Journal of General Virology (2008), **89**, 2122–2131 DOI 10.1099/vir.0.2008/001719-0

Phylogeographical structure and evolutionary history of two Buggy Creek virus lineages in the western Great Plains of North America

Abinash Padhi,[1]† Amy T. Moore,[1] Mary Bomberger Brown,[1]‡ Jerome E. Foster,[1]§ Martin Pfeffer,[2] Kathryn P. Gaines,[1] Valerie A. O'Brien,[1] Stephanie A. Strickler,[1]¶ Allison E. Johnson[3] and Charles R. Brown[1]

Correspondence
Charles R. Brown
charles-brown@utulsa.edu

[1]Department of Biological Sciences, University of Tulsa, Tulsa, OK 74104, USA

[2]Bundeswehr Institute of Microbiology, Neuherbergstrasse 11, 80937 Munich, Germany

[3]Department of Biology, St Olaf College, St Olaf, MN 55057, USA

Buggy Creek virus (BCRV) is an unusual arbovirus within the western equine encephalitis complex of alphaviruses. Associated with cimicid swallow bugs (*Oeciacus vicarius*) as its vector and the cliff swallow (*Petrochelidon pyrrhonota*) and house sparrow (*Passer domesticus*) as its amplifying hosts, this virus is found primarily in the western Great Plains of North America at spatially discrete swallow nesting colonies. For 342 isolates collected in Oklahoma, Nebraska, Colorado and North Dakota, from 1974 to 2007, we sequenced a 2076 bp region of the 26S subgenomic RNA structural glycoprotein coding region, and analysed phylogenetic relationships, rates of evolution, demographical histories and temporal genetic structure of the two BCRV lineages found in the Great Plains. The two lineages showed distinct phylogeographical structure: one lineage was found in the southern Great Plains and the other in the northern Great Plains, and both occurred in Nebraska and Colorado. Within each lineage, there was additional latitudinal division into three distinct sublineages. One lineage is showing a long-term population decline. In comparing sequences taken from the same sites 8–30 years apart, in one case one lineage had been replaced by the other, and in the other cases there was little evidence of the same haplotypes persisting over time. The evolutionary rate of BCRV is in the order of $1.6–3.6\times10^{-4}$ substitutions per site per year, similar to that estimated for other temperate-latitude alphaviruses. The phylogeography and evolution of BCRV could be better understood once we determine the nature of the ecological differences between the lineages.

Received 28 February 2008
Accepted 13 May 2008

INTRODUCTION

Buggy Creek virus (BCRV, family *Togaviridae*) is an unusual alphavirus known mostly from the western Great Plains of North America. A single-stranded, positive-sense RNA virus that is part of the western equine encephalitis antigenic complex, BCRV is vectored primarily by the swallow bug (Hemiptera, Cimicidae, *Oeciacus vicarius*), an ectoparasite of the colonially nesting cliff swallow (Passeriformes, Hirundinidae, *Petrochelidon pyrrhonota*). The closely related Fort Morgan virus (FMV; Hayes *et al.*, 1977; Calisher *et al.*, 1980) reported from Colorado is a strain of BCRV (Pfeffer *et al.*, 2006) that is also vectored by swallow bugs. Vertebrate amplifying hosts for BCRV include cliff swallows and introduced house sparrows (Passeriformes, Passeridae, *Passer domesticus*) that occupy nests in swallow colonies (Scott *et al.*, 1984; Hopla *et al.*, 1993). The transmission cycle for this virus is not well understood. BCRV appears to be capable of overwintering in bugs at temperate latitudes as far north as southern North Dakota (Strickler, 2006; C. R. Brown, A. T. Moore & S. A. Strickler, unpublished data). This virus is not known

†Present address: Center for Infectious Disease Dynamics, Department of Biology, 208 Mueller Laboratory, The Pennsylvania State University, University Park, PA 16802, USA.

‡Present address: Tern and Plover Conservation Partnership, University of Nebraska, 3310 Holdrege St, Lincoln, NE 68583, USA.

§Present address: Department of Preclinical Sciences, Faculty of Medical Sciences, University of the West Indies, St Augustine, Trinidad.

¶Present address: Department of Zoology, University of Oklahoma, Norman, OK 73019, USA.

The GenBank/EMBL/DDBJ accession numbers of the sequences reported in this paper are DQ451557–DQ451599 and EU708019–EU708315.

to cause disease in humans, but large numbers of nestling house sparrows are infected at some sites, and mortality among those infected can be high (V. A. O'Brien and C. R. Brown, unpublished data).

Because the wingless swallow bugs are sedentary during much of the year and confined to cliff swallow colony sites, the spatial foci for BCRV occurrence are discrete and predictable (Brown *et al.*, 2001; Moore *et al.*, 2007). When bugs do move (on birds) from site to site, the extent of this movement can be estimated based on known movement probabilities of cliff swallows (Brown *et al.*, 2007, 2008). Consequently, spatial structure of BCRV population genetics can be studied at a finer scale than in most arboviruses vectored by insects and amplified by birds.

An earlier phylogenetic analysis of BCRV isolates from primarily a single year (2001) revealed the existence of two major virus lineages (Pfeffer *et al.*, 2006). Within these lineages, we found some degree of spatial structure, with isolates taken from colony sites closer together more alike phylogenetically than isolates from more distant localities (Pfeffer *et al.*, 2006). These analyses indicate that BCRV exhibits an unusually high degree of spatial structure for an alphavirus, many of which are relatively genetically homogeneous over large geographical regions (Weaver *et al.*, 1993; Cilnis *et al.*, 1996; Poidinger *et al.*, 1997; Brault *et al.*, 1999; Sammels *et al.*, 1999). The two sympatric BCRV lineages differ from each other by >6 % at the nucleotide level (Pfeffer *et al.*, 2006), which is about twice that seen between the geographically disparate lineages of the related Western equine encephalitis virus (WEEV; Nagata *et al.*, 2006).

Little is known about the geographical range or the extent of evolutionary differences between the two BCRV lineages, other than that both co-circulate in some of the same swallow colony sites in south-western Nebraska. Thus, one objective of this study was to determine the phylogeographical distribution of the two lineages in the western Great Plains and the extent to which they co-occur over broader geographical regions. We also examined rates of molecular evolution of the two lineages, using nucleotide substitution data, as a way to infer whether these lineages may be under different selective pressures potentially brought about by, for example, differences in host use or geography.

The predictable occurrence of BCRV at specific cliff swallow colony sites also allows us to examine the temporal distribution of virus haplotypes. Understanding how the distribution of genetic variants in arboviruses changes with time is important in predicting where epizootics or epidemics may occur (Holmes, 2004; Moya *et al.*, 2004), but there are few studies that have systematically sampled a given virus in the same locality with high enough virus prevalence to determine the extent of haplotype stability. If certain haplotypes are replacing others and increasing in abundance, this may mean that these variants are favoured in some environments. In this study, we used data spanning 33 years to examine temporal structure in the population genetics of BCRV and the evolutionary histories of the two lineages.

As in the earlier analysis of BCRV (Pfeffer *et al.*, 2006), here we used sequence data for a 2076 bp region of the viral subgenomic 26S RNA, including the entire E2 gene. The E2 and E3 genes in alphaviruses code for glycoproteins that are responsible for cell receptor binding (Navaratnarajah & Kuhn, 2007). This region of the genome is most sensitive to selection brought about by the immune systems of different hosts (Strauss & Strauss, 1994; Powers *et al.*, 2001; Pfeffer *et al.*, 2006), and the E2 region also determines infection of invertebrate vectors (Brault *et al.*, 2002). If there are functional differences among virus isolates that reflect varying levels of adaptation to cell receptors of different hosts or vectors, they are likely to be expressed in the E2 gene. In addition, the E2 gene has been used in other alphaviruses to infer phylogeographical patterns (Oberste *et al.*, 1998; Kramer & Fallah, 1999; Kondig *et al.*, 2007), enabling a comparison of BCRV with related viruses. Our analyses were aimed at better understanding the spatial and temporal genetic structure and processes of evolutionary change in an alphavirus that is closely related to WEEV, a widely distributed virus with medical and veterinary significance (Reisen & Monath, 1989; Weaver *et al.*, 1997).

METHODS

Study organisms. BCRV isolates collected in this study were from swallow bug vectors. The swallow bug is an ectoparasite primarily of cliff swallows and is found throughout the bird's wide geographical range (Brown & Brown, 1995). Swallow bugs are nest-based parasites that overwinter in cliff swallows' nests or in the cracks and crevices of the nesting substrate near the nests. They are haematophagous, feeding on the birds mostly at night, and they travel on the adult birds only briefly (Brown & Brown, 1996, 2004, 2005). Swallow bugs are long-lived and begin to reproduce as soon as they feed in the spring. The birds do not use all of the colony sites in a given year (Brown & Brown, 1996), and the bugs seem to be adapted to withstanding long periods of host absence (Smith & Eads, 1978; Rannala, 1995). The bugs also parasitize house sparrows that nest in some cliff swallow colonies (Hopla *et al.*, 1993; Brown *et al.*, 2001).

Cliff swallows are highly colonial passerines that breed throughout most of western North America (Brown & Brown, 1995). They build gourd-shaped mud nests and attach them to the vertical faces of cliff walls, rock outcrops or artificial sites such as the eaves of buildings or bridges. Cliff swallows are migratory, wintering in southern South America, and have a relatively short breeding season in North America from April to July.

House sparrows were introduced into North America from Europe in the late 1800s and are found in all parts of the USA (Lowther & Cink, 1992). House sparrows usurp active cliff swallow nests and will occupy them until the nests fall from the substrate. Numbers of sparrows vary among colony sites, with some colonies having none and others having only sparrows. House sparrows are non-migratory and resident around the swallow colonies throughout the year.

Study sites. Our principal study site was in south-western Nebraska, centred at the Cedar Point Biological Station (41°13′ N, 101°39′ W)

near Ogallala, in Keith County, along the North and South Platte Rivers and including portions of Deuel, Garden, Lincoln and Morrill counties. Cliff swallows have been studied there since 1982 (Brown & Brown, 1996). Approximately 170 cliff swallow colony sites are in the 200×60 km study area; about a third of these are not occupied by swallows in any given year. The study site is described by Brown & Brown (1996) and Brown *et al.* (2008).

Collections of swallow bugs for virus isolation in the primary Nebraska study area were done each year between 1998 and 2006. We made additional collections of bugs from near Dickinson, Stark County, south-western North Dakota, in June 2006 and April 2007; from near Scottsbluff, Scottsbluff County, far western Nebraska in July 2006; from Weldona and the Bijou Bridge near Fort Morgan, Morgan County, north-eastern Colorado, in June 2006 (where the Fort Morgan strain of BCRV was first isolated; Hayes *et al.*, 1977); from near Ault, Weld County, north-eastern Colorado, in July 2001 (Pfeffer *et al.*, 2006); and from the original collection site of BCRV in Grady County, west-central Oklahoma, in July 2006. We also had sequence data for reference strains of BCRV and FMV from the Centers for Disease Control's diagnostic collection, isolated from swallow bugs and sera of house sparrows in 1981–1983 in Grady County, Oklahoma, and in 1974–1977 in Morgan and Logan counties, Colorado, respectively (Pfeffer *et al.*, 2006).

Field collections of bugs. At sites with active cliff swallow nests, swallow bugs were collected from the outsides of the nests by brushing bugs off nests into a wide-mouthed collecting jar. At inactive colony sites, we removed nests to expose bugs on the substrate behind the nests and harvested additional bugs by sorting through the mud nest fragments. We sampled bugs from throughout a colony site (in parts where nests were accessible). Bugs were sorted into pools of 100 individuals while alive and frozen immediately at -70 °C. Additional details on field sampling and collecting are given in Moore *et al.* (2007) and Brown *et al.* (2007).

Virus isolation and sequencing. All isolates from 1998 to 2003 were identified by plaque assay on Vero cells and sequences were obtained on the first or second Vero passage (see Pfeffer *et al.*, 2006 for details). Beginning in 2004, positive samples were first identified by RT-PCR (Moore *et al.*, 2007) and later confirmed by plaque assay. Some RT-PCR-positive samples from 2004 to 2006 that either did not generate plaques (see Moore *et al.*, 2007), or were not tried on Vero cells, were sequenced directly from the RT-PCR product.

Viral RNA was extracted from 100 µl of the infectious pre-cleared supernatant of a Vero cell passage or from the bug-pool homogenates using the Qiagen Qiamp Mini Viral RNA extraction kit. Five microlitres of the eluted RNA suspension was used as a template in an alphavirus RT-PCR to amplify a 2076 bp fragment of the structural glycoprotein region within the subgenomic 26S RNA (Pfeffer *et al.*, 2006; Brown *et al.*, 2008).

Phylogenetic analyses. A total of 177 unique BCRV sequences, each comprising 2076 bp from the capsid (189 bp), the E3 (183 bp), the E2 (1269 bp), the 6K (165 bp) and the E1 (270 bp) regions, were found among the 342 isolates sequenced. Unique sequences were identified using Collapse version 1.2 (Posada, 2004). Maximum-likelihood (ML) and Bayesian inference (BI) phylogenies were reconstructed for (i) all three codon positions and (ii) using the first- and second-codon positions only. PAUP* 4.0b10 (Swofford, 2002) and MRBAYES version 3.1 (Huelsenbeck & Ronquist, 2001) were used for ML and BI analyses, respectively. The appropriate nucleotide substitution models for the first-, second- and third-codon and for the first- and second-codon datasets were selected by the hierarchical likelihood ratio test (hLRT) in MODELTEST version 3.5 (Posada & Crandall, 1998). The ML tree was reconstructed with the nearest neighbour index tree searching method in PAUP* 4.0b10 (Swofford, 2002). Nodal supports were estimated with 100 non-parametric bootstrap replicates using PAUP. Bayesian analyses were performed by running four simultaneous Metropolis-coupled Markov chain Monte Carlo (MCMC) simulations for 50×10^6 generations and sampling every 100 generations. All trees below the observed stationary level were discarded, resulting in a 'burn-in' of 50 000 generations. The convergence of trees was checked using Tracer version 1.4 (Rambaut & Drummond, 2007), and the resulting trees were used to generate a majority consensus tree with posterior probability values. Mean genetic distances among the sublineages of BCRV were estimated based on the Tamura–Nei model of nucleotide substitution implemented in MEGA version 4 (Tamura *et al.*, 2007). Using the same program, standard errors of the distance estimates were calculated with 1000 non-parametric bootstrap replicates. We estimated genetic distances between the two BCRV lineages and WEEV (GenBank accession nos DQ393793 and DQ393794) and Highlands J virus (HJV; GenBank accession no. AF339476) using the same Tamura–Nei model and the same 2076 bp region of the genome in WEEV and HJV.

We used a Bayesian MCMC approach implemented in BEAST version 1.4.4 (Drummond & Rambaut, 2007) that utilizes the number and temporal distribution of genetic differences among viruses sampled at different times to estimate the evolutionary change for each lineage (Drummond *et al.*, 2002, 2006). The overall substitution rate (nucleotide substitutions per site per year) of each lineage was estimated using the Bayesian skyline model, with a relaxed (variable) molecular clock (uncorrelated log-normal model) implemented in BEAST. Bayesian skyline plots with 10 grouped intervals were reconstructed to infer demographical history. The Bayesian skyline plot is a powerful new method of estimating past population dynamics that does not rely on pre-specified parametric models and works well with small sample sizes (Drummond *et al.*, 2005); this approach is being used increasingly in studies of virus population dynamics (e.g. Moratorio *et al.* 2007; Romano *et al.* 2007; Kitchen *et al.*, 2008). Phylogenies were evaluated using a chain length of 100 million (lineage A) and 150 million (lineage B) states under the HKY85 + Γ_4 substitution model, with uncertainty in the data reflected in the 95 % high-probability density (HPD) intervals.

To examine potential temporal structure of BCRV, we calculated the proportion of isolates of each unique haplotype occurring by year and superimposed these on the first-, second- and third-codon phylogeny. To control for site-related differences, we also compared sequence similarity between (i) isolates collected from Colorado at the same colony site in 1974–1975 versus 2006, (ii) isolates collected from Oklahoma at the same colony site in 1981–1983 versus 2006, and (iii) isolates collected from Nebraska at three of the same colony sites in 1998 versus 2006.

RESULTS

For the first-, second- and third-codon position dataset, the most appropriate nucleotide substitution model was the Tamura–Nei model with equal frequencies (TrNef) with the proportion of invariable sites (I)=0.5084, and the gamma distribution shape parameter (G)=0.8526. For the first- and second-codon position dataset, the Kimura two-parameter (K80) model with I=0.7024 and G=0.7466 was the best fit. The first, second and third codon-based ML tree revealed the existence of two distinct lineages of BCRV (A and B) with three distinct sublineages for each (A1–A3 and B1–B3) in the western Great Plains. Isolates belonging to both lineages A and B were circulating within Nebraska

(*n*=280 isolates) and Colorado (*n*=22); samples from North Dakota (*n*=19) contained only lineage B; and in Oklahoma (*n*=21) only lineage A was found (Fig. 1). Consistent with the first-, second- and third-codon tree, analyses based on only the first- and second-codon positions revealed the same lineages and sublineages within each (Fig. 2). However, many of the unique sequences defined in the first-, second- and third-codon tree were collapsed within the first- and second-codon tree, indicating that the majority of the variation among the isolates within each sublineage could be attributed to the third-codon position. Both analyses (Figs 1 and 2) showed

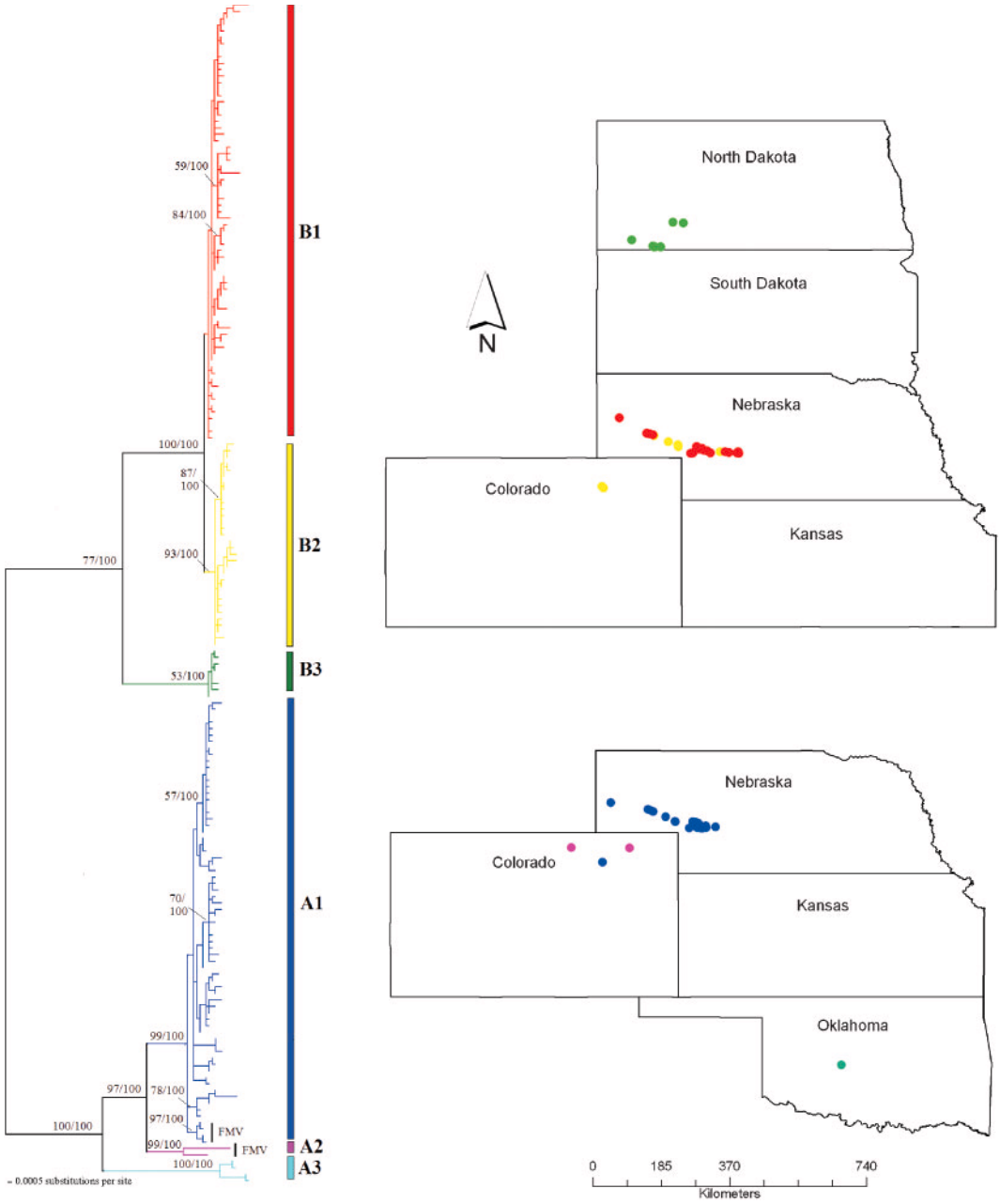

Fig. 1. ML tree based on the first-, second- and third-codon positions of a 2076 bp structural glycoprotein coding region showing phylogenetic relationships among 177 unique isolates of BCRV from 342 isolates. The phylogeny revealed two distinct lineages (A and B) and sublineages (A1–A3 and B1–B3) associated with their geographical distribution within the Great Plains. Coloured dots indicate collecting sites (often multiple isolates were found at a site) and are keyed to their respective branches in the phylogeny. The phylogenetic position of FMV isolates is also shown. Nodal support (ML/BI) is given at the base of each node. The midpoint rooting of the tree is for purposes of clarity.

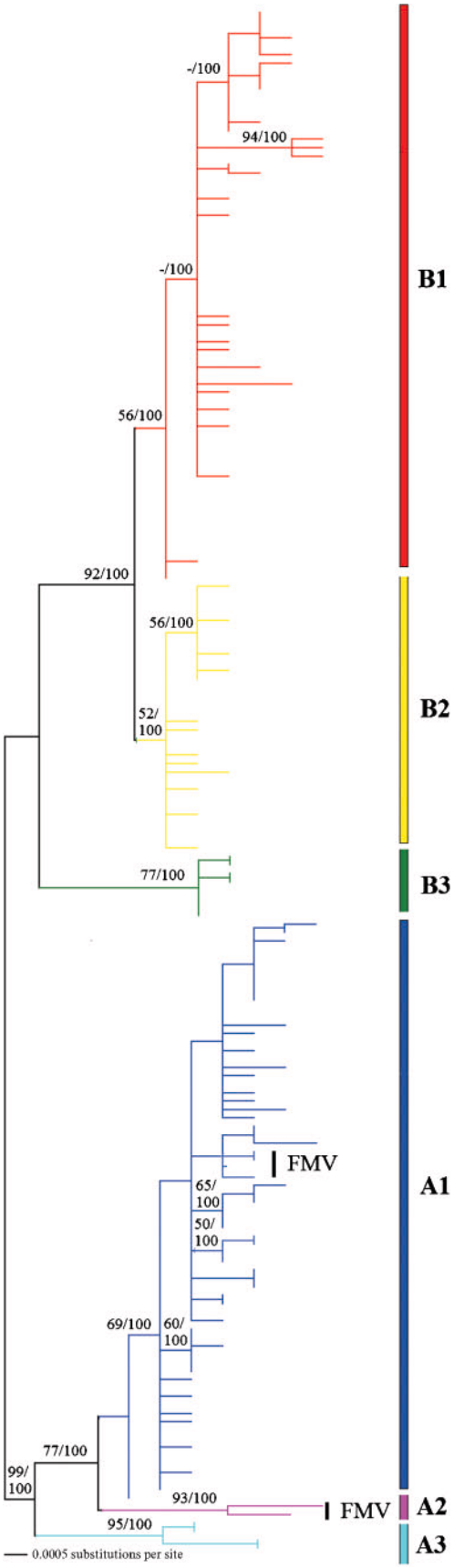

Fig. 2. ML tree based on the first- and second-codon positions only (1384 bp), showing phylogenetic relationships among the 177 unique BCRV isolates of Fig. 1. Many of the unique isolates from Fig. 1 collapsed here, suggesting that within each lineage, variation among the isolates is due mostly to rapid evolution of the third-codon position. The phylogenetic position of FMV isolates is also shown. Nodal support (ML/BI) is given at the base of each node. The midpoint rooting of the tree is for purposes of clarity.

that the Fort Morgan strain of BCRV belonged to two sublineages (A1 and A2).

Estimated genetic distances (in percentage of nucleotides) among the different sublineages (Table 1) reflected the sublineages' geographical proximity. For example, the genetic distance between sublineages A1 and A2 was smaller than that between A1 and A3 or between A2 and A3 (Table 1). The genetic distance between sublineages B1 and B2 was smaller than that between B1 and B3 or between B2 and B3 (Table 1). The genetic distances between BCRV lineage A and WEEV and between BCRV lineage A and HJV were 45.1 ± 1.9 and 46.6 ± 1.82, respectively. BCRV lineage B showed 46.01 ± 1.9 and 44.9 ± 1.82 genetic distances from WEEV and HJV, respectively.

Based on the relaxed clock assumption, the evolutionary rate for BCRV lineage A was 2.641×10^{-4} (95 % HPD 1.664×10^{-4}–3.642×10^{-4}) substitutions per site per year and for lineage B was 1.807×10^{-4} (95 % HPD 1.666×10^{-4}–2.480×10^{-4}) substitutions per site per year. The times to the most recent common ancestor (TMRCA) for the sublineages of A were 80.334 (95 % HPD 40.809–137.653) years and for the sublineages of B was 59.322 (95 % HPD 14.778–124.444) years. Lineage A has undergone an overall population decline in the western Great Plains (Fig. 3a), whereas lineage B showed an increase that began about 15 years ago, followed by a relatively sharp

Table 1. Percentage genetic distances, based on the Tamura–Nei nucleotide substitution model of sequence evolution, among different sublineages within each lineage of BCRV.

Distance estimates and standard errors are listed below and above the diagonal, respectively.

Sublineage	A1	A2	A3
A1	–	0.3	0.4
A2	2.0	–	0.4
A3	3.5	3.6	–
Sublineage	**B1**	**B2**	**B3**
B1	–	0.1	0.4
B2	0.6	–	0.4
B3	2.9	2.9	–

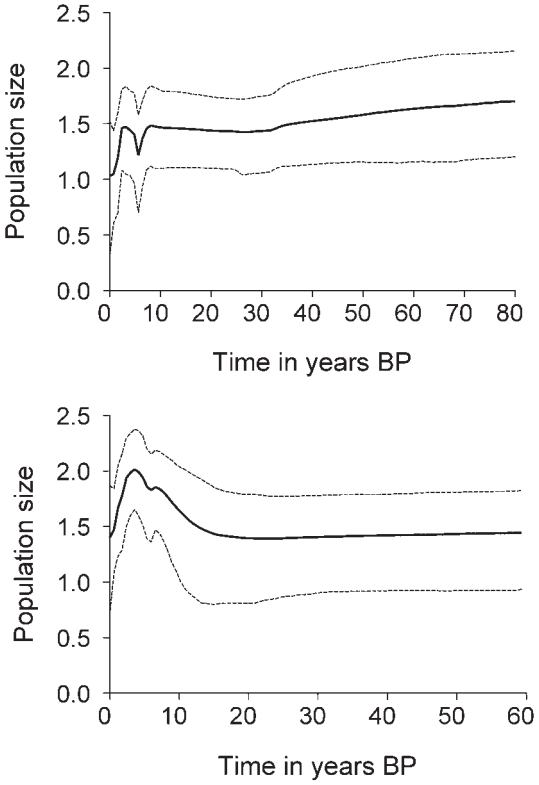

Fig. 3. Bayesian skyline plots depicting population size in relation to time in years (before present, BP) for BCRV lineages A (upper) and B (lower). Population size is the mean estimate (Ne×g; product of the effective population size and the generation length in years), expressed on a logarithmic scale (*y* axis). The dotted lines give the 95 % HPD interval of the estimates. Population size was estimated over the time since the sublineages of A and B diverged from their respective common ancestors.

return recently to population levels characteristic of those before the increase (Fig. 3b).

Within lineages A1 and B1 especially, which came largely from sites in the south-western Nebraska study area, there was little evidence for temporal clustering of multiple isolates of the same haplotype (Fig. 4). However, detection of temporal clustering can be affected by unequal geographical sampling between years, especially if there is spatial structure. Among sites that were represented both early and later in the study, the Colorado samples (n=6) from 1974 to 1977 ('FMV') belonged exclusively to lineage A, whereas isolates from the same colony site in 2006 were all lineage B (n=15). Oklahoma samples all taken from the same bridge in 1981–1983 and again in 2006, though of the same lineage, had diverged to the point that no haplotypes from the 1980s were found 25 years later at the same site (see lineage A3 in Fig. 4). For three Nebraskan colony sites sampled 8 years apart, only 14.3 % of the haplotypes detected at those sites (n=21) were shared between 1998 and 2006 (Table 2).

DISCUSSION

The analyses reported here suggest that the two lineages of BCRV in the western Great Plains have latitudinally distinct geographical ranges that overlap in some areas and somewhat different demographical histories. Within each lineage, there is clear spatial structure both regionally (this study) and at a more local level (Pfeffer *et al.*, 2006; Brown *et al.*, 2008). There is some evidence that over the past 25–30 years, the predominant haplotypes at certain sites have changed, and in one case one lineage apparently replaced the other. Consistent with other analyses (Powers *et al.*, 2001; Pfeffer *et al.*, 2006), we found strong support for Fort Morgan virus being a strain of BCRV confined to north-eastern Colorado.

The genetic distances between the two BCRV lineages and their most closely related congeners (HJV and WEEV) suggest that the two main BCRV lineages are of roughly similar age and do not permit conclusions as to whether one is ancestral to the other. Despite the monophyly of each BCRV sublineage (Fig. 1), the genetic distances among different sublineages were relatively small; those among the sublineages of B were smaller than those among the sublineages of A (Table 1). This finding suggests that the sublineages have only diverged recently from a common ancestor, a conclusion also supported by estimated TMRCA values of 80 and 59 years for the A and B sublineages, respectively. That the current sublineages evolved relatively recently is perhaps not surprising given the fast rates of evolution in RNA viruses (Moya *et al.*, 2004). The phylogenetic analyses based on the first- and second-codon positions confirm that the sublineages are distinct and suggest that the divergence of the sublineages is probably due mostly to changes in the first- and second-codon regions. The genetic difference between the sympatric sublineages A1 and A2 was equivalent to that seen among geographically separated lineages of the related WEEV in California using the same E2 region of the genome (Kramer & Fallah, 1999), suggesting that BCRV sublineages, even within the same geographical area, are relatively well differentiated genetically.

BCRV shows strong phylogeographical structure within the western Great Plains, with a more southerly lineage (A) and a more northerly lineage (B) both showing three sublineages that are themselves somewhat latitudinally distinct. Lineages A and B co-circulate within Colorado and Nebraska, at times within swallow bugs from the same cliff swallow colony site. The available data suggest that north-eastern Colorado and south-western Nebraska are near the northerly and southerly limits of the ranges of lineages A and B, respectively, although this conclusion is tentative until additional samples are taken from other localities within the Great Plains. Little work has been done on BCRV at other locations. We searched for BCRV in Texas in 2006, finding none by RT-PCR in 48 100-bug pools analysed from south central Texas and 39 from east Texas. Recently, BCRV of an undetermined lineage was

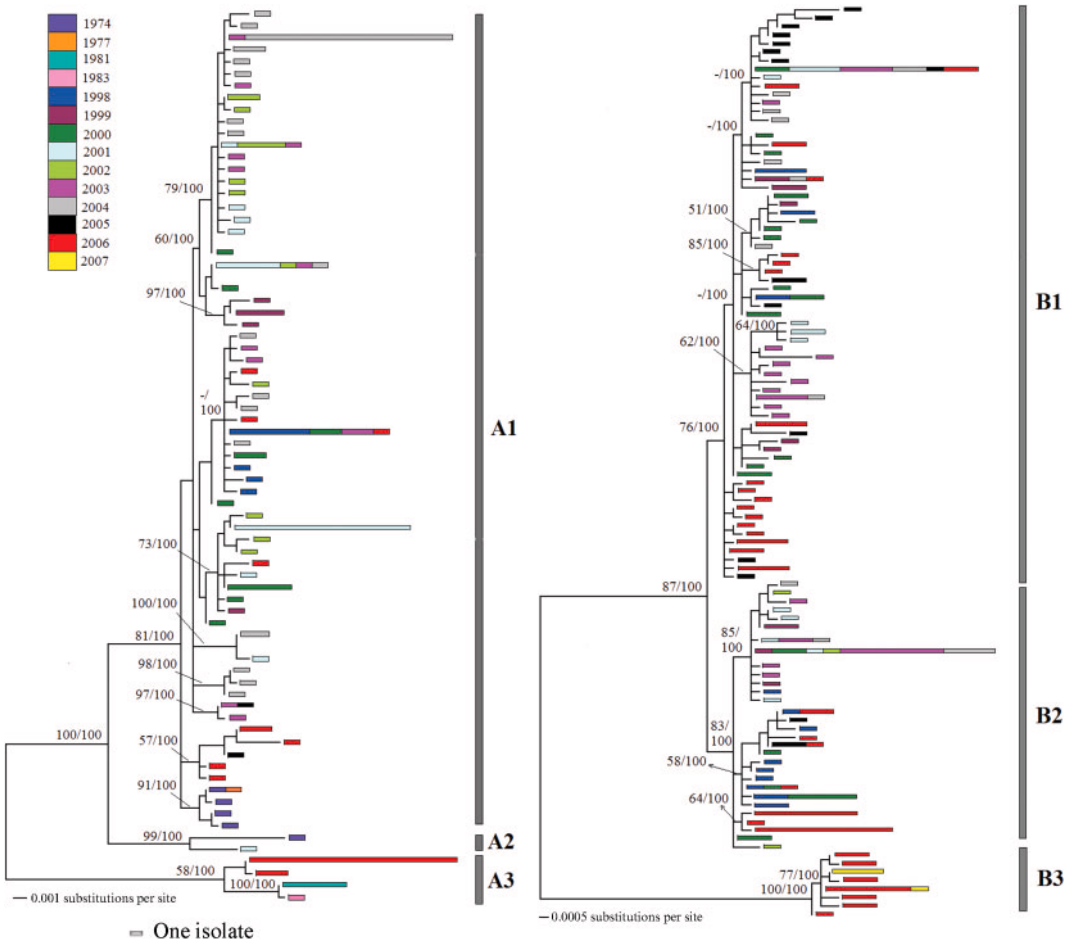

Fig. 4. ML tree of BCRV lineages A (left) and B (right) showing the number of isolates from different years for each unique sequence.

isolated from swallow bugs in California (W. K. Reisen & A. C. Brault, personal communication). The virus is apparently associated closely with cliff swallows and their ectoparasitic bugs, and it should be searched for throughout the bird's wide North American breeding range. In particular, the lineage distributions would be clarified with BCRV samples from Kansas and South Dakota because, based on the phylogeographical pattern (Fig. 1), those regions might contain any of the lineages/sublineages except perhaps B3 and A3, respectively.

Our results for BCRV generally match those found for other alphaviruses, many of which exhibit multiple lineages. There are both geographically distinct lineages of WEEV in North America and ones that apparently co-circulate in certain areas (e.g. California), although much

Table 2. Extent of shared haplotypes among BCRV isolates at three colony sites in south-western Nebraska sampled in 1998 and again in 2006.

Colony site	Lineage	1998		2006		Shared between years
		Total isolates	Unique to that year	Total isolates	Unique to that year	
02	B1	6	3	4	2	0
53	B2	10	7	5	2	2
88	A1	7	2	3	2	1

of what we know about the geographical distribution of the WEEV lineages is based on single or a few isolates taken in some cases many years apart (Weaver *et al.*, 1997; Kramer & Fallah, 1999). Sindbis virus (SINV) and Ross River virus (RRV) both have multiple lineages, with RRV lineages showing (as in BCRV) partial geographical separation (Lindsay *et al.*, 1993) and SINV lineages exhibiting both temporal and geographical separation (Norder *et al.*, 1996; Sammels *et al.*, 1999). Chikungunya virus in Africa has multiple lineages that are largely separate geographically (Powers *et al.*, 2000). In HJV, some evidence indicated multiple lineages within North America, but conclusions were hampered by small sample sizes (Cilnis *et al.*, 1996). In contrast, Eastern equine encephalitis virus in North America and Barmah Forest virus in Australia each exist as a single, genetically conserved lineage with relatively little geographical and temporal structure (Poidinger *et al.*, 1997; Brault *et al.*, 1999).

In general, the selective pressures causing the evolution and differentiation of alphavirus strains are not well known despite considerable speculation as to various ecological factors that might be responsible (Weaver *et al.*, 1992, 1994). In the case of BCRV, the latitudinal differences in the distribution of the two lineages may reflect adaptation to different climatic regimes. Climate is the most obvious ecological variable that differs among these geographical regions. Although we would need information on how the lineages differ in their non-structural genes to address potential differences in temperature tolerances, BCRV in general appears well suited to overwintering at relatively high temperate latitudes. It has been isolated routinely from bugs in abandoned swallow nests in mid winter in Nebraska when temperatures can be below freezing for extended periods (Strickler, 2006). Perhaps lineage B is better suited to surviving prolonged winter cold in northerly localities such as North Dakota. All isolates in this study came from swallow bugs, so the phylogeographical structure is apparently not related to geographical differences in vector type or abundance. Similarly, the presumed vertebrate hosts for BCRV (cliff swallows and house sparrows) are found throughout the Great Plains, with no latitudinal differences in host availability known for either of these species.

We did not find statistical evidence for a difference in the evolutionary rate of the two lineages, as measured by nucleotide substitutions per site per year. However, the upper bound (of the 95 % HPD) for lineage A was considerably higher than for lineage B. If lineage B has a slower evolutionary rate, this could be because the relatively harsh climatic regime of the more northerly locations (e.g. North Dakota) constrains its evolution. In other alphaviruses, tropical lineages have higher evolutionary rates than their more temperate lineages (Weaver *et al.* 1994; Cilnis *et al.* 1996), perhaps because lower ambient temperatures in temperate areas reduce activity of the viruses' poikilothermic vectors and thereby shorten the transmission season and number of replication cycles (Weaver *et al.*, 1997).

Collectively the evolutionary rates measured for each BCRV lineage were broadly similar to those reported for other RNA arboviruses and for alphaviruses in particular, in which evolutionary rates are relatively low and presumably constrained by the viruses' alternating replication cycles in invertebrate vectors and vertebrate hosts (Weaver *et al.*, 1992, 1997, 1999; Cooper & Scott, 2001; Jenkins *et al.*, 2002). Interestingly, the evolutionary rates for both BCRV lineages were about twice that of the related alphavirus, HJV, which also occurs in strictly temperate latitudes of North America but is presumably vectored only by mosquitoes (Cilnis *et al.*, 1996). This might reflect vector differences between swallow bugs and mosquitoes. Because swallow bugs undergo five instar stages and are longer-lived than mosquitoes, BCRV might have more replication cycles in its vector prior to transmission to a vertebrate host, leading to its higher evolutionary rate relative to strictly mosquito-borne alphaviruses.

Within the BCRV lineages, evolutionary rates (more akin to those of WEEV; Weaver *et al.*, 1997) are sufficiently fast that the predominant haplotypes at a site are diverging over periods of 25 years and apparently in some cases in as few as 8 years. This sort of temporal structure may reflect in part the sedentary nature of the swallow bug, the virus's principal vector, with the virus not being widely dispersed to or from a given site between years and thus not being likely to continually mix with nearby haplotypes. Virulence of this virus may have also changed with time. Hopla *et al.* (1993) reported over 250 isolates of BCRV from its type locality in west central Oklahoma in the early 1980s, all of which were identified by plaque assay (and thus all were cytopathic on vertebrate cells). In contrast, only one of the BCRV isolates we identified by RT-PCR from the same site ~25 years later (n=32) exhibited any plaque growth (C. R. Brown, A. T. Moore, V. A. O'Brien & N. Komar, unpublished data), and thus most would not have been identified as even being present using 1980s plaque-assay screening.

The gradual long-term population decline of lineage A may reflect its replacement in certain areas by lineage B. For example, we found only lineage A in north-eastern Colorado in 1974–1977, but only lineage B at the same site 30 years later. This sort of ecological replacement would be most likely in areas of range overlap, and thus we might see a similar replacement of lineage A by lineage B in Nebraska over time. Interestingly, we did not detect any FMV in our Colorado samples in 2001–2006, suggesting that this strain may have been replaced by other BCRV lineages since it was last reported in the mid 1970s. Additional sampling from Colorado is needed to confirm this conclusion.

Lineage B showed evidence of population growth, beginning about 15 years ago, at about the same time lineage A began a relatively steep population decline, although lineage B has also declined in recent years. The decline in population sizes of both lineages since about 2001 is unexplained but could be related to the unusually cold

winter of 2000–2001 (http://www.hprcc.unl.edu/nebraska/winter2000-2001.html). Harsh winter conditions that year may have reduced virus and/or vector populations to well below average. A severe drought cycle began in the Great Plains in about 2000 that persisted through to 2006 (Sauchyn *et al.*, 2003; Anonymous, 2006), which may have also affected virus or vector populations in unknown ways. Rapid extinction of lineages is known in alphaviruses: in WEEV, one of two North American lineages that diverged from a common ancestor only about 100 years ago has not been detected since 1972, suggesting that it may have disappeared (Weaver *et al.*, 1997).

The phylogeography and evolution of BCRV in the western Great Plains could probably be understood better once we determine how these lineages differ from each other ecologically. In Nebraska where both occur, there is increasing evidence that the lineages may differ in their cytopathogenicity, in what sort of colony sites they are most likely to occur in, to what extent their genetic diversity at a site is affected by immigration of infected vectors to that site, when in the season they are most prevalent, and possibly in the extent to which they replicate in vertebrate hosts (Brown *et al.*, 2008; V. A. O'Brien & C. R. Brown, unpublished data). The phylogeographical and demographical patterns we document here may either reflect the emergence of two distinct strains of BCRV in the Great Plains, each adapted to different ecological niches, or portend the replacement of one by the other.

ACKNOWLEDGEMENTS

Plaque assays of the BCRV isolates were done in Nicholas Komar's laboratory at the Division of Vector-Borne Infectious Diseases, Centers for Disease Control and Prevention. For field and laboratory assistance, we thank Susan Beckett, Jillian Blackwell, Eric Edwards, Jennifer Klaus, Sarah Knutie, Matt Moore, Cheryl Ormston, Nicholas Panella, Sunita Quick, Sara Robinson, Rajni Sethi, Karen Winans and Gudrun Zöller. The School of Biological Sciences at the University of Nebraska-Lincoln allowed us to use the Cedar Point Biological Station, and the Union Pacific Railroad and the Robert Clary, Dave Knight and Loren Soper families allowed us access to land. D. B. Francy, Cluff Hopla, Gordon Smith, and W. D. Sudia collected and isolated the FMV and BCRV strains from the 1970s and 1980s. This work was supported by grants from the National Institutes of Health (R01-AI057569) and the National Science Foundation (IBN-9974733, DEB-0075199, DEB-0514824).

REFERENCES

Anonymous (2006). *Kansas Drought Report: 2006 Report No. 4.* Topeka, KS: Kansas Water Office.

Brault, A. C., Powers, A. M., Villarreal, C. L., Lopez, R. N., Cachon, M. F., Gutierrez, L. F. L., Kang, W., Tesh, R. B., Shope, R. E. & Weaver, S. C. (1999). Genetic and antigenic diversity among eastern equine encephalitis viruses from North, Central, and South America. *Am J Trop Med Hyg* **61**, 579–586.

Brault, A. C., Powers, A. M. & Weaver, S. C. (2002). Vector infection determinants of Venezuelan equine encephalitis virus reside within the E2 envelope glycoprotein. *J Virol* **76**, 6387–6392.

Brown, C. R. & Brown, M. B. (1995). Cliff swallow (*Hirundo pyrrhonota*). In *Birds of North America*, no. 149. Edited by A. Poole & F. Gill. Philadelphia, PA: Academy of Natural Sciences and Washington DC: The American Ornithologists' Union.

Brown, C. R. & Brown, M. B. (1996). *Coloniality in the Cliff Swallow: the Effect of Group Size on Social Behavior.* Chicago: University of Chicago Press.

Brown, C. R. & Brown, M. B. (2004). Empirical measurement of parasite transmission between groups in a colonial bird. *Ecology* **85**, 1619–1626.

Brown, C. R. & Brown, M. B. (2005). Between-group transmission dynamics of the swallow bug, *Oeciacus vicarius. J Vector Ecol* **30**, 137–143.

Brown, C. R., Komar, N., Quick, S. B., Sethi, R. A., Panella, N. A., Brown, M. B. & Pfeffer, M. (2001). Arbovirus infection increases with group size. *Proc R Soc Lond B* **268**, 1833–1840.

Brown, C. R., Brown, M. B., Moore, A. T. & Komar, N. (2007). Bird movement predicts Buggy Creek virus infection in insect vectors. *Vector Borne Zoonotic Dis* **7**, 304–314.

Brown, C. R., Brown, M. B., Padhi, A., Foster, J. E., Pfeffer, M., Moore, A. T. & Komar, N. (2008). Host and vector movement affects genetic diversity and spatial structure of Buggy Creek virus (Togaviridae). *Mol Ecol* **17**, 2164–2173.

Calisher, C. H., Monath, T. P., Muth, D. J., Lazuick, J. S., Trent, D. W., Francy, D. B., Kemp, G. E. & Chandler, F. W. (1980). Characterization of Fort Morgan virus, an alphavirus of the western equine encephalitis virus complex in an unusual ecosystem. *Am J Trop Med Hyg* **29**, 1428–1440.

Cilnis, M. J., Kang, W. & Weaver, S. C. (1996). Genetic conservation of Highlands J viruses. *Virology* **218**, 343–351.

Cooper, L. A. & Scott, T. W. (2001). Differential evolution of eastern equine encephalitis virus populations in response to host cell type. *Genetics* **157**, 1403–1412.

Drummond, A. J. & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* **7**, 214.

Drummond, A. J., Nicholls, G. K., Rodrigo, A. G. & Solomon, W. (2002). Estimating mutation parameters, population history, and genealogy simultaneously from temporally spaced sequence data. *Genetics* **161**, 1307–1320.

Drummond, A. J., Rambaut, A., Shapiro, B. & Pybus, O. G. (2005). Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol Biol Evol* **22**, 1185–1192.

Drummond, A. J., Ho, S. Y. W., Phillips, M. J. & Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. *PLoS Biol* **4**, e88.

Hayes, R. O., Francy, D. B., Lazuick, J. S., Smith, G. C. & Gibbs, E. P. J. (1977). Role of the cliff swallow bug (*Oeciacus vicarius*) in the natural cycle of a western equine encephalitis-related alphavirus. *J Med Entomol* **14**, 257–262.

Holmes, E. C. (2004). The phylogeography of human viruses. *Mol Ecol* **13**, 745–756.

Hopla, C. E., Francy, D. B., Calisher, C. H. & Lazuick, J. S. (1993). Relationship of cliff swallows, ectoparasites, and an alphavirus in west-central Oklahoma. *J Med Entomol* **30**, 267–272.

Huelsenbeck, J. P. & Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754–755.

Jenkins, G. M., Rambaut, A., Pybus, O. G. & Holmes, E. C. (2002). Rates of molecular evolution in RNA viruses: a quantitative phylogenetic analysis. *J Mol Evol* **54**, 156–165.

Kitchen, A., Miyamoto, M. M. & Mulligan, C. J. (2008). Utility of DNA viruses for studying human host history: case study of JC virus. *Mol Phylogenet Evol* **46**, 673–682.

Kondig, J. P., Turell, M. J., Lee, J. S., O'Guinn, M. L. & Wasieloski, L. P., Jr (2007). Genetic analysis of South American eastern equine encephalomyelitis viruses isolated from mosquitoes collected in the Amazon basin region of Peru. *Am J Trop Med Hyg* **76**, 408–416.

Kramer, L. D. & Fallah, H. M. (1999). Genetic variation among isolates of western equine encephalomyelitis virus from California. *Am J Trop Med Hyg* **60**, 708–713.

Lindsay, M. D. A., Coelen, R. J. & Mackenzie, J. S. (1993). Genetic heterogeneity among isolates of Ross River virus from different geographical regions. *J Virol* **67**, 3576–3585.

Lowther, P. E. & Cink, C. L. (1992). House sparrow. In *Birds of North America*, no. 12. Edited by A. Poole, P. Stettenheim & F. Gill. Philadelphia, PA: Academy of Natural Sciences and Washington, DC: The American Ornithologists' Union.

Moore, A. T., Edwards, E. A., Brown, M. B., Komar, N. & Brown, C. R. (2007). Ecological correlates of Buggy Creek virus infection in *Oeciacus vicarius*, southwestern Nebraska, 2004. *J Med Entomol* **44**, 42–49.

Moratorio, G., Costa-Mattioli, M., Piovani, R., Romero, H., Musto, H. & Cristina, J. (2007). Bayesian coalescent inference of hepatitis A virus populations: evolutionary rates and patterns. *J Gen Virol* **88**, 3039–3042.

Moya, A., Holmes, E. C. & Gonzalez-Candelas, F. (2004). The population genetics and evolutionary epidemiology of RNA viruses. *Nat Rev Microbiol* **2**, 279–288.

Nagata, L. P., Hu, W.-G., Parker, M., Chau, D., Rayner, G. A., Schmaltz, F. L. & Wong, J. P. (2006). Infectivity variation and genetic diversity among strains of western equine encephalitis virus. *J Gen Virol* **87**, 2353–2361.

Navaratnarajah, C. K. & Kuhn, R. J. (2007). Functional characterization of the Sindbis virus E2 glycoprotein by transposon linker-insertion mutagenesis. *Virology* **363**, 134–147.

Norder, H., Lundström, J. O., Kozuch, O. & Magnius, L. O. (1996). Genetic relatedness of Sindbis virus strains from Europe, the Middle East, and Africa. *Virology* **222**, 440–445.

Oberste, M. S., Weaver, S. C., Watts, D. M. & Smith, J. F. (1998). Identification and genetic analysis of Panama-genotype Venezuelan equine encephalitis virus subtype ID in Peru. *Am J Trop Med Hyg* **58**, 41–46.

Pfeffer, M., Foster, J. E., Edwards, E. A., Brown, M. B., Komar, N. & Brown, C. R. (2006). Phylogenetic analysis of Buggy Creek virus: evidence for multiple clades in the western Great Plains, United States of America. *Appl Environ Microbiol* **72**, 6886–6893.

Poidinger, M., Roy, S., Hall, R. A., Turley, P. J., Scherret, J. H., Lindsay, M. D., Broom, A. K. & Mackenzie, J. S. (1997). Genetic stability among temporally and geographically diverse isolates of Barmah Forest virus. *Am J Trop Med Hyg* **57**, 230–234.

Posada, D. (2004). *Collapse: Describing Haplotypes from Sequence Alignments*. Version 1.2. Vigo, Spain: University of Vigo. Available from http://darwin.uvigo.es/software/collapse.html

Posada, D. & Crandall, K. A. (1998). MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**, 817–818.

Powers, A. M., Brault, A. C., Tesh, R. B. & Weaver, S. C. (2000). Re-emergence of chikungunya and o'nyong-nyong viruses: evidence for distinct geographical lineages and distant evolutionary relationships. *J Gen Virol* **81**, 471–479.

Powers, A. M., Brault, A. C., Shirako, Y., Strauss, E. G., Kang, W., Strauss, J. H. & Weaver, S. C. (2001). Evolutionary relationships and systematics of the alphaviruses. *J Virol* **75**, 10118–10131.

Rambaut, A. & Drummond, A. J. (2007). *Tracer v1.4*. Available from http://beast.bio.ed.ac.uk/Tracer.

Rannala, B. H. (1995). *Demography and genetic structure in island populations*. PhD thesis. New Haven, Connecticut: Yale University.

Reisen, W. K. & Monath, T. P. (1989). Western equine encephalomyelitis. In *The Arboviruses: Epidemiology and Ecology*, vol. 5, pp. 89–137. Edited by T. P. Monath. Boca Raton, FL: CRC Press.

Romano, C. M., de Melo, F. L., Corsini, M. A. B., Holmes, E. C. & Zanotto, P. M. A. (2007). Demographic histories of ERV-K in humans, chimpanzees and rhesus monkeys. *PLOS One* **2**, e1026, doi:10.1371/journal.pone.0001026

Sammels, L. M., Lindsay, M. D., Poidinger, M., Coelen, R. J. & Mackenzie, J. S. (1999). Geographic distribution and evolution of Sindbis virus in Australia. *J Gen Virol* **80**, 739–748.

Sauchyn, D. J., Stroich, J. & Beriault, A. (2003). A paleoclimatic context for the drought of 1999–2001 in the northern Great Plains of North America. *Geogr J* **169**, 158–167.

Scott, T. W., Bowen, G. S. & Monath, T. P. (1984). A field study on the effects of Fort Morgan virus, an arbovirus transmitted by swallow bugs, on the reproductive success of cliff swallows and symbiotic house sparrows in Morgan County, Colorado, 1976. *Am J Trop Med Hyg* **33**, 981–991.

Smith, G. C. & Eads, R. B. (1978). Field observations on the cliff swallow, *Petrochelidon pyrrhonota* (Vieillot), and the swallow bug, *Oeciacus vicarius* Horvath. *J Wash Acad Sci* **68**, 23–26.

Strauss, J. H. & Strauss, E. G. (1994). The alphaviruses: gene expression, replication, and evolution. *Microbiol Rev* **58**, 491–562.

Strickler, S. A. (2006). *Winter ecology of a bird-associated arbovirus*. MSc thesis. Tulsa, OK: University of Tulsa.

Swofford, D. L. (2002). PAUP*: phylogenetic analysis using parsimony (*and other methods), version 4. Sunderland, MA: Sinauer Associates.

Tamura, K., Dudley, J., Nei, M. & Kumar, S. (2007). MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol* **24**, 1596–1599.

Weaver, S. C., Rico-Hesse, R. & Scott, T. W. (1992). Genetic diversity and slow rates of evolution in New World alphaviruses. *Curr Top Microbiol Immunol* **176**, 99–117.

Weaver, S. C., Bellew, L. A., Gousset, L., Repik, P. M., Scott, T. W. & Holland, J. J. (1993). Diversity within natural populations of eastern equine encephalomyelitis virus. *Virology* **195**, 700–709.

Weaver, S. C., Hagenbaugh, A., Bellew, L. A., Gousset, L., Mallampalli, V., Holland, J. J. & Scott, T. W. (1994). Evolution of alphaviruses in the eastern equine encephalomyelitis complex. *J Virol* **68**, 158–169.

Weaver, S. C., Kang, W., Shirako, Y., Rumenapf, T., Strauss, E. G. & Strauss, J. H. (1997). Recombinational history and molecular evolution of western equine encephalomyelitis complex alphaviruses. *J Virol* **71**, 613–623.

Weaver, S. C., Brault, A. C., Kang, W. & Holland, J. J. (1999). Genetic and fitness changes accompanying adaptation of an arbovirus to vertebrate and invertebrate cells. *J Virol* **73**, 4316–4326.