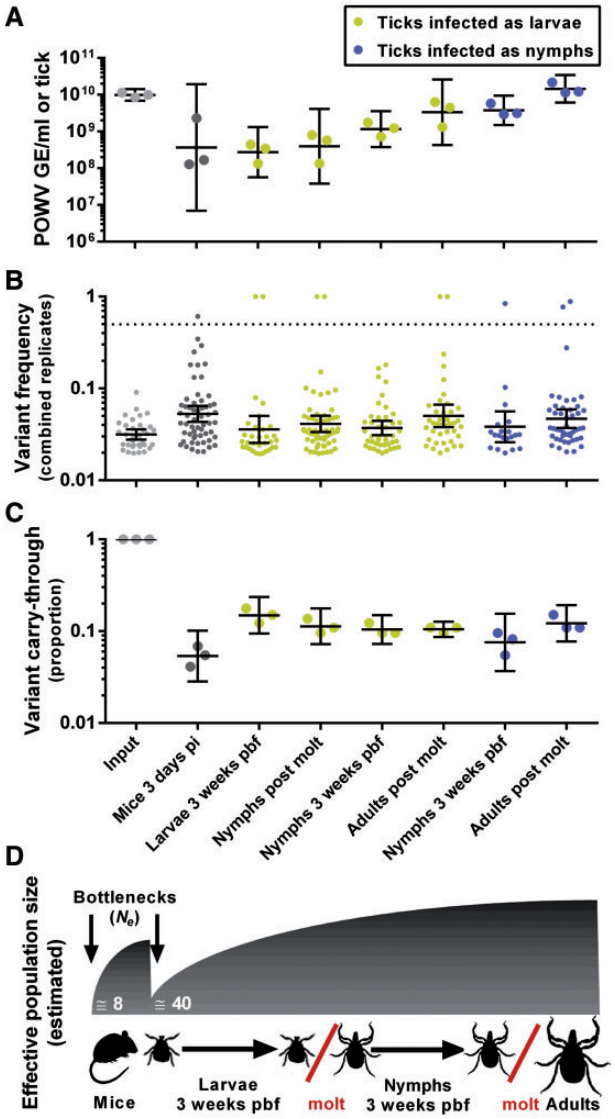

Figure 2. Bottlenecks occur during horizontal but not transstadial transmission. (**A**) POWV genome equivalents (GE) per ml input virus, ml mouse serum, or whole tick bodies were determined by qRT-PCR (geometric mean with 95% CI). (**B**) The POWV variants detected in the intrahost populations were plotted by frequency of occurrence. (**C**) The proportion of input variant sites carried through transmission was determined from each intrahost population. All data are presented as the geometric mean with 95% CI. The yellow data points are from transmission pathway 1 (ticks infected as larvae) and the blue data points are from transmission pathway 2 (ticks infected as nymphs) (Fig. 1). (**D**) The effective population sizes were calculated by F_{ST} (Equation 5) for mouse (1–31 founding POWV genomes) and tick infection (1–89 founding POWV genomes), but were estimated for transstadial transmission using the POWV GE per tick (A), variant frequency distributions (B), and variant carry-through (C).

Table 1. POWV consensus changes during horizontal and transstadial transmission.

Host (replicate)	Genome position	Nucleotide change	Coding region	S/N[a]	Amino acid change	iSNV freq
Mouse 3dpi (A)	4652	C→A	NS3	S		0.61
Larva 3wpbf (A)	1961	G→T	E	S		1
Larva 3wpbf (C)	564	G→A	prM	N	G→R	1
Nymph pm (A)	5474	T→C	NS3	S		1
Nymph pm (B)	1281	G→T	E	N	V→L	1
Adult pm (C)	1376	C→T	E	S		1
	4844	A→G	NS3	S		1
*Nymph 3wpbf (C)	5725	A→G	NS3	N	Q→R	0.84
*Adult pm (A)	4529	A→G	NS2B	S		0.89
*Adult pm (B)	4043	G→T	NS2A	S		0.77

[a]Synonymous or nonsynonymous variant
*Ticks infected as nymphs
dpi, days post inoculation; wpbf, weeks post blood feeding; pm, post molt; iSNV freq, intrahost single nucleotide variant frequency; NS3, nonstructural protein 3; E, envelope protein; prM, premembrane protein; NS2B, nonstructural protein 2B; NS2A, nonstructural protein 2A

521× coverage. To account for this, we compared our analysis of intrahost single nucleotide variants (iSNVs) and length variants (iLVs, including single and double insertions and deletions) using frequency cut-offs of 0.02, 0.05, and 0.10 (Supplementary Fig. S2). All major findings were consistent with regards to the frequency cut-off used and therefore, a frequency cut-off of 0.02 was used for all of the main text figures. Coverage was variable across the viral genome (likely due to differences in GC content—Dohm et al. 2008), but consistent between populations (Supplementary Fig. S1B). In addition, we obtained incomplete coverage of the 5′ and 3′ untranslated regions in most samples. Therefore, we limited analysis to the coding sequence.

Intrahost variants were used to analyze population structures and to assess population bottlenecks during transmission. Virus genetic diversity (i.e. the presence of variants >0.2 frequency) increased during mouse infection, and decreased during horizontal transmission to both larvae and nymphs (Fig. 2B). The initial point of tick infection during bloodfeeding also introduced fixed or nearly fixed POWV mutations into the populations. During transstadial transmission, however, the populations diversified very slowly and while there was an increase in the number of variants >0.2 frequency and fixed mutations, these increases were not significant (Fig. 2B). Genetic bottlenecks were first assessed by the number of POWV variants that carried-through from the input populations during transmission to mice and ticks (Fig. 2C). We found that only a small proportion of variants from the input virus were detected in the mice and ticks. However, the proportional carry-through of the input variants in the ticks did not decrease during transstadial transmission. We further assessed transmission bottlenecks by calculating the effective population size (N_e) of POWV genomes using the genetic variance within and between populations caused by drift of neutral alleles (Equation 5—Monsion et al. 2008). From the input virus, we found that ∼8 (1–31 95% CI) POWV genomes initiated infection in the mice and ∼40 (1–89 95% CI) POWV genomes initiated infection in the ticks, confirming the presence of tight (i.e. narrow) bottlenecks at those transmission points. We were not able to calculate N_e during transstadial transmission because the viral populations are not in direct ancestral descent (i.e. the adult tick POWV populations did not come directly from the nymphs that we sampled). Based on POWV GE (Fig. 2A), variant frequency distributions (Fig. 2B), and variant carry-through (Fig. 2C) (i.e. variants found in both pre- and post-metamorphosis samples) we determined that the effective population size remained large (i.e. wide bottlenecks) during transstadial transmission (Fig. 2D).

3.2 POWV genetic diversity and selection

Ten consensus changes (frequency >0.5) were detected in the POWV populations as compared with the input virus. Of these, seven were synonymous, and tended to occur singly in individual specimens (except for adult post molt replicate C which had two synonymous substitutions, Table 1). Though some of these high frequency mutations may have arose through positive selection, we predict that their occurrence was mostly due to genetic drift during horizontal transmission. We analyzed intrahost viral populations using several genetic diversity indices. Virus genetic complexity (i.e. the uncertainty associated with randomly sampling an allele, measured by Shannon entropy) was significantly higher in mice than ticks (Fig. 3A, $P < 0.05$, Kolmogorov–Smirnov comparison of cumulative distributions). Though genetic distance (i.e. the accumulation of mutations compared with the input consensus sequence) was not significantly different between the hosts, distance was more variable in the ticks likely due to additional genetic drift during initial tick infection. Complexity and distance were not significantly different among tick life stages (Fig. 3A, Kruskal–Wallis–Dunn's corrections), regardless if they were infected as larvae or nymphs (Supplementary Fig. S3). Therefore, genetic diversity then remained relatively stable during transstadial transmission. Intra-tick nucleotide diversity was 0.00007–0.00015 (95% CI of mean), similar to field-collected *Ix. scapularis* (0.00006–0.000012—Brackney et al. 2010; $P = 0.4$, unpaired *t*-test).

Selection was measured by the variant frequency compared with the intrahost consensus sequence, as opposed to the frequency compared with the input consensus sequence used for genetic diversity analysis. We first examined intrahost selection pressures by calculating the ratio of nonsynonymous (d_N) to synonymous (d_S) substitutions per coding sequence site (d_N/d_S) (Fig. 3, Supplementary Fig. S3G). As indicated by d_N/d_S ratios <1, purifying selection was very strong in mice but became quite variable ($d_N/d_S = 0.1–1.5$) during tick transstadial transmission. We also examined selection for host codon usage specifically based on dicucleotide bias for CpG and UpA (Kunec and Osterrieder 2016). Mice display a strong underrepresentation of CpG dinucleotide usage compared with *Ix.*

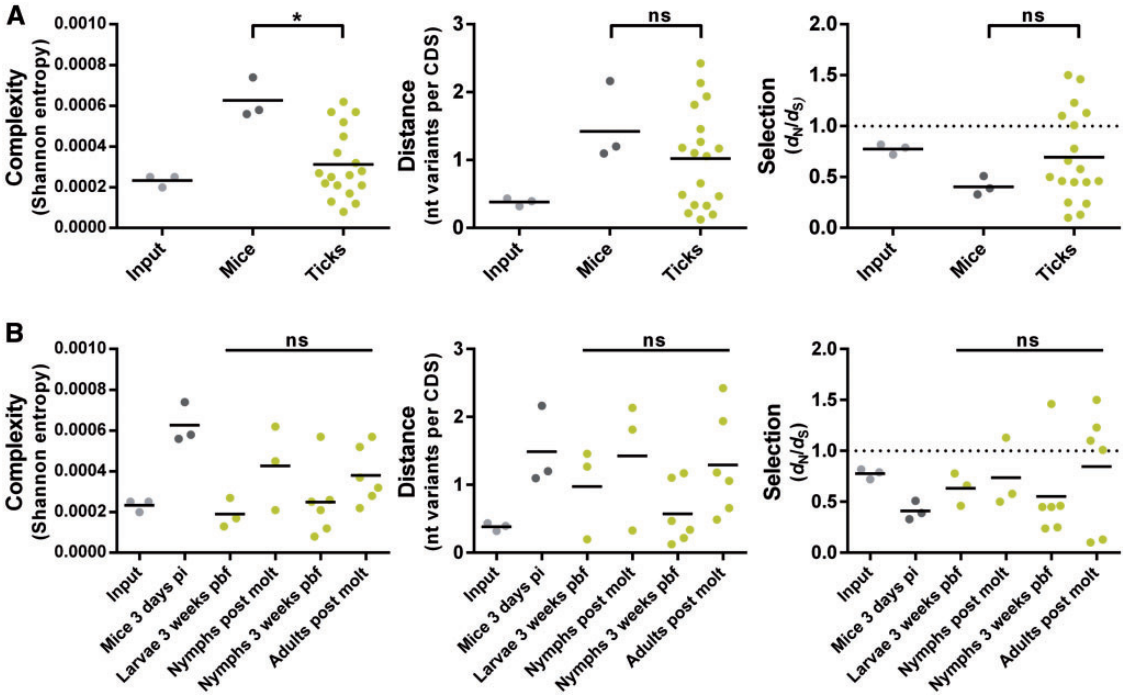

Figure 3. Viral genetic complexity decreases during mouse-to-tick transmission. Intrahost virus genetic diversity and selection were compared (**A**) between mice and ticks (Kolmogorov–Smirnov) and (**B**) among tick life stages (Kruskal–Wallis–Dunn's corrections). Genetic complexity (left panel) is the uncertainty associated with randomly sampling an allele at a site and was measured by Shannon entropy (Equation 1). Genetic distance (middle panel) is the number of nucleotide changes per coding sequence compared with the input consensus. Selection was estimated by the ratio of nonsynonymous (d_N) to synonymous (d_S) substitutions per coding sequence site (d_N/d_S) within each population was used to measure the strength of purifying selection (strong <1 [Equations 6 and 7]). All data are shown as the mean.

scapularis, while UpA is about equally under-represented in both hosts (Simmen 2008; Kunec and Osterrieder 2016). Therefore, if POWV was adapting toward the host's preferred codon usage, we would expect to see an increased CpG usage during replication in ticks and decreased usage in mice, as demonstrated for insect-specific and no known vector flaviviruses, respectively (Lobo et al. 2009; Blitvich and Firth 2015). The viral CpG usage did increase from adult ticks, but not significantly compared with intra-mouse populations (Supplementary Fig. S4). Therefore, adapting to host dinucleotide usage is unlikely a significant source of selection during this relatively short evolutionary time frame.

3.3 Tick small RNA response to POWV infection

Ixodes scapularis vsRNA responses were characterized in all tick life stages in order to determine whether vsRNA targeting influences POWV evolution. sRNA libraries from all but four of the previously described samples were sequenced (Supplementary Table S2) and 19–32 nt reads were aligned to the POWV genome (allowing for one mismatch). The percentage of POWV-derived sRNA reads varied between samples (Supplementary Table S2) and was highest in adult ticks that had been infected as nymphs (4.27 ± 0.76%). However, size distribution of vsRNAs was nearly identical between larvae, nymphs and adult ticks (Fig. 4A–C), all of which predominantly generated 22 nt vsRNAs with little strand bias (65% genome-derived, 35% antigenome-derived). Only a few 24–32 nt vsRNA reads, indicative of putative virus-derived piRNAs (vpiRNAs), were detected in all of the life stages. 19–23 nt vsRNAs were distributed along the POWV genome with the highest frequency reads derived from the structural region of the genome (Fig. 4D). This rather striking bias of vsRNA targeting for this part of the genome was seen on both the genome and antigenome across all samples and life stages. Increased vsRNA targeting of the structural region did not simply correlate with higher abundance of viral 5′ total RNA (Supplementary Fig. S1B). Interestingly, a single read at position 635 (TGGATTGTTTTTGCCGAGGTGT) was highly over-represented. To exclude the misleading detection of a tick-derived sRNA, the sequence was compared to the nonredundant nucleotide database on NCBI (i.e. BLASTn) and the only hit with 100% coverage and identity was POWV (no *Ix. scapularis* sequences matched more than 16 nt). One hypothesis to explain the strong 5′ bias was that high GC content could be partially responsible for an over-representation of specific reads, however, no correlation was observed between GC content and read abundance (Supplementary Fig. S5A-B). A comparison of base composition between the 22 nt vsRNAs derived from one hot spot region (genome position 337–1144 (Fig. 5C)) and two cold spot regions (genome position 6227–7031 (Fig. 5D) and 7500–8306 (Fig. 5E) of an adult tick (*Adult pm (A)) showed no striking nucleotide bias in vsRNAs aligning to the regions.

In contrast, 24–32 nt vsRNAs were scattered evenly across the genome with a few hot and cold spots (Fig. 4E). These larger vsRNAs were mainly derived from the positive sense of the genome as would be consistent with piRNAs; however, no nucleotide bias as usually seen in ping-pong amplification cycling of piRNAs (U at position 1 in antisense; A at position 10 in sense) was found in *Ix. scapularis* 24–32 nt vsRNAs (Supplementary Fig. S6A–C). 5′ read distance plots also showed no pattern of ping-pong amplification (data not shown) comparable to those previously reported in insects (Vodovar et al. 2012). The GC content of 24–32 nt vsRNAs was high (64 ± 1.79%), possibly suggesting these were degradation products that were preferentially sequenced due to high GC content. A potential problem could

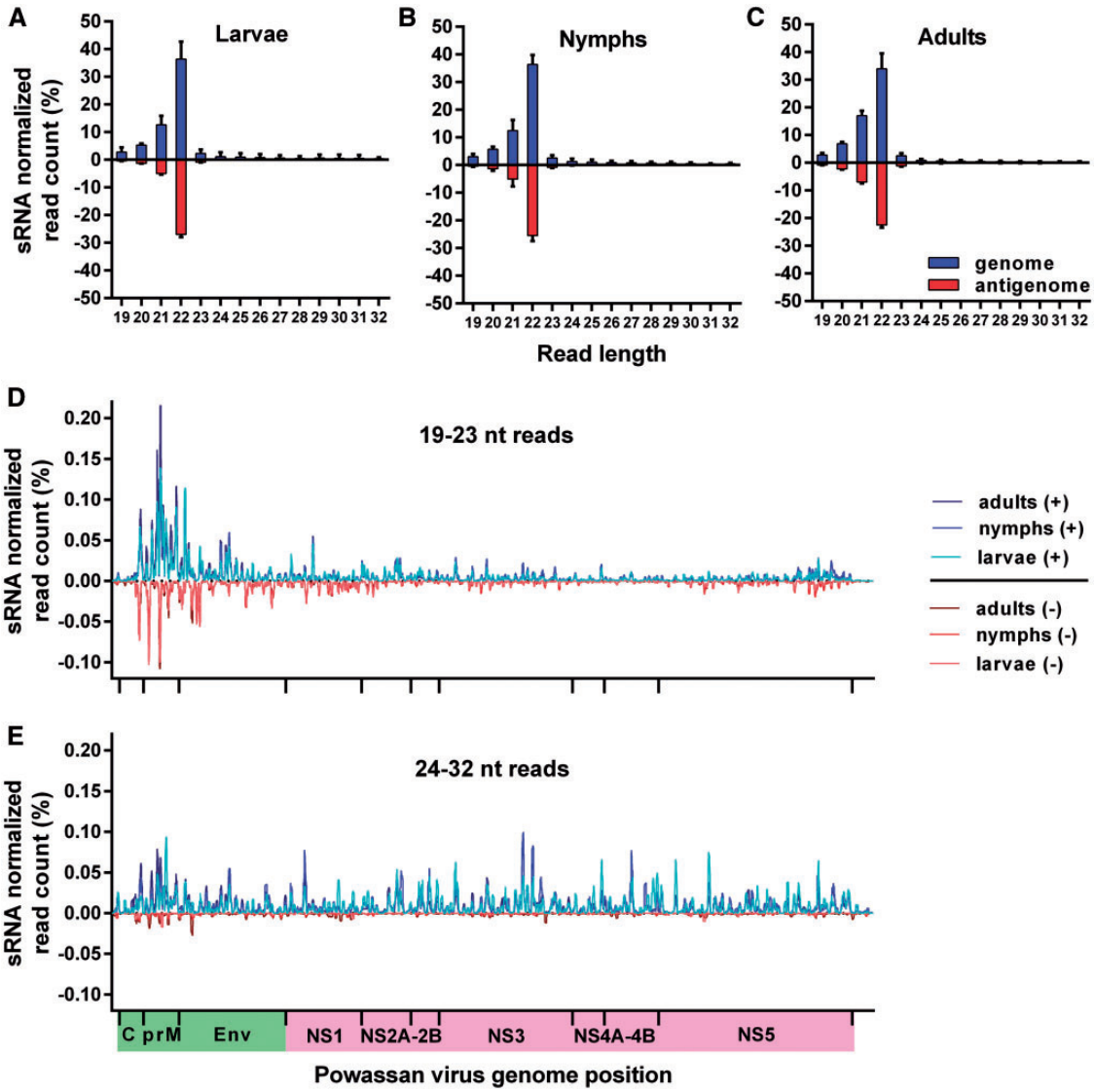

Figure 4. All *Ix. scapularis* life stages predominantly generate 22 nt vsRNAs derived from the structural region of the POWV genome and antigenome. Size distribution of vsRNA reads was combined from all three larval (**A**), five nymph (**B**), and six adult (**C**) samples. vsRNA reads derived from the genome and antigenome are depicted in blue and red, respectively. Data are shown as the mean with 95% CI. Positional targeting of 19–23 nt vsRNAs (**D**) and 24–32 nt vsRNAs (**E**) was plotted along the length of the POWV genome (blue colors) and antigenome (red colors). Data are shown as the mean of the individual life stages. The structural region of the genome is shaded in green, the nonstructural region in pink.

have been an accidental removal of most sRNAs larger than 24 nt during size selection, yet FASTQC analysis of our data showed that a considerable number of 24–32 nt reads were sequenced (Supplementary Fig. S6D–F).

We next aligned our sRNA data to the *Ix. scapularis* transcriptome (Wikel strain predicted transcript sequences, IscaW1.4 geneset) and four individual *Ix. scapularis* retrotransposons (AY682794, XM_002433619, XM_002415368, and GEFM0100 0081.1) in order to determine if any endogenous piRNAs can be detected in our samples. However, no *bona fide* piRNAs with the classic nucleotide bias associated with ping-pong amplification were found—the majority of the sRNAs aligning to the *Ix. scapularis* transcriptome were 22 nt in length (Supplementary Fig. S6G). 24–32 nt sRNAs found aligning to the *Ix. scapularis* transcriptome or retrotransposons had no nucleotide composition typical of piRNAs. However, some notable nucleotide position biases were observed—antisense 27 nt sRNA reads aligning to the *Ix. scapularis* transcriptome had a strong bias for an adenine at position 27 and a bias away from a thymine at position 1 (Supplementary Fig. S6I). Our data show that while 24–32 nt reads were sequenced, no vpiRNAs or endogenous piRNAs containing signatures of the prototypical ping-pong amplification were detected.

3.4 Virus-derived sRNA and intrahost POWV genetic diversity

It has been demonstrated, in mosquitoes, that RNAi can be a significant driver of arbovirus genetic diversification (Brackney et al. 2009, 2015). Therefore, we used the matched sRNA and total RNA data to determine how the *Ix. scapularis* RNAi response to POWV infection contributed to intra-tick viral genetic diversity (Fig. 5). First we combined the positive and negative strand 19–23 nt vsRNA data and plotted the mean normalized

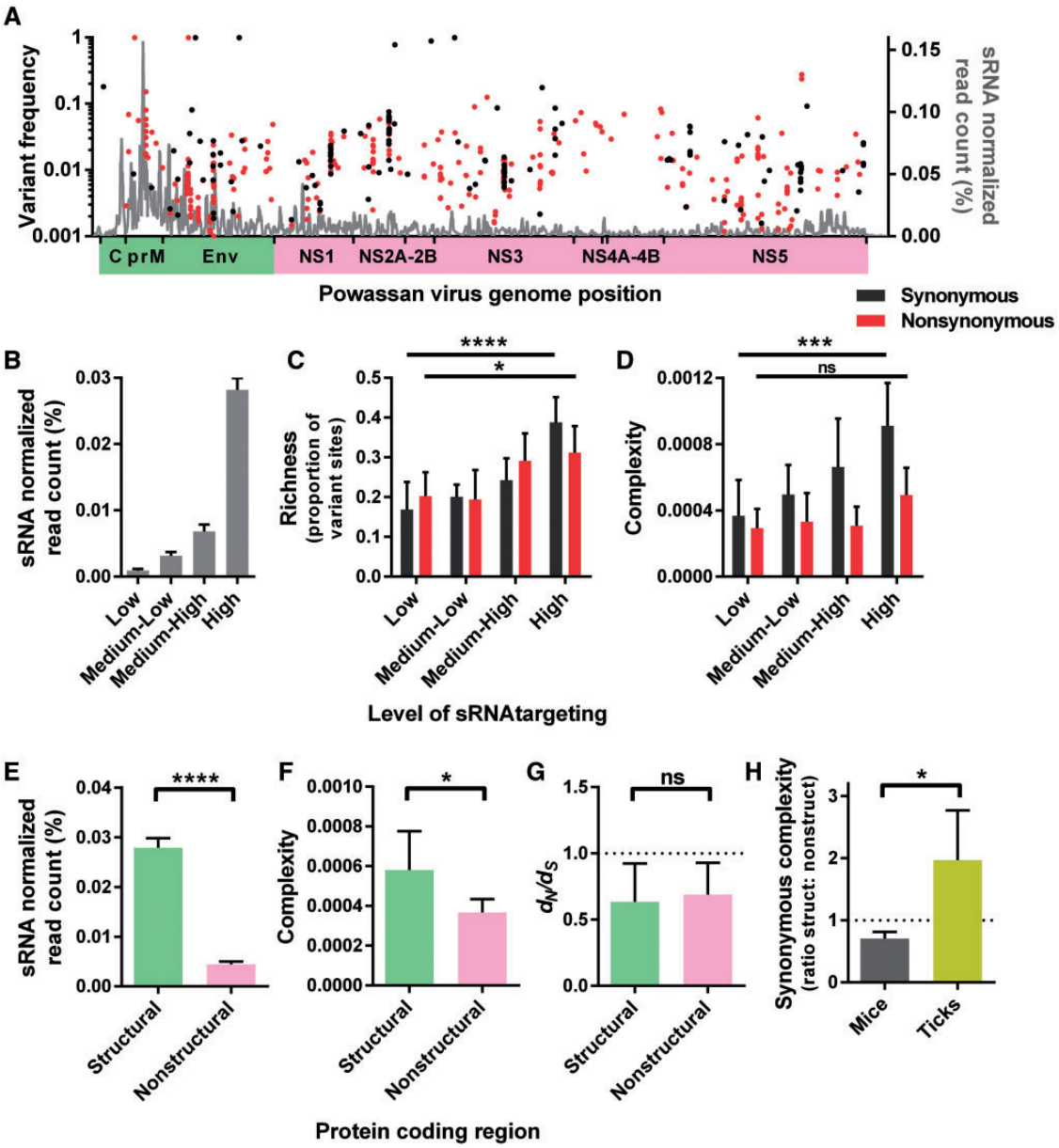

Figure 5. The effects of RNAi-mediated genetic diversification are constrained by selection. (**A**) The positive and negative strand 19–23 nt positional normalized read count profiles were combined from each life stage and replicate (grey lines, right *y*-axis) and were plotted with each POWV variant by frequency (including variants <0.02 frequency, excluding variants >0.02 frequency detected in the input virus, left *y*-axis). Black and red points are synonymous and nonsynonymous virus variants, respectively. (**B**) For each intrahost sample, the positional vsRNA normalized read counts were sorted into quartiles by percent (high, medium-high, medium-low, and low), each containing 2,562 genome positions. The data are represented by the mean with 95% CI normalized read count per quartile genome position. To examine the influence of vsRNA targeting on viral genetic diversity, intrahost viral genetic (**C**) richness, the number of variant sites and (**D**) complexity, measured by Shannon entropy (Equation 1) using only the predicted synonymous and nonsynonymous sites for each variant type, was calculated within each vsRNA read count quartile from the total RNA NGS data. These calculations were performed separately for synonymous (black) and nonsynonymous (red) mutations to determine if they are impacted differently by RNAi. Linear trends between the vsRNA quartiles were assessed by repeated measures one-way ANOVA post-tests. (**E**) The normalized vsRNA read counts, (**F**) complexity, and (**G**) d_N/d_S (Equations 6 and 7) were calculated separately for the structural and nonstructural POWV protein coding regions to assess the influence of RNAi-mediated diversifying selection in parts of the genome (paired *t*-test). (**H**) The ratios of structural (struct) to nonstructural (nonstruct) protein coding region complexity from only synonymous variants were compared between intra-mouse and intra-tick virus populations to determine if diversifying selection in the structural region was host-specific (Kolmogorov–Smirnov test). All data in bar graphs are shown as mean with 95% CI (****P <0.0001; ***P <0.001; *P <0.05; ns, not significant).

distribution profiles along with the POWV variant frequencies from each tick (Fig. 5A). Here, we included viral variants with frequencies <0.02 (since the previous 0.02 cut-off was set by the lower coverage obtained from mouse serum), but excluded variants >0.02 frequency detected in the input virus so that the variants primarily represent changes that occurred during replication within the hosts. vsRNA targeting was concentrated in the protein coding regions, but variants occurred throughout the genome. Next, we sorted the vsRNA abundance data from the protein coding sequence into quartiles based on the percent

of normalized reads along the genome position (low, medium-low, medium-high, high targeting, Fig. 5B). Each quartile contained 2,562 sites. To look for associations between the level of vsRNA targeting and viral genetic diversification, we calculated richness and complexity for the intra-tick populations within each quartile and did the analysis separately for synonymous and nonsynonymous mutations (Fig. 5C and D). Intra-quartile synonymous and nonsynonymous mutational complexity was calculated using only the predicted synonymous and nonsynonymous sites, respectively, as determined by DnaSP (Librado and Rozas 2009). Synonymous sites significantly increased in genetic richness and complexity from low to high targeting (repeated measures one-way ANOVA post-test for linear trend, richness $P < 0.0001$, complexity $P < 0.001$), but nonsynonymous sites only significantly increased in richness ($P < 0.05$). Therefore, the level of vsRNA targeting was more associated with an increase in synonymous genetic diversity than nonsynonymous. However, the proportional difference in genetic diversity of either mutation type from low to high targeting was not equivalent to the difference in the vsRNA read counts per site (Fig. 5B).

The structural protein region, containing the capsid, premembrane, and envelope protein coding sequences, contained significantly more aligned vsRNA reads per site ($P < 0.0001$, paired t test, Fig. 5E) and greater genetic complexity ($P < 0.05$, paired t test, Fig. 5F) than the nonstructural protein region. Again, the abundance of vsRNA and amount of complexity were not directly proportional, likely because purifying selection was strong ($d_N/d_S < 1$, Fig. 5G). We also compared the ratios of structural to nonstructural complexity formed only by synonymous mutations from intra-mouse and intra-tick populations (Fig. 5H). In mice, the viral populations had significantly lower structural to nonstructural synonymous complexity compared with populations in ticks ($P < 0.05$, Kolmogorov–Smirnov). This suggests that the diversifying selection targeting the structural region of POWV is limited to ticks.

4 Discussion

4.1 Viral demographics during tick-borne transmission

We created a laboratory model to assess the micro-evolutionary processes shaping POWV populations. This *in vivo* model allowed us to compare the intrahost population structure and dynamics during horizontal transmission from mice to ticks as well as during transstadial transmission between tick life stages. It should be noted that while this manuscript refers to POWV throughout, it is specific to POWV lineage 2 (DTV) and therefore, conclusions drawn from this study may not be directly transferrable to POWV lineage 1 which is maintained in an enzootic transmission cycle between *Ix. cookei* and *Ix. marxi* and medium-sized woodland rodents such as woodchucks. Using this model, we found that POWV populations isolated from mouse serum had higher variant frequencies and increased complexity compared to POWV populations isolated from ticks. However, it should be noted that we obtained low coverage from our mouse serum samples. Despite this shortcoming, we have reason to believe that our results are accurate for two reasons: (1) the variant-calling software used in these studies, VPhaser2 (Macalalad et al. 2012; Yang et al. 2013), has previously been shown to have high sensitivity (96%; ability to identify true SNVs) and a low false positive rate (0.1%) when detecting SNVs at a frequency of 0.02 from 100× coverage and (2) when the data were analyzed using highly conservative frequency cut-off thresholds (i.e. 0.05 and 0.1) the results were consistent to those obtained at 0.02 (Supplementary Fig. S2).

The demographics (i.e. population size and genetic composition) of many viruses rapidly change during transmission due to the establishment of infection by only a few viral particles (e.g. Ali et al. 2006; Wang et al. 2010; Varble et al. 2014). When N_e is small, such as during a severe transmission bottleneck, evolution is primarily influenced by the random processes of genetic drift which act to homogenize viral populations and can decrease viral fitness (Chao 1990; Duarte et al. 1992; Grubaugh et al. 2016). Similar to mosquitoes (Ciota et al. 2012; Forrester et al. 2012; Gutierrez et al. 2015; Grubaugh et al. 2016; Lequime et al. 2016), our data suggest that severe bottlenecks occur during horizontal transmission of POWV from mice to ticks. This conclusion is supported by our estimations that only one to eighty-nine POWV genomes initiated infection of the ticks. While we were unable to directly assess potential bottlenecks associated with tick transmission to mice, using intraperitoneal inoculation as a proxy, our N_e estimates suggest that severe genetic bottlenecks accompany tick to mouse transmission as well. It is unknown, however, if the amount of virus utilized in these studies (i.e. 10^4 PFU) accurately mimics the amount of virus transmitted over prolonged periods of tick feeding or if tick saliva would alter population demographics or our N_e estimates. Therefore, additional studies will be required to address these questions. Another consideration when assessing the data is our use of BALB/C mice (*Mus musculus*) as the vertebrate host. In nature, POWV lineage II is maintained in an enzootic transmission cycle between *Ix. scapularis* and the white-footed mouse, *Peromyscus leucopus*. While both mice, BALB/C mice are more closely related to rats than they are to *Peromyscus* (Bedford and Hoekstra 2015) and therefore it is unknown if BALB/C mice are an accurate model for *P. leucopus*. Interestingly, the demographics of WNV populations in young chickens do not significantly differ from wild-caught robins, crows or sparrows (Grubaugh et al. 2015). Therefore, if anything like WNV in birds, BALB/C mice may be a suitable surrogate for *P. leucopus* when assessing POWV population demographics in a mammalian host.

As opposed to mosquitoes that can take multiple bloodmeals as adults, ixodid ticks only take one bloodmeal per life stage (Evans 1992). Consequently, virus infecting an immature tick (i.e. larvae or nymph) must persist through molting to be subsequently transmitted. Our data suggest that during transstadial transmission there are relaxed (i.e. wide) bottlenecks that allow most of the viral variants to persist between the tick life stages and thereby limiting the effects of genetic drift and founder's effect during this process. The reasons for this observation are unknown, but may be related to the very poorly understood process of molting. During molting, it appears that many of the internal tissues are reorganized in the new life stage (Evans 1992). One hypothesis explaining the relaxed bottlenecks associated with transstadial transmission is that populations are able to freely circulate throughout this seemingly disordered process. As a result, POWV populations would not encounter the same anatomical barriers to transmission as WNV populations encounter during the extrinsic incubation period (EIP) of adult mosquitoes (Grubaugh et al. 2016). This is supported by two observations: (1) once ticks become infected, POWV is likely transstadially transmitted with 100% efficiency (data not shown) and (2) newly emerged life stages (i.e. nymphs or adults) are able to transmit POWV immediately upon refeeding (Costero and Grayson 1996; Ebel and Kramer 2004). Alternatively, our observations could be an artifact of sampling

whole ticks. Molting can take months to complete, this timeframe is considerably longer than the average EIP associated with arboviruses in mosquitoes (i.e. ~7–14 days). Consequently, POWV populations may similarly traverse anatomical barriers prior to the completion of molting, replicate and reach stasis prior to emergence of the next life stage. If POWV populations located in the midgut diverticulae (i.e. the initial site of infection) comprise a disproportionate majority of the POWV populations in ticks, our analysis may miss unique subpopulations in other tissue compartments (Grubaugh et al. 2016). Clearly additional research is warranted.

4.2 Tick RNAi response to viral infection

The rationale for characterizing the RNAi response to POWV infection was twofold, (1) to determine if POWV RNA is targeted by the RNAi response during a natural infection *in vivo* and (2) to determine if RNAi targeting promoted POWV diversification in ticks. The results of our vsRNA profiling have largely confirmed *in vivo* what other studies have previously observed *in vitro* (Schnettler et al. 2014). Notably, vsRNAs were predominantly 22 nt in length and originated from both the genome and antigenome with a strong targeting bias toward the 5′-end and to a lesser degree the 3′-end of the genome. In addition, there was a subpopulation of 24–32 nt vsRNAs, yet these lacked signatures of ping-pong amplification typically associated with piRNAs. This observation may be due to the fact that *Ix. scapularis* lacks an Ago3 homologue which is integral for ping-pong amplification (Schnettler et al. 2014). These characteristics are consistent with findings in cell culture (Schnettler et al. 2014), but they are quite different than those observed in other arbovirus–arthropod systems (Brackney et al. 2009; Schnettler et al. 2013) and further experimentation will be required to elucidate the significance of these observations.

Antiviral RNAi promotes diversification by selecting for rare variants that can avoid sequence complementarity, thus driving variants towards maximum complexity (i.e. 0.5 frequency). Similar to mosquitoes (Brackney et al. 2009; Brackney et al. 2015), we find an association between vsRNA targeting and genetic diversification in ticks that is not present in mice. We further demonstrate that RNAi is more likely to promote diversification of synonymous mutations rather than nonsynonymous. Resistance to nonsynonymous diversification, in all likelihood, is mediated by selective constraint against low fitness (or lethal) amino acid substitutions. Interestingly, despite an observed positive correlation between increased RNAi targeting and increased mutational diversity at synonymous sites, this relationship was not proportional. The degree of RNAi targeting far exceeded that of complexity at synonymous sites. This likely reflects a conflict between intra-tick RNAi-mediated diversification and selective constraint. Synonymous changes conserve the amino acid sequence, but can change the codon, and in the correct circumstances, can disrupt translation and protein function (Kimchi-Sarfaty et al. 2007). While we did not observe a codon usage shift towards the preference in ticks, purifying selection may be acting to remove some synonymous mutations through constraints in codon usage (Haas et al. 1996; Aragones et al. 2008). Furthermore, synonymous mutations could alter the genomic RNA tertiary structure and consequently impact the fitness of the mutated genome thereby imposing a selective constraint on the accumulation of synonymous changes (Pirakitikulr et al. 2016). Resistance to genetic diversification is most clearly demonstrated when comparing the structural and nonstructural protein coding regions. Significantly more vsRNAs were produced from the structural regions than from the nonstructural; yet, the differences in genetic complexity were not proportional. It appears that the structural protein region of the genome is less likely to create RNAi-escape mutants due to other fitness trade-offs. For the tick, this may be an advantage. One hypothesis is that RNAi targeting of this conserved region may be more effective at controlling viral replication and virus-induced pathogenesis. Regardless, further examination on the role of the tick RNAi pathway on POWV diversification is needed.

4.3 Comparison to mosquito-borne viruses

Many different forces act to influence the evolutionary outcomes of arboviruses, including bottlenecks (genetic drift) (Ciota et al. 2012; Forrester et al. 2012; Gutierrez et al. 2015; Grubaugh et al. 2016; Lequime et al. 2016), control of viral replication (purifying selection) (Grubaugh et al. 2015), viral adaptation (positive selection) (Brault et al. 2007; Moudy et al. 2007; Tsetsarkin et al. 2007, 2014; Stapleford et al. 2014), and the arthropod RNAi response (diversifying selection) (Brackney et al. 2009, 2015). In the WNV system, evolution acts in a way to constrain genetic diversity in birds (Jerzak et al. 2008; Grubaugh et al. 2015) and promote diversity in mosquitoes (Jerzak et al. 2005, 2007; Grubaugh et al. 2016). These patterns are strikingly different in the POWV system where diversity is greater in mice than in ticks. The reason for the differences between the two systems is unknown. One factor contributing to this discrepancy could be that POWV populations were analyzed from a single mouse tissue (i.e. sera) and multiple tissues from the tick (i.e. whole body). However, if this were a contributing factor, we would expect that the combined diversity of the POWV subpopulations derived from multiple tick tissues would be greater than the diversity associated with populations derived from a single mouse tissue. Alternatively, biological and physiological differences between the vertebrate hosts (mammals vs. birds) and the invertebrate vectors (ticks vs. mosquitoes) may present unique environmental challenges during arboviral infection. It could be that purifying selection is inherently strong within birds and ticks while relaxed in mammals and mosquitoes. This is supported by the observation that dengue virus, another mosquito-borne virus, accumulates slightly more diversity in the mammalian host (i.e. humans) than in *Aedes aegypti* mosquitoes (Lin et al. 2004; Sim et al. 2015). Experiments evaluating arbovirus demographics within non-natural hosts and vectors are needed to gain insight into the mechanistic underpinnings of these observations.

4.4 Model for tick-borne virus evolution

We propose a model for tick-borne virus evolution based on the *Ix. scapularis* life cycle (Fig. 6). Following horizontal transmission, POWV rapidly diversifies within the mammalian host. Subsequently, POWV encounters a severe bottleneck during transmission to ticks thereby facilitating genetic drift and founder's effect. Combined with the diversification in the mammal, the majority of the genetic change (~one to two mutations per genome) occurs during these few days of horizontal transmission. Once in the tick, the viral populations remain proportionally stable during transstadial transmission as a result of relaxed bottlenecks. As opposed to mosquitoes, viral purifying selection remains strong in ticks. To that effect, POWV resists much of the pressure to diversify in response to RNAi. In mosquitoes, RNAi is a significant driving force for arbovirus

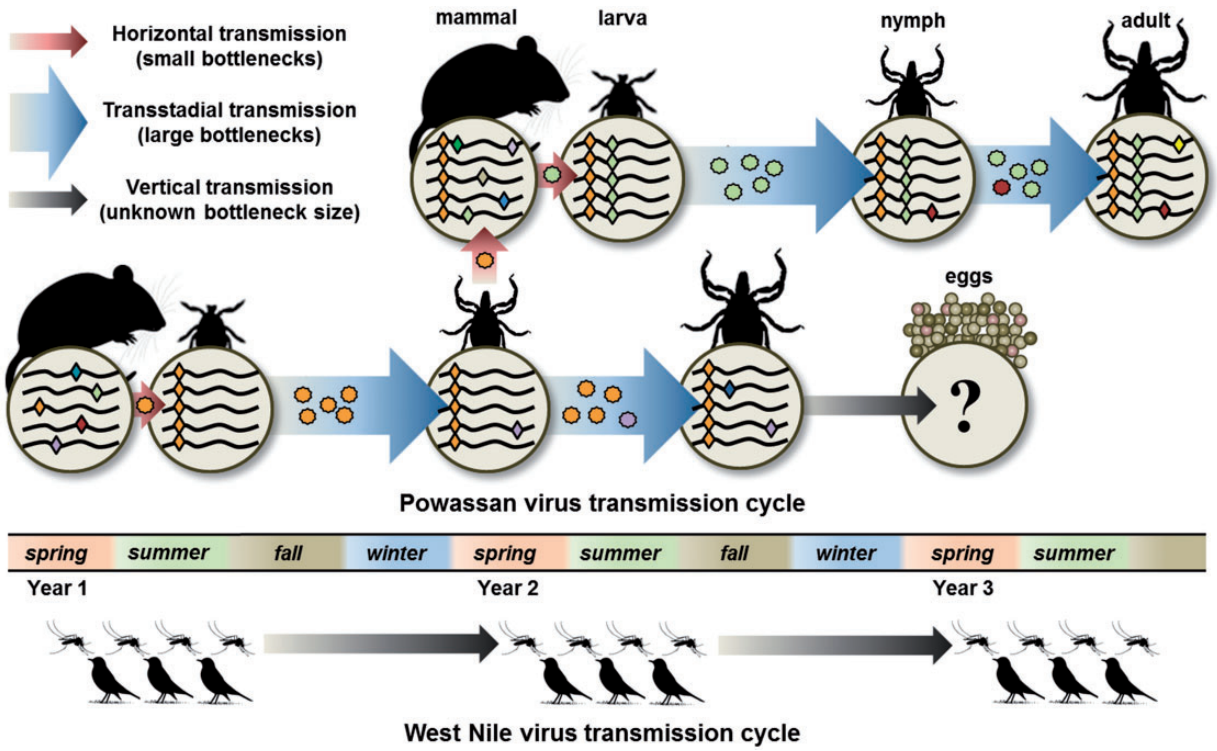

Figure 6. Model for tick-borne virus evolution based on the *Ix. scapularis* life cycle. Tick bloodfeeding is necessary for development and reproduction and occurs once per life stage (i.e. as larvae, nymphs, or adults), which is once every few months. Bloodfeeding also provides opportunities for horizontal virus transmission (red arrows), diversification, and emergence. In mice or other mammals, the virus diversifies. Only 1–89 viral genomes from the mammalian host establish primary infection in the ticks during bloodfeeding, representing a narrow bottleneck and randomly introducing one to two mutations per genome (genetic drift). In ticks, strong purifying selection constrains the effects of RNAi-mediated diversification, resulting in minimal genetic change. Combined with relaxed bottlenecks during transstadial transmission (blue arrows), the virus evolves slowly during the long periods of systemic tick infection. It is unknown how vertical transmission from adult to offspring (grey arrows) alters genetic diversity. In general, the amount of genetic diversity added to the viral populations is largely dependent upon the number of host switches. These events in ticks are rare compared to mosquito-borne viruses, such as West Nile virus which can be transmitted between mosquitoes and birds five to ten times per season. The *Ix. scapularis* life cycle is based on Ostfeld et al. (1996) and Nonaka et al. (2010).

diversification (Brackney et al. 2009, 2015), but the strong purifying selection in ticks constrains this diversity, especially from nonsynonymous mutations.

In our system, we observed that the majority of POWV genetic divergence occurred during horizontal transmission. However, its slow rate of evolution in nature indicates that horizontal transmission is infrequent (Brackney et al. 2008; Brackney et al. 2010; Pesko et al. 2010). We previously predicted that the rates of POWV evolution were $\sim 4 \times 10^{-5}$ to 2×10^{-4} substitutions per site per year (Pesko et al. 2010), or ~0.5–2 mutations per genome per year. In this study we observed ~one to two mutations during horizontal transmission. Therefore, based on that estimate, we expect that transmission between hosts occurs on average about once a year. This is consistent with the expected frequency of horizontal transmission during a typical *Ix. scapularis* life cycle where it can take between 16 and 20 months from its first bloodmeal (as a larva) to its last (as an adult) (Ostfeld et al. 1996; Nonaka et al. 2010). In addition, viral population divergence may occur through alternate routes of inter-tick transmission not directly assessed here. For example, co-feeding and vertical transmission are predicted to be critical components of maintaining POWV persistence in nature (Nonaka et al. 2010). Ticks infected by co-feeding on a non-viremic host presumably impose similar bottlenecks on viral populations as horizontal transmission. Similarly, POWV infection of the germ-line would presumably impose severe genetic bottlenecks during vertical transmission. Both routes likely factor into the rates of tick-borne virus evolution and will require further experimentation.

The slow rates of tick-borne virus evolution (Pesko et al. 2010; Uzcategui et al. 2012) are likely due to a combination of genetic constraint in ticks and few opportunities for interhost transmission. Mosquito-borne viruses evolve two- to fivefold faster (Jenkins et al. 2002; Snapinn et al. 2007), certainly influenced by the genetic diversity produced within the vector (Jerzak et al. 2005, 2007; Grubaugh et al. 2016), but also by the number of horizontal transmission cycles completed per season (Fig. 6). Overall, it is clear that arbovirus ecology can directly influence evolution and continued experimental evolution studies such as these will be necessary to fully understand its contribution.

Supplementary data

Supplementary data are available at *Virus Evolution* online.

Acknowledgements

The authors would like to thank Mike Vasil (The Connecticut Agricultural Experiment Station) for maintaining the tick colonies and assistance with the mouse infections and transmission studies. The authors thank T. Eike and E. Petrilli (Colorado State University) for technical

support and J. Fauver, J. Weger-Lucarelli, and M. Stenglein (Colorado State University) for their useful discussions.

Conflict of interest: None declared.

References

Ali, A., et al. (2006) 'Analysis of genetic bottlenecks during horizontal transmission of Cucumber mosaic virus', *Journal of Virology*, 80/17: 8345–50.

Alizon, S., and Fraser, C. (2013) 'Within-host and between-host evolutionary rates across the HIV-1 genome', *Retrovirology*, 10: 49.

Aragones, L., Bosch, A., and Pinto, R. M. (2008) 'Hepatitis A virus mutant spectra under the selective pressure of monoclonal antibodies: codon usage constraints limit capsid variability', *Journal of Virology*, 82/4: 1688–700.

Balashov Iu, S., Grigor'eva, L. A., and Oliver, J. (1998) Interspecies cross of ixodes ticks from the group Ixodes ricinus-persulcatus. *Doklady Akademii nauk/[Rossiiskaia akademii nauk]*361/5: 712–4.

Barry, G., et al. (2013) 'Gene silencing in tick cell lines using small interfering or long double-stranded RNA', *Experimental & Applied Acarology*, 59/3: 319–38.

Bedford, N. L., and Hoekstra, H. E. (2015) Peromyscus mice as a model for studying natural variation. *eLife*;4: e06813.

Blitvich, B. J., and Firth, A. E. (2015) 'Insect-specific flaviviruses: a systematic review of their discovery, host range, mode of transmission, superinfection exclusion potential and genomic organization', *Viruses*, 7/4: 1927–59.

Brackney, D. E., Beane, J. E., and Ebel, G. D. (2009) 'RNAi targeting of West Nile Virus in mosquito midguts promotes virus diversification', *PLoS Pathogen*, 5/7: e1000502.

——, ——, and ——, et al. (2008) 'Stable prevalence of Powassan virus in *Ixodes scapularis* in a northern Wisconsin focus', *The American Journal of Tropical Medicine and Hygiene*, 79/6: 971–3.

——, ——, and ——, et al. (2010) 'Homogeneity of Powassan virus populations in naturally infected *Ixodes scapularis*', *Virology*, 402/2: 366–71.

——, ——, and ——, et al. (2015) 'Modulation of flavivirus population diversity by RNA interference', *Journal of Virology*, 89/7: 4035–9.

Brault, A. C., et al. (2007) 'A single positively selected West Nile viral mutation confers increased virogenesis in American crows', *Nature Genetics*, 39/9: 1162–6.

Chao, L. (1990) 'Fitness of RNA virus decreased by Muller's ratchet', *Nature*, 348/6300: 454–5.

Ciota, A. T., et al. (2012) 'Quantification of intrahost bottlenecks of West Nile virus in *Culex pipiens* mosquitoes using an artificial mutant swarm', *Infection, Genetics and Evolution* 12/3: 557–64.

Costero, A., and Grayson, M. A. (1996) 'Experimental transmission of Powassan virus (Flaviviridae) by *Ixodes scapularis* ticks (Acari:Ixodidae)', *The American Journal of Tropical Medicine and Hygiene*, 55/5: 536–46.

de la Fuente, J., et al. (2007) 'RNA interference for the study and genetic manipulation of ticks', *Trends in Parasitological*, 23/9: 427–33.

Dohm, J. C., et al. (2008) 'Substantial biases in ultra-short read data sets from high-throughput DNA sequencing', *Nucleic Acids Research*, 36/16: e105.

Duarte, E., et al. (1992) 'Rapid fitness losses in mammalian RNA virus clones due to Muller's ratchet', *Proceedings of the National Academy of Sciences of the United States of America*, 89/13: 6015–9.

Ebel, G. D. (2010) 'Update on Powassan virus: emergence of a North American tick-borne flavivirus', *Annual Review of Entomology*, 55: 95–110.

——, and Kramer, L. D. (2004) 'Short report: duration of tick attachment required for transmission of Powassan virus by deer ticks', *The American Journal of Tropical Medicine and Hygiene*, 71/3: 268–71.

——, and ——, et al. (2000) 'Enzootic transmission of deer tick virus in New England and Wisconsin sites', *The American Journal of Tropical Medicine and Hygiene*, 63/1–2: 36–42.

Evans, G. O. (1992) *Principles of Acarology*. Cambridge, UK: University Press.

Fauci, A. S., and Morens, D. M. (2016) 'Zika virus in the Americas—yet another arbovirus threat', *The New England Journal of Medicine*, 374/7: 601–4.

Forrester, N. L., et al. (2012) 'Vector-borne transmission imposes a severe bottleneck on an RNA virus population', *PLoS Pathogens*, 8/9: e1002897.

Fumagalli, M., et al. (2013) 'Quantifying population genetic differentiation from next-generation sequencing data', *Genetics*, 195/3: 979–92.

Fusaro, A., et al. (2011) 'Phylogeography and evolutionary history of reassortant H9N2 viruses with potential human health implications', *Journal of Virology*, 85/16: 8413–21.

Garcia, S., et al. (2005) 'Nairovirus RNA sequences expressed by a Semliki Forest virus replicon induce RNA interference in tick cells', *Journal of Virology*, 79/14: 8942–7.

——, et al. (2006) 'Viral suppressors of RNA interference impair RNA silencing induced by a Semliki Forest virus replicon in tick cells', *Journal of General Virology*, 87/7: 1985–9.

Grabherr, M. G., et al. (2011) 'Full-length transcriptome assembly from RNA-Seq data without a reference genome', *Nature Biotechnology*, 29/7: 644–52.

Grubaugh, N. D., et al. (2015) 'Experimental evolution of an RNA virus in wild birds: evidence for host-dependent impacts on population structure and competitive fitness', *PLoS Pathogen*, 11/5: e1004874.

——, et al. (2016) 'Genetic drift during systemic arbovirus infection of mosquito vectors leads to decreased relative fitness during host switching', *Cell Host & Microbe*, 19/4: 481–92.

Gutierrez, S., et al. (2015) 'Demographics of natural oral infection of mosquitos by venezuelan equine encephalitis virus', *Journal of Virology*, 89/7: 4020–2.

Haas, J., Park, E. C., and Seed, B. (1996) 'Codon usage limitation in the expression of HIV-1 envelope glycoprotein', *Current Biology*, 6/3: 315–24.

Heinze, D. M., Gould, E. A., and Forrester, N. L. (2012) 'Revisiting the clinal concept of evolution and dispersal for the tick-borne flaviviruses by using phylogenetic and biogeographic analyses', *Journal of Virology*, 86/16: 8663–71.

Holmes, E. C. (2009) 'The evolutionary genetics of emerging viruses', *Annual Review of Ecology, Evolution, and Systematics*, 40: 353–72.

Jenkins, G. M., et al. (2002) 'Rates of molecular evolution in RNA viruses: a quantitative phylogenetic analysis', *Journal of Molecular Evolution*, 54/2: 156–65.

Jerzak, G., et al. (2005) 'Genetic variation in West Nile virus from naturally infected mosquitoes and birds suggests quasispecies structure and strong purifying selection', *Journal of General Virology*, 86: 2175–83.

——, et al. (2007) 'The West Nile virus mutant spectrum is host-dependant and a determinant of mortality in mice', *Virology*, 360/2: 469–76.

——, et al. (2008) 'Genetic diversity and purifying selection in West Nile virus populations are maintained during host switching', *Virology*, 374/2: 256–60.

Jones, L. D., et al. (1997) 'Transmission of louping ill virus between infected and uninfected ticks co-feeding on mountain hares', *Medical and Veterinary Entomology*, 11/2: 172–6.

Jukes, T. H., and Cantor, C. R. (1969) 'Evolution of protein molecules', in H. N. Munro(ed.) *Mammalian Protein Metabolism*, pp 71–123. New York: Academic Press.

Kimchi-Sarfaty, C., et al. (2007) 'A "silent" polymorphism in the MDR1 gene changes substrate specificity', *Science*, 315/5811: 525–8.

Kovalev, S. Y., and Mukhacheva, T. A. (2014) Tick-borne encephalitis virus subtypes emerged through rapid vector switches rather than gradual evolution. *Ecology and Evolution*4/22: 4307–16.

Kramer, L. D., Styer, L. M., and Ebel, G. D. (2008) 'A global perspective on the epidemiology of West Nile virus', *Annual Review of Entomology*, 53: 61–81.

Kryazhimskiy, S., and Plotkin, J. B. (2008) 'The population genetics of dN/dS', *PLoS Genetics*, 4/12: e1000304.

Kunec, D., and Osterrieder, N. (2016) 'Codon pair bias is a direct consequence of dinucleotide bias', *Cell Reports*, 14/1: 55–67.

Langmead, B., et al. (2009) 'Ultrafast and memory-efficient alignment of short DNA sequences to the human genome', *Genome Biology*, 10/3: R25.

Lee, W.-P., et al. (2014) 'MOSAIK: a hash-based algorithm for accurate next-generation sequencing short-read mapping', *PLoS One*, 9/3: e90581.

Lequime, S., et al. (2016) 'Genetic drift, purifying selection and vector genotype shape dengue virus intra-host genetic diversity in mosquitoes', *PLoS Genetics*, 12/6: e1006111.

Li, H., et al. (2009) 'The sequence alignment/map format and SAMtools', *Bioinformatics*, 25/16: 2078–9.

Librado, P., and Rozas, J. (2009) 'DnaSP v5: a software for comprehensive analysis of DNA polymorphism data', *Bioinformatics*, 25/11: 1451–2.

Lin, S. R., et al. (2004) 'Study of sequence variation of dengue type 3 virus in naturally infected mosquitoes and human hosts: implications for transmission and evolution', *Journal of Virology*, 78/22: 12717–21.

Liu, Q., et al. (2014) 'Severe fever with thrombocytopenia syndrome, an emerging tick-borne zoonosis', *The Lancet Infectious Diseases*, 14/8: 763–72.

Lobo, F. P., et al. (2009) 'Virus-host coevolution: common patterns of nucleotide motif usage in Flaviviridae and their hosts', *PLoS One*, 4/7: e6282.

Macalalad, A. R., et al. (2012) 'Highly sensitive and specific detection of rare variants in mixed viral populations from massively parallel sequence data', *PLoS Computational Biology*, 8/3: e1002417.

Mansfield, K. L., et al. (2009) 'Tick-borne encephalitis virus—a review of an emerging zoonosis', *Journal of General Virology*, 90/Pt 8: 1781–94.

McMullan, L. K., et al. (2012) 'A new phlebovirus associated with severe febrile illness in Missouri', *The New England Journal of Medicine*, 367/9: 834–41.

Monsion, B., et al. (2008) 'Large bottleneck size in cauliflower mosaic virus populations during host plant colonization', *PLoS Pathogen*, 4/10: 7.

Moudy, R. M., et al. (2007) 'A newly emergent genotype of West Nile virus is transmitted earlier and more efficiently by Culex mosquitoes', *The American Journal of Tropical Medicine and Hygiene*, 77/2: 365–70.

Nei, M., and Gojobori, T. (1986) 'Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions', *Molecular Biology and Evolution*, 3/5: 418–26.

Nonaka, E., Ebel, G. D., and Wearing, H. J. (2010) 'Persistence of pathogens with short infectious periods in seasonal tick populations: the relative importance of three transmission routes', *PLoS One*, 5/7: e11745.

Ostfeld, R. S., Hazler, K. R., and Cepeda, O. M. (1996) 'Temporal and spatial dynamics of *Ixodes scapularis* (Acari: Ixodidae) in a rural landscape', *Journal of Medical Entomology*, 33/1: 90–5.

Pesko, K. N., et al. (2010) 'Molecular epidemiology of Powassan virus in North America', *Journal of General Virology*, 91/Pt 11: 2698–705.

Pirakitikulr, N., et al. (2016) 'The coding region of the HCV genome contains a network of regulatory RNA structures', *Molecular Cell*, 62/1: 111–20.

Randolph, S. E. (2011) 'Transmission of tick-borne pathogens between co-feeding ticks: Milan Labuda's enduring paradigm', *Ticks and Tick-Borne Diseases*, 2/4: 179–82.

Schnettler, E., et al. (2013) 'RNA interference targets arbovirus replication in Culicoides cells', *Journal of Virology*, 87/5: 2441–54.

———, et al. (2014) 'Induction and suppression of tick cell antiviral RNAi responses by tick-borne flaviviruses', *Nucleic Acids Research*, 42/14: 9436–46.

Sim, S., et al. (2015) 'Tracking dengue virus intra-host genetic diversity during human-to-mosquito transmission', *PLoS Neglected Tropical Diseases*, 9/9: e0004052.

Simmen, M. W. (2008) 'Genome-scale relationships between cytosine methylation and dinucleotide abundances in animals', *Genomics*, 92/1: 33–40.

Simpson, G. G. (1944) *Tempo and Mode in Evolution*. New York: Columbia University Press, vol. xviii, 237 pp.

Snapinn, K. W., et al. (2007) 'Declining growth rate of West Nile virus in North America', *Journal of Virology*, 81/5: 2531–4.

Stapleford, K. A., et al. (2014) 'Emergence and transmission of arbovirus evolutionary intermediates with epidemic potential', *Cell Host & Microbe*, 15/6: 706–16.

Telford, S. R., et al. (1997) 'A new tick-borne encephalitis-like virus infecting New England deer ticks, *Ixodes dammini*', *Emerging Infectious Diseases*, 3/2: 165–70.

Tsetsarkin, K. A., et al. (2007) 'A single mutation in chikungunya virus affects vector specificity and epidemic potential', *PLoS Pathogen*, 3/12: 1895–906.

———, et al. (2014) 'Multi-peaked adaptive landscape for chikungunya virus evolution predicts continued fitness optimization in *Aedes albopictus* mosquitoes', *Nature Communication*, 5: 4084.

Uzcategui, N. Y., et al. (2012) 'Rate of evolution and molecular epidemiology of tick-borne encephalitis virus in Europe, including two isolations from the same focus 44 years apart', *Journal of General Virology*, 93/Pt 4: 786–96.

Varble, A., et al. (2014) 'Influenza A virus transmission bottlenecks are defined by infection route and recipient host', *Cell Host & Microbe*, 16/5: 691–700.

Vodovar, N., et al. (2012) 'Arbovirus-derived piRNAs exhibit a ping-pong signature in mosquito cells', *PloS One*, 7/1: e30861.

Wang, G. P., et al. (2010) 'Hepatitis C virus transmission bottlenecks analyzed by deep sequencing', *Journal of Virology*, 84/12: 6218–28.

Weaver, S. C., and Lecuit, M. (2015) 'Chikungunya virus and the global spread of a mosquito-borne disease', *The New England Journal of Medicine*, 372/13: 1231–9.

Woelk, C. H., and Holmes, E. C. (2002) 'Reduced positive selection in vector-borne RNA viruses', *Molecular Biology and Evolution*, 19/12: 2333–6.

Yang, X., et al. (2013) 'V-Phaser 2: variant inference for viral populations', *BMC Genomics*, 14: 674.